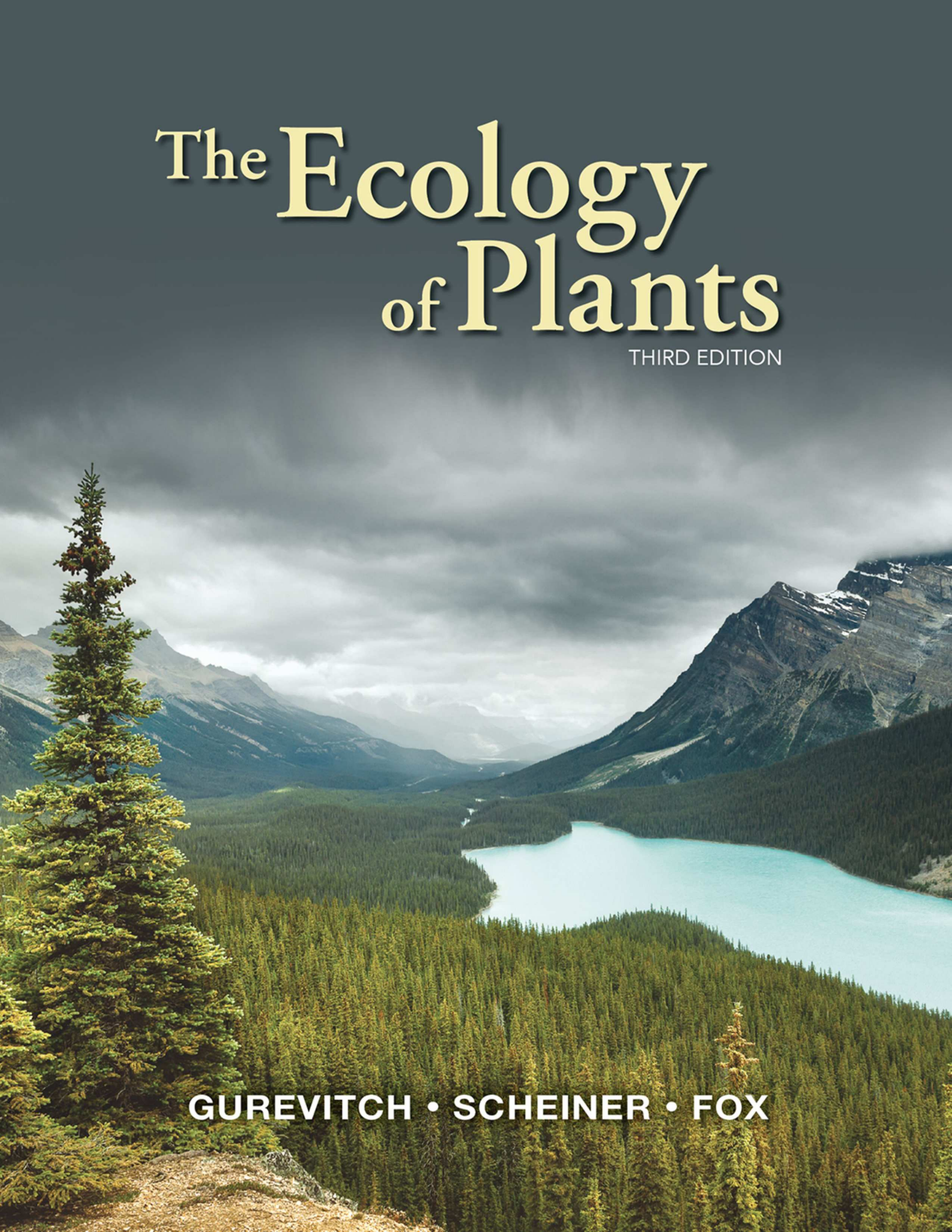
# The Ecology of Plants

THIRD EDITION



GUREVITCH • SCHEINER • FOX

# The Ecology of Plants

THIRD EDITION

# The Ecology of Plants

THIRD EDITION

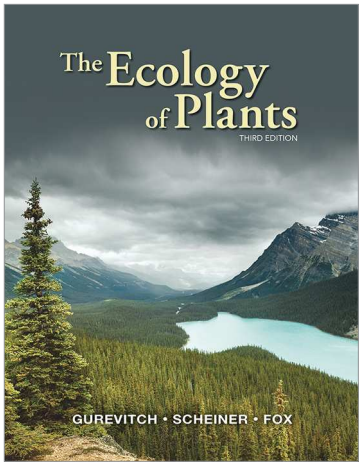
**Jessica Gurevitch**
*Stony Brook University*

**Samuel M. Scheiner**

**Gordon A. Fox**
*University of New Mexico*
*University of South Florida*

SINAUER ASSOCIATES

NEW YORK OXFORD
OXFORD UNIVERSITY PRESS




"Shades of Green"
© Anthony Robin/www.anthonyrobin.com

### About the Cover

Peyto Lake and surrounding slopes in Banff National Park, Alberta, Canada (image reversed). The trees in the foreground are *Picea engelmannii* (Engelmann spruce); the high elevation forests in the photo also include *Abies lasiocarpa* (subalpine fir), and likely *Pinus contorta* (Lodgepole pine) and *Pinus albicaulis* (whitebark pine), interspersed with alpine meadows. Treeline is clearly visible, as is a sliver of a glacier. The surreal color of the lake is due to pulverized silt or "rock flour" ground by glaciers, and washed into the lake.

### Back Cover

Top left: *Epilobium angustifolium* (fireweed, Onagraceae)

Top center: *Peonia suffruticosa* (Chinese tree peony, Paeoniaceae)

Top right: *Chondrosum gracile* or *Bouteloua gracilis* (blue grama, Poaceae)

Center left: *Crinum americanum* (swamp lily, Amaryllidaceae)

Center: *Banksia paludosa* (marsh or swamp banksia, Proteaceae)

Bottom left: *Gazania lichtensteinii* (yellow calendula or geelgousblom, Asteraceae)

Bottom center: *Leucospermum reflexum* (rocket pincushion, or perdekop-speldekussing, Proteaceae)

### Frontispiece

Spatial pattern in spring wildflowers at Namaqua National Park, South Africa. Photo courtesy of Gordon Fox.

*The Ecology of Plants*, Third Edition
Oxford University Press is a department of the University of Oxford. It furthers the University's objective of excellence in research, scholarship, and education by publishing worldwide. Oxford is a registered trade mark of Oxford University Press in the UK and certain other countries.

Published in the United States of America by Oxford University Press

198 Madison Avenue, New York, NY 10016, United States of America

© 2021, 2007, 2002 Oxford University Press

Sinauer Associates is an imprint of Oxford University Press.

For titles covered by Section 112 of the US Higher Education Opportunity Act, please visit www.oup.com/us/he for the latest information about pricing and alternate formats.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, without the prior permission in writing of Oxford University Press, or as expressly permitted by law, by license, or under terms agreed with the appropriate reproduction rights organization.

Inquiries concerning reproduction outside the scope of the above should be sent to the Rights Department, Oxford University Press, at the address above.

You must not circulate this work in any other form and you must impose this same condition on any acquirer.

Address editorial correspondence to:

Sinauer Associates
23 Plumtree Road
Sunderland, MA 01375 U.S.A.

Address orders, sales, license, permissions, and translation inquiries to:

Oxford University Press U.S.A.
2001 Evans Road
Cary, NC 27513 U.S.A.
Orders: 1-800-445-9714

Library of Congress Cataloging-in-Publication Data

Names: Gurevitch, Jessica, 1952– author. | Scheiner, Samuel M., 1956– author. | Fox, Gordon A., 1952– author.

Title: The ecology of plants / Jessica Gurevitch, Samuel M. Scheiner, Gordon A. Fox.

Description: Third edition. | New York : Sinauer Associates/Oxford University Press, [2021] | Includes bibliographical references and index.

Identifiers: LCCN 2020005892 (print) | LCCN 2020005893 (ebook) | ISBN 9781605358291 (paperback) | ISBN 9781605358307 (epub)

Subjects: LCSH: Plant ecology.

Classification: LCC QK901 .G96 2021 (print) | LCC QK901 (ebook) | DDC 581.7--dc23

LC record available at https://lccn.loc.gov/2020005892

LC ebook record available at https://lccn.loc.gov/2020005893

6 5 4 3 2 1
Printed in the United States of America

From the Authors:

*This book is dedicated to Andrew Sinauer,*
*who planted the seed and nurtured it into being.*

JG

*In memory of my parents, Esther and Louis Gurevitch,*
*and my teachers and students, for educating me.*

SMS

*To my mentors, Mike Wade, Jim Teeri, and Conrad Istock,*
*whose fingerprints are all over this book.*

GAF

*To Kathy, who has been with me for the whole journey.*

# Brief Contents


CHAPTER 1 The Science of Plant Ecology 1

**PART I Individuals and Their Environments 19**

CHAPTER 2 Photosynthesis and Light 21
CHAPTER 3 Water Relations and Thermal Energy Balance 53
CHAPTER 4 Soil and Terrestrial Plant Life 83
CHAPTER 5 Ecosystem Processes 111

**PART II From Individuals to Populations 141**

CHAPTER 6 Individual Growth and Reproduction 143
CHAPTER 7 Plant Life Histories 177
CHAPTER 8 Population Structure, Growth, and Decline 199
CHAPTER 9 Evolution: Processes and Change 231

**PART III Population Interactions and Communities 259**

CHAPTER 10 Competition and Other Plant Interactions 261
CHAPTER 11 Herbivory and Other Trophic Interactions 297
CHAPTER 12 Community Diversity and Structure 333
CHAPTER 13 Community Dynamics and Succession 371
CHAPTER 14 Local Abundance, Diversity, and Rarity 397

**PART IV From Landscapes to Planet Earth 419**

CHAPTER 15 Landscapes: Pattern and Scale 421
CHAPTER 16 Climate, Plants, and Climate Change 447
CHAPTER 17 Paleoecology 495
CHAPTER 18 Biomes and Physiognomy 513
CHAPTER 19 Global Biodiversity Patterns, Loss, and Conservation 543

# Contents


## 1 The Science of Plant Ecology 1

1.1 Ecology Is a Science 2

Where does scientific knowledge come from? 2

Scientific research involves objectivity, subjectivity, choice, and chance 5

Observational studies detect and quantify patterns 5

Experiments are central to research 5

In ecology, "controls" are what you are using for baseline comparisons 8

How do we test theories? 10

Studies can lead to specific results but contribute to general understanding 12

Science is ultimately consistent, but getting to consistency is a challenge 12

1.2 Ecological Phenomena Are Heterogeneous in Many Ways 12

1.3 Plant Ecology Has Developed through the Interaction of Observation, Measurement, Analysis, Technology, and Theory 14

Plant ecology is situated in the more general theoretical framework of ecology 16

Ecology has a range of subdisciplines 17

Science is a human endeavor 18

# PART I Individuals and Their Environments 19

## 2 Photosynthesis and Light 21

2.1 Photosynthesis Is the Engine of Life on Earth 22

BOX 2A The Discovery and Elucidation of Photosynthetic Carbon Reduction 26

2.2 Photosynthesis Is Affected by the Environment and by Plant Adaptations 28

The amount of light available limits photosynthesis 28

Carbon uptake is limited by the ways plants respond to their environments 31

Photosynthetic rates can vary among species in different habitats 32

2.3 There Are Three Photosynthetic Pathways: $C_3$, $C_4$, and CAM 33

$C_3$ photosynthesis is the most common and original type of photosynthesis 33

BOX 2B Photorespiration 34

$C_4$ photosynthesis is a specialized adaptation for rapid carbon uptake in warm, bright environments 35

BOX 2C Stable Isotopes and Photosynthesis 37

Crassulacean acid metabolism (CAM photosynthesis) is a specialized adaptation for minimizing water loss but at the cost of reduced photosynthesis and slow growth 38

2.4 $C_3$ Photosynthesis Is the Foundation for the Evolution of $C_4$ and CAM 39

$C_4$ and CAM evolved from $C_3$ photosynthesis many different times in many different plant families 39

Photosynthesis first evolved about 2.5 billion years ago and has continued to evolve over Earth's history 39

2.5 $C_3$, $C_4$, and CAM Plants Each Have Distinct Growth Forms, Phenology, and Distributions 42

The three photosynthetic types dominate in different habitats and differ in growth form 42

$C_3$ and $C_4$ plants grow most actively in different seasons 43

$C_3$, $C_4$, and CAM species have different geographic distributions 44

2.6 Plants Possess Many Different Adaptations to Their Light Environments 46

Many plants can detect the length of daylight and how it is changing seasonally 46

Leaves grown in sunlit and shaded conditions can differ in structure and function 47

BOX 2D Blue Color and Iridescence, Structural Coloration, and Anthocyanin Pigments 50

## 3 Water Relations and Thermal Energy Balance 53

The ancestors of modern plants evolved to live in terrestrial environments 54

3.1 Water Potential Provides a Framework for Understanding How Plants Interact with Water in Their Environment 55

BOX 3A Measuring Photosynthesis, Transpiration, and Water Potential 56

3.2 Water Moves through a Soil-Plant-Atmosphere Continuum 57

3.3 Plants Manage Transpiration and Water Loss 59

Plants have different strategies for adapting to water availability 60

Water use efficiency is a measure of carbon gain versus water loss 62

Plants have different adaptations for coping with reduced water availability 62

Plants have complex physiological adaptations to drought 65

The anatomy and physiology of stomata shape plant responses to water loss 67

Leaf anatomy can be adaptive for survival and growth in arid environments 68

Roots, stems, and their tissues have adaptations for controlling plant water relations 71

3.4 Everything in the Universe Has a Thermal Energy Balance 75

Radiant energy is always being exchanged between plants and their surroundings 76

BOX 3B Why the Sky Is Blue and the Setting Sun Is Red 77

Energy flows between plants and air, water and soil via conduction and convection 77

Water loss is accompanied by latent heat loss 78

Putting it all together: what determines leaf and whole-plant temperature? 78

Plants may have adaptations to extreme temperature regimes 80

4 Soil and Terrestrial Plant Life 83

4.1 Soils Have Distinct and Varied Composition, Characteristics, and Structure 84

It takes many thousands of years to create soil 85

BOX 4A Serpentine Soils 85

Soil texture determines many of the properties of soils that affect plants 89

Soil pH has profound but indirect effects 92

Soils are characterized by horizons—layers with distinctive properties 94

BOX 4B Soil Conservation Is a Major Global Environmental Issue 96

Soils are the unique product of living organisms acting on soil parent material 97

4.2 The Rhizosphere Is a Unique Environment Created by Roots and Their Interaction with Microorganisms 98

4.3 Water Moves through the Soil to Reach Plants 99

4.4 The Basic Building Blocks of Plants are C, H, and O from Air and Water, and Macronutrients and Micronutrients from the Soil 101

The stoichiometry of elements in plants and soils regulates many ecological processes 104

Nitrogen is often the limiting nutrient for plant growth 104

In some plants nitrogen comes from fixation by symbiotes 105

BOX 4C Symbioses and Mutualisms 105

Phosphorus is limiting for plant growth in many environments 108

5 Ecosystem Processes 111

5.1 Ecosystem Processes Set the Stage for Life in a Salt Marsh 112

5.2 Ecosystem Pools and Fluxes Form Cycles of Nutrients and Energy 114

BOX 5A Biogeochemical Cycles: Quantifying Pools and Fluxes 114

5.3 Carbon Is the Foundation of Life on Earth 116

Productivity measures how carbon moves between living things and their nonliving environment 116

Carbon is stored in the living and nonliving components of ecosystems 119

Soil food webs are the recycling engine of terrestrial ecosystems 122

Soil organic matter revisited: Bacteria are an essential component of soil organic matter 125

Primary productivity can be measured or estimated in a variety of ways 125

BOX 5B Using Remote Sensing and Eddy Covariance Methods to Estimate NPP 126

5.4 The Nitrogen Cycle Is an Essential Component of Ecosystems 127

Bacteria mineralize organic N to inorganic forms taken up by plants 129

Nitrogen is lost from ecosystems through denitrification and leaching 131

Decomposition can immobilize soil nitrogen when $NO_3^-$ or $NH_4^+$ are sequestered in bacterial biomass 132

5.5 Nitrogen Deposition and Acid Precipitation Can Alter Ecosystems 132

BOX 5C The Haber Process, the Green Revolution, and Nitrogen in Ecosystems 134

5.6 Cycles of Phosphorus and Other Elements Play Important Roles in Ecosystems 135

Microorganisms make phosphorus available for plants 135

Sulfur is critical for certain plant compounds 136

Calcium is necessary for many plant processes and structures 136

5.7 The Water Cycle Is Central to Life and Climate 137

Water cycles at local and regional scales 137

Local water cycles can affect global cycles 139

# PART II From Individuals to Populations 141

## 6 Individual Growth and Reproduction 143

6.1 Growth Begins with Seed Germination 144

6.2 Plants Grow by Adding Repeated Units to Their Bodies 144

6.3 Plant Growth Affects Resource Acquisition 146

Shoot architecture determines light interception 146

The growth of clonal plants affects their ability to take up patchy resources 147

6.4 Plants Reproduce both Sexually and Asexually 149

Many plants reproduce vegetatively 149

Some plants produce seeds asexually 150

The sexual life cycles of plants involve alternation of generations 150

6.5 The Movement of Pollen Is an Important Aspect of a Plant's Life Cycle 153

The pollen of many plants is moved by the wind 153

Visual displays are important for attracting animal visitors 154

Animal visitors are attracted to plants with floral odors or acoustic guides 156

Plants often need to limit unwanted visits 158

How strongly are floral characteristics associated with particular pollinators? 158

BOX 6A Specialized Plants and Pollinators 159

BOX 6B Some Complex Plant-Pollinator Interactions 160

Aquatic plants have special adaptations for pollination 161

BOX 6C Is There a Pollinator Crisis? 161

6.6 Plants Have Complicated Mating Systems 162

Inbreeding is mating between close relatives 162

Plants may vary in gender 163

BOX 6D Pollination Experiments 164

Competition occurs among plants and among pollen grains 165

Most mating is between neighboring individuals 166

Plants may mate preferentially with individuals with similar phenotypes 166

Fitness can depend on a population's composition 167

Mating systems have other important consequences 168

6.7 Fruit and Seed Characteristics Affect Dispersal across Space and Time 170

The structures of seeds and fruits affect their dispersal 170

Plants can disperse across time via seed banks 174

## 7 Plant Life Histories 177

7.1 Trade-Offs Are a Central Cause of Variation in Life History Patterns 178

Trade-offs are difficult to measure 178

An important trade-off is in the size and number of seeds 179

7.2 Evolution Acts on the Schedule of Survival and Reproduction 181

How long a plant lives and when it does its growing is part of its life history strategy 182

7.3 Several Theories Address Life History Strategies 184

Demographic life history theory is based on evolutionary principles 184

*r*- and *K*-selection theory was influential in earlier thinking about life histories 184

*r*- and *K*-strategy theory was extended to the ecology of plants 185

Grime's triangular model focuses on the ecological conditions favoring different life history strategies 185

Demographic life history theory has been tied to patterns of reproductive allocation 186

Other theories of life history strategies are based on examining plant traits 188

7.4 Year-to-Year Variation in the Environment Shapes Life Histories 189

Among-year demographic variation reduces fitness 189

Bet-hedging strategies can reduce the variance in fitness 190

Seed germination is triggered by many factors 191

Masting can result in synchronization of flowering among individuals 192

7.5 Phenology Is the Within-Year Schedule of Growth and Reproduction 193

The timing of growth is driven by both abiotic and biotic factors 194

The timing of reproduction may be due to abiotic factors 195

The timing of reproduction may be due to biotic factors 196

## 8 Population Structure, Growth, and Decline 199

8.1 Plant Biology Creates Special Issues for Population Studies 200

BOX 8A Genets and Ramets: What Is an Individual? 201

8.2 Plant Populations Are Structured by Age, Size, and Developmental Stage 202

Plant population structure is complicated because plants can change size or form at variable rates 203

8.3 Studying Population Growth Usually Involves Models of Changes in Population Structure 205

Life cycle graphs are useful models of plant demography and its relationship to data acquisition 206

Estimating vital rates can be done several ways 207

BOX 8B How to Construct a Life Table 208

BOX 8C Borrowing the Mark-Recapture Method from Animal Ecology 209

BOX 8D Obtaining Data for Survival Studies 210

There are several approaches to building models for structured populations 211

BOX 8E Constructing Matrix Models 212

Analyzing demographic models gives information on population growth rates and population composition 213

BOX 8F Demography of an Endangered Cactus 213

BOX 8G Multiplying a Population Vector by a Matrix 214

Measuring lifetime reproduction gives us the net reproductive rate of the population 215

Reproductive value is the contribution of each stage to population growth 215

BOX 8H Reproductive Value 216

Sensitivity and elasticity indicate how individual matrix elements affect population growth 217

Life table response experiments can examine the demographic differences among populations 218

BOX 8I How Do Changes in Transition Probabilities Affect the Population Growth Rate? 219

Ecologists are beginning to study demography at larger spatial scales 219

There are additional approaches to modeling plant demography 220

Plant populations are heterogeneous 220

8.4 Demographic Studies of Long-Lived Plants Require Creative Methods 221

8.5 Population Growth Fluctuates Randomly over Time 222

There are two general types of random variation 223

Random fluctuations reduce long-term growth rates 225

Studying variable population growth requires data recorded over many years 227

8.6 Demographic Models Have Strengths and Limitations 228

## 9 Evolution: Processes and Change 231

9.1 Natural Selection Is a Primary Cause of Evolutionary Change 232

Variation in phenotype is necessary for natural selection 233

Three conditions are necessary for evolution by natural selection 234

9.2 Heritability Measures the Genetic Basis of Phenotypic Variation 235

Heritability is a measure of resemblances among relatives 235

BOX 9A A Simple Genetic System and the Resemblance of Relatives 237

BOX 9B Using Genes to Track Pollen and Seeds and to Identify Species 238

Phenotypic variation can be partitioned into genetic and nongenetic components 238

The environment can interact with the genome to determine the phenotype 239

Genotypes are often nonrandomly distributed among environments 240

9.3 Patterns of Adaptation Are the Result of Natural Selection 240

Heavy-metal tolerance is an example of genetic differentiation 241

Adaptation to different light conditions is an example of adaptive plasticity 243

Environmental effects can extend over generations 245

Phenotypic plasticity is important for understanding other ecological concepts 245

9.4 Natural Selection Can Occur at Levels Other Than the Individual 246

9.5 Other Processes Can Cause Evolutionary Change 247

Mutation, migration, and sexual reproduction are processes that increase genetic variation 248

Genetic drift is a process that decreases genetic variation 248

These evolutionary processes have important conservation implications 250

9.6 Evolutionary Processes Can Affect Variation among Populations 250
9.7 Ecotypes Are Different Forms of a Species That Are Adapted to Different Environments 250
9.8 Natural Selection Can Cause the Origin of New Species 254
9.9 Adaptation and Speciation Can Happen through Hybridization 256

PART III Population Interactions and Communities 259

10 Competition and Other Plant Interactions 261

10.1 Individuals Compete for Limited Resources 262
What are the mechanisms of resource competition? 263
Resource competition often depends on plant size 266
Plant competition frequently occurs between seedlings 266
Seedling competition can lead to self-thinning 269
10.2 There Are Several Approaches to Experiments for Studying Competition 270
How we quantify competition can affect experimental results 270
Competition experiments were originally conducted in greenhouse or garden environments 271
10.3 Interactions among Species Range from Competition to Facilitation 273
Different theories attempt to explain how competitive trade-offs lead to strategies 274
Are there fixed competitive hierarchies? 275
BOX 10A Plant Traits and the Worldwide Leaf Economic Spectrum: Attempts to Simplify Understanding of Plant Diversity 276
Does allelopathy between species explain patterns in nature? 276
Plants can change the environment to the advantage of other plants 279
Competitive exclusion sometimes determines species distributions 281
10.4 Competition and Facilitation May Vary along Environmental Gradients 282
There are conflicting models of how productivity affects the importance of competition and facilitation 282
Experimental evidence provides a mixed picture about the roles of competition and facilitation along productivity gradients 284
Research syntheses provide some help in interpreting the evidence 286
Can we resolve the conflicting results? 287
BOX 10B Research Synthesis, Systematic Reviews, and Meta-Analysis: Tools for Summarizing Results across Studies 289
Models of plant competition can help us to better understand competitive processes and the role of competition in species coexistence 289
Modern coexistence theory is a framework for understanding how competition affects coexistence 290
Models within the framework of modern coexistence theory have stimulated research and discovery 291
New research can extend our understanding of coexistence 293

11 Herbivory and Other Trophic Interactions 297

11.1 The Effects of Herbivores on Individual Plants Depend on What Is Eaten 298
11.2 Herbivores Can Alter Plant Population Composition and Dynamics 300
Herbivores can change where plants grow 302
Herbivory on seeds has both negative and positive consequences for plant populations 303
People use insect herbivores for biological control 303
11.3 Herbivores Affect Plant Communities in Different Ways 305
Herbivore behavior can change plant community composition 305
Herbivory might result in apparent competition among plants 308
Domesticated and introduced herbivores can shape plant communities 308
How important is herbivory in shaping the natural world? 310
11.4 Plants Defend Themselves against Herbivores by Different Means 310
Plants use a variety of physical defenses to protect themselves 310
Plants have evolved a wide range of chemical defenses against herbivores 312
Plant chemical defenses can be constant or be induced by herbivory 314
Evolutionary consequences of plant-herbivore interactions 316
11.5 Plants Are Involved in Many Kinds of Trophic Interactions 317
Some plants are parasites of other plants 317

11.6 Plants Interact with Pathogens, Endophytes, and Mycorrhizae in Complex Ways 318

BOX 11A "Broken" Tulips and the Tulip Mania of the 1600s 319

Plants are attacked by many different disease-causing organisms 319

BOX 11B Effects of Plant Disease on Humans: Potato Blight and the Irish Potato Famine (the Great Famine) 319

Plants have immediate defenses and long-term evolutionary responses to pathogens 321

Pathogens can shape plant populations and communities 322

Plant pathogens can interact in complex ways with other organisms 324

Endophytes are symbiotic organisms that live inside plant cells 324

Mycorrhizae are essential for terrestrial life 325

Arbuscular mycorrhizae and ectomycorrhizae are the two most ecologically important groups 326

Specialized mycorrhizal interactions include those associated with the Ericaceae and Orchidaceae 328

Mycorrhizae function in other ways in addition to nutrient uptake 329

Are mycorrhizal fungi mutualists or parasites? 329

The influence of mycorrhizae can depend on plant-plant interactions as well as on soil nutrients 330

12 Community Diversity and Structure 333

12.1 There Are Many Ways of Thinking about Communities 334

BOX 12A Communities, Taxa, Guilds, and Functional Groups 335

The debate between Henry Gleason and Frederic Clements shaped modern ideas about plant communities 336

Today's ecologists have a different perspective on the issues in contention 338

BOX 12B A Deeper Look at Some Definitions: Abiotic Factors and Emergent Properties 340

The concept of communities is useful but has often been debated 340

12.2 Biodiversity Describes Variation in Biological Organisms and Systems 341

Biodiversity metrics can be built from different types of information 342

Inventory diversity is the variation of types of objects 342

BOX 12C A Unified Measure of Diversity 344

Differentiation diversity is the variation among units 345

Phylogenetic diversity is variation in evolutionary relationships 345

Functional diversity is variation in traits 347

Different types of biodiversity information can be combined 349

12.3 Communities Can Be Measured in Many Ways 349

Measuring species richness can involve simple sampling procedures or complex mathematical estimates 349

There are many ways to sample communities 354

One measure of a plant community is its physiognomy 357

Long-term studies are important for measuring communities 357

BOX 12D The Long-Term Ecological Research Network 358

12.4 Plant Communities Can Be Compared by Many Methods 358

Non-numerical techniques were the first methods for comparing communities 359

Communities can be compared by single factors using univariate techniques 360

Most community comparisons use multivariate techniques 360

12.5 Communities Are Distributed across Landscapes 362

Ordination is a group of techniques for describing landscape patterns 362

Patterns of species difference among communities are caused by variation in the environment 364

What types of data are used? 365

Classification is an alternative approach to describing communities in a landscape 366

13 Community Dynamics and Succession 371

13.1 Conflicting Theories Have Attempted to Explain the Mechanisms of Succession 372

Are communities dynamic mosaics or regulated by predictable processes? 372

BOX 13A History of the Development of Modern Succession Theory 373

Scientific understanding can be influenced by methodology 374

13.2 Successional Change Has Three General Causes 377

Disturbance size affects which species can colonize 377

Fire can cause disturbance 379

Wind can cause disturbance 381

Water can cause disturbance 381

Animals can cause disturbance 382

Earthquakes and volcanoes can cause disturbance 382

BOX 13B The Dust Bowl of the 1930s 383

Disease can cause disturbance 384
Humans can cause disturbance 384

13.3 Which Species Are Available for Colonization Affects Succession 385
The dispersal capacity of species affects their colonization capability 385
Species can emerge from the propagule pool 387

13.4 Species Performance Determines the Pattern of Successional Change 390
Species vary in their life histories 390
Species interactions are central to species replacement during succession 391
Resource availability can change during succession 392

13.5 The Pathway of Succession Can Vary 393
Succession may or may not be predictable 393
Understanding successional processes is critical for community restoration 394

13.6 Ecologists Have Reconsidered the Concept of Climax 395

14 Local Abundance, Diversity, and Rarity 397

14.1 Are Dominant Species Competitively Superior? 398
There are many ways to be rare but few ways to be common 398
Being rare can vary over space and time 399
What makes a species common or rare? 402

14.2 Biological Invasions Are a Worldwide Concern 403
Why do some species become invasive? 404
What makes a community susceptible to invasion? 405
Efforts have been made to integrate explanations for invasiveness and susceptibility to invasion 409
Invasive species may alter many community properties and threaten biodiversity 410

14.3 Species Richness and Abundances Differ Greatly among Communities 411
Abundance curves illustrate community structure graphically 411
Productivity and diversity are related in complex ways within communities 412
Trade-offs and specialization contribute to diversity in heterogeneous environments 414
Disturbances might maintain community diversity 415

14.4 Does Increased Diversity Enhance Community Productivity or Stability? 416
Community dominance and diversity can affect ecosystem processes 417
Diversity has been hypothesized to increase stability 417
Diversity, rarity, and commonness vary with spatial extent 417

PART IV From Landscapes to Planet Earth 419

15 Landscapes: Pattern and Scale 421

15.1 Understanding Scale Is Critical to Understanding Ecological Processes 422
Patterns and processes can vary with scale 422
Scale interacts with environmental heterogeneity 424
Processes and patterns may vary as grain and extent change 425
Spatial pattern and scale can be analyzed using graphical and statistical methods 426

15.2 Landscape Ecology Involves Measuring Spatial Patterns and Looking at Their Effects 427
Defining patches is a key step in measuring patterns 427

BOX 15A Differentiating Vegetation Based on Spectral Quality 428
Patches can be quantified by their sizes, shapes, and spatial arrangement 429
Spatial patterns determine many ecological processes 430
How one analyzes landscape data affects whether the landscape appears to be continuous or discrete 430

15.3 Ecological Processes Occur across Landscapes 431
Island biogeography theory 431
Ecologists have debated whether there is a set of rules that determines how communities are put together 433
Metapopulation theory 434

BOX 15B Metapopulation Models 435
Demographic processes occur across landscapes 436
Metacommunity theory 437

15.4 Ecological Processes at the Level of Landscapes Is Important for Plant Conservation 439
Fragmentation of landscapes is a major threat to biodiversity 439
Key landscape characteristics are edges, connectivity, and nestedness 443
Ecological theory can help guide reserve design 445

## 16 Climate, Plants, and Climate Change 447

16.1 There Are Important Differences between Climate and Weather 448

16.2 The Kinetic Energy of Molecules Determines Heat and Temperature 448

The sun's angle is the main factor determining the radiant energy received at Earth's surface 451

There are long-term cycles in Earth's path around the sun that affect radiant energy at Earth's surface 455

16.3 Precipitation Patterns Vary across the Earth 457

Global patterns are determined by air moving in three dimensions at huge spatial scales 457

BOX 16A The Coriolis Effect 460

Continental-scale movement of air and water explain regional differences in snow and rain 464

Seasonal variation in precipitation is an important component of climate 465

The El Niño Southern Oscillation affects rainfall at large spatial scales and intermediate time scales 468

Temperature and rainfall predictability affect plant ecology and evolution 470

16.4 Anthropogenic Global Climate Change Is Caused by Humans and Is Affecting Vegetation 471

The global carbon cycle is central to Earth's climates 472

Increasing atmospheric $CO_2$ has direct effects on plants 474

The greenhouse effect warms the Earth due to greenhouse gases 475

BOX 16B The Ozone Hole and the Greenhouse Effect 476

16.5 Humans Are Changing the Global Carbon Cycle 477

Fossil fuel combustion is the most important factor changing the greenhouse effect 477

Deforestation and land use change also affect climate 481

16.6 Agriculture Is a Major Source of Greenhouse Gases 482

16.7 Global Climate Change Is Already Occurring 483

16.8 Large Changes Are Predicted for Earth's Climates, but Some Impacts Can Still Be Mitigated 486

BOX 16C Understanding Past Climates and Predicting Future Climates 486

16.9 Changing Climates Are Affecting Species and Ecological Systems 489

16.10 Responses to Ongoing and Predicted Climate Change 492

## 17 Paleoecology 495

17.1 Plants Invaded the Land in the Paleozoic Era 496

17.2 The Mesozoic Era Was Dominated by Gymnosperms and Saw the Origin of the Angiosperms 499

Gymnosperms were the first group of dominant seed plants 499

The breakup of Pangaea happened as the angiosperms rose to dominance 501

The boundary between the Cretaceous and Tertiary periods resulted in big changes in the flora and fauna 503

17.3 The Cenozoic Era Was Dominated by Angiosperms 503

17.4 Many Different Methods Are Used to Uncover the Past 504

17.5 Vegetation Change in the Recent Past Has Been Dominated by the Waxing and Waning of Glaciers 505

At the glacial maximum, climates and habitats were very different from today 506

Modern plant communities began to appear as the glaciers retreated 508

Climatic fluctuations of the recent past continue to shape the vegetation 510

## 18 Biomes and Physiognomy 513

18.1 Vegetation Can Be Categorized by Its Structure and Function 514

Plant physiognomy varies across the globe 514

Forests are closed canopy systems dominated by trees 516

Tree line defines the edge between treed and treeless landscapes 518

Grasslands and woodlands dominate in areas of lower precipitation 518

Shrublands and deserts are found in very dry or cool regions 519

18.2 Biomes with Similar Vegetation Forms May Be the Result of Convergent Evolution 520

18.3 Moist Tropical Forests 523

Tropical rainforest 523

Tropical montane forest 526

18.4 Seasonal Tropical Forests and Woodlands 526

Tropical deciduous forest 526

Thorn forest 527

Tropical woodland 527

18.5 Temperate Deciduous Forest 528

18.6 Other Temperate Forests and Woodlands 529

Temperate rainforest 529
Temperate evergreen forest 530
Temperate woodland 531
**18.7** Taiga 532
**18.8** Temperate Shrubland 533
**18.9** Grasslands 534
Temperate grassland 534
Tropical savanna 536
**18.10** Deserts 537
Hot desert 537
Cold desert 538
**18.11** Alpine and Arctic Vegetation 539
Alpine grassland and shrubland 539
Tundra 540

## 19 Global Biodiversity Patterns, Loss, and Conservation 543

**19.1** Biodiversity Varies Enormously across the Earth 544
Global biodiversity increases toward the tropics 545
**19.2** What Explains Global Biodiversity Patterns? 546
Explanations for the latitudinal diversity gradient include energy, water, and environmental heterogeneity, but all explanations have limitations 546
**BOX 19A** The Fynbos and the Cape Region of Africa Have Some of the World's Highest Plant Diversity 547
There are also regional and global patterns of β-diversity 550
**19.3** There Are Distinctive Regional and Continental Patterns of Plant Biodiversity 550
Continents at the same latitudes differ in species diversity 551
Transition zones may have higher diversity due to overlaps in species' distributions 553
Mountains and mountainous regions have distinct but complex patterns of species diversity 554
**19.4** Regional Diversity and Local Diversity Can Influence One Another 556
Endemism, isolation, and global biodiversity hotspots 557
**19.5** Patterns of Species Diversity May Be Explained in General Terms 561
Null models and the neutral theory of biodiversity and biogeography pose a different approach to explaining patterns of species diversity 562
Other explanations have been posed to explain variation in biodiversity, but patterns are scale dependent 562
**BOX 19B** Explaining Diversity along Ecological Gradients 564
**19.6** Biodiversity Is Rapidly Being Lost Globally 566
What is being lost? 566
Biodiversity is threatened by human activity 567
Does human domination require a new definition of the biomes? 570
Both rare and common species face threats in a range of communities 570
Human population growth and land use contribute to biodiversity loss 570
**19.7** Ecosystem Services Are One Way of Quantifying the Benefits of Natural Systems to Humans 572
Why should anyone care about plant biodiversity? 572
Conservation and restoration of biodiversity: a ray of hope? 573

## Glossary G-1

## Index I-1

# Preface

This book grew out of an informal chat at a conference. A long time ago (1994) in what seems like a galaxy far away, at the joint meeting of the Society for the Study of Evolution and American Society of Naturalists at the University of Georgia, one of us (GAF) was browsing the books on display by Sinauer Associates. Andy Sinauer struck up a conversation that at some point included the question "Who might be a good person to write a textbook on plant ecology?" GAF's immediate answer was something like, "I don't know," while he thought, "Please, not me!" Later that same day, in another of the casual chats that occur at academic conferences, GAF happened to mention this conversation to SMS, who expressed a similar feeling. A while later, as the two walked along, they returned to the topic, and one of them said something like "Well, we could do it if we involved someone else; how about JG?" Eventually the three of us went to Andy and told him that maybe we were interested after all. Andy encouraged us, but he also made clear what we'd have to do: write him a proposal convincing him that we knew what we were doing, and then actually go and write the book. We looked at one another with a bit of dismay and trepidation (well-founded, as it turned out).

Eventually we created the first edition. Along the way, some colleagues told us we were insane to undertake the project (too much time and effort, too little in the way of professional or financial reward, they cautioned). Many other colleagues and friends encouraged us. Our goal was to provide a comprehensive, readable textbook for an upper-level course in plant ecology, emphasizing a conceptual approach to the subject and an evolutionary focus. Evolutionary biology is essential to how we as scientists think about ecology, and we incorporated an evolutionary perspective throughout the book, as well as including a short introduction to the subject. We think we did that again and brought everything up to date, fixed some errors, added color illustrations, and published our second edition in 2006. The book brought us into contact with many students, instructors, and scientists we would not have had the opportunity to engage with otherwise, and we are grateful for that. Eventually it became apparent that we really were way overdue for bringing the book up to date, so 14 years later we are pleased to present a third edition.

Books don't write (or revise) themselves, and this book is certainly a collaborative effort. The order of authors' names must necessarily be printed in a linear fashion, and in most cases this implies the order of their contributions. In the case of this book, a circle would be more appropriate. The three authors of this book all contributed in multiple overlapping ways to the book; our contributions were different, but not greater or lesser. This book could not have been written by any one or two of us, and it very strongly reflects all our contributions and differing perspectives. It also offers a taste of our various senses of humor, and we hope that it provides at least a few chuckles to students and instructors wading through this sometimes intense and rather dense compilation of information and ideas.

We have been delighted with the response to the second edition of this book and have received many positive and useful comments from both students using the book and professors who have adopted it for their courses. The third edition is different in a number of respects. Most notably, 14 years have elapsed since the previous edition, and much has been learned in plant ecology as well as in ecology more broadly. We have endeavored to include new developments and new ideas, as well as new evaluations of older work. We hope that this edition will be useful in helping young ecologists make their way through the enormous literature of plant ecology and that we are effective in sharing our continued excitement about the discipline and our love for the natural world. We have added many new illustrations and photos and have updated and redrawn many others. In addition to including work published since the second edition, we have also reorganized and consolidated material and have developed certain sections to include a fuller treatment; for example, the material on how to think about and quantify diversity has been updated and consolidated into a single chapter; the explanations of Earth's climate and climate change have been integrated, sharpened, corrected, and we hope made clearer. Other, less central material has been deleted or shortened.

We assume that students using this book will have had an introductory course in biology, but they may or may not have had advanced biology courses and perhaps have not taken a course in general ecology. Recognizing that plant ecology may be the only ecology course a student will take, we have broadly covered the field of ecology, from individual plants through populations and communities, to large scale patterns and global issues. Thus, we strive to be comprehensive, albeit from a uniquely plant perspective. While topics are introduced at a basic level, there is sufficient depth, coverage, and leads to further references and information on the topics for more advanced students as well.

Plant ecology touches and builds on many subject areas that may not be covered in a typical introductory biology course. Therefore, we include background information that might be considered beyond the subject of plant ecology in its strictest sense. For example, we introduce aspects of plant anatomy and physiology, integrating the information on these subjects when we address herbivory and ecosystem ecology. We include common names, family affinities, and photos or drawings to make species more familiar to students. We discuss soils and belowground interactions, paleoecology, evolution, climate, and nutrient cycling in greater depth than might ordinarily be expected in an ecology text, and we address global climate change from the perspective of both the roles and responses of plants and those of people. Every college textbook is a reflection not only of the subject but of what the authors think is important and interesting, and this one is unabashedly so.

Ecology can be taught in many different sequences: it is conceptually a "hypertext" subject rather than a strictly linear one in which one topic clearly builds on the other and leads to the next one. For example, one can begin with ecophysiology of individuals and proceed to the global ecosystem; but the reverse order is equally valid. While we present the topics in a fairly conventional order starting from individuals and moving to global ecology, we recognize that other orders are equally logical and that different instructors cover the topics in a different order. In the classic film *The Wizard of Oz*, Dorothy reaches a crossroad and wonders aloud which way to go. The Scarecrow (who is still mounted on a post) points one way and says, "That way is a very nice way." Then he adds, pointing in the opposite direction, "It's pleasant down that way too." And so it is in ecology, including plant ecology. To facilitate those different approaches, we provide abundant cross-references for topics introduced or covered in other chapters. This book should be usable, therefore, in courses that begin with biomes, for instance, rather than with the ecology of individual plants.

Science has a language of its own. Acquiring that language can sometimes be daunting. Throughout the book we have placed words that may be unfamiliar in bold, and we have defined them in the text and in the Glossary. Scientific terminology may be tedious to learn, but it performs a necessary function: providing a concise and precise vocabulary that facilitates clarity and communication. In some cases, though, these definitions are presented not because we approve of the proliferation of jargon in ecology, but because these terms are commonly used, and students need to be familiar with them to understand the scientific literature.

Throughout the book we have provided an entry to the scientific literature through the use of examples and key references, incorporating key classic references as well as new literature and papers we think should be well known into the text itself. This edition has a longer bibliography than the previous editions, not only because more has been published, but because we believe strongly that science comes from work published in the scientific literature, and familiarity with this foundation is essential for students. Because of the large number of references, they have been collected into a searchable PDF available online at oup.com/he/gurevitch3e. An appreciation of both classic and contemporary work also helps convey some of the sense of plant ecology as a vibrant, dynamic, and exciting field of study.

Rather than presenting scientific information as a static collection of "facts," we attempt to portray the history and ongoing process of scientific study and discovery. By doing so, we hope to convey some of the excitement and turmoil that that process often involves, while showing how scientists learn how nature works. We extensively rewrote Chapter 1 to provide a stronger (and more modern) introduction to the philosophy of science, the theoretical underpinnings of the field, and the history of plant ecology—topics we think are essential parts of the education of ecologists. Because science is a human endeavor, we show the face of science by including photos of some of the important scientists (both classical and contemporary) whose work we discuss.

With the same goal, we include the first names of scientists whose work we discuss. While this is an unconventional format, we feel that it not only makes science more human, but also reveals the wonderful diversity of those doing important work in plant ecology. It adds something, somehow, in reading about the highly cited work of Waloff and Richards (1977), to learn that the first author was Nadia Waloff and to find out that she was a "formidable chain-smoking Russian entomologist" at Silwood Park of the Imperial College of London in the mid-twentieth century (Michael Crawley, unpublished); and to see, beyond the many Davids and Johns and Jameses, names that include Camille, Katherine, Valerie, Lynn, and Suzanne, and also Vigdis, Xianzhong, Mohamed, Akio, Ignacio, Govindan, Avi, Nerre Awana, and Staffan.

## Acknowledgments

Plant ecology is also a global endeavor. As authors and scientists, the three of us have learned a great deal and benefited enormously from interaction with our colleagues and friends in many other countries. Our travels and sojourns internationally have been invaluable in expanding our understanding and knowledge about plant ecology and the natural world, and in providing the opportunity to take many of the photographs in this book. JG and GAF in particular owe their thanks to the Stellenbosch Institute for Advanced Studies for hosting us for an invaluable visit in 2014. We are delighted to know that students from many

different countries have learned from the previous editions of this book. We hope that this edition will reach many more people in more places in the future.

For the third edition, we received comments, reviews, and corrections from Laura Aldrich-Wolfe, Peter Alpert, Mario Bretfeld, Cynthia Chang, Rebecca Cook, Jeffrey D. Corbin, Robert D. Cox, Michael Fleming, Zachariah K. Fowler, Suzanne Koptur, Daniel Laughlin, Diane Marshall, David McKenzie, Kerrie Sendall, Jeffrey Stone, Sarah M. Swope, Amy Trowbridge, Alexandra Wright, and several anonymous reviewers; we are much indebted to all of them, as well as to our colleagues and students who have offered comments and suggestions and pointed out errors on previous editions and along the way as we worked on this one. JG thanks Alan Robock, who helped clarify many issues and who answered many pesky questions about climate for this book and from whom she learned a great deal about the complex subject of climate science. Graham Chapman and his colleagues contributed at least one joke. Any errors, flaws, and oversights that remain are of course ours.

Textbooks are much more than just the words they contain; a well-produced textbook also includes illustrations that are attractive and instructive, has a useful index, and is laid out and assembled in a way that makes it appealing, readable, and accessible. Sinauer Associates and its new parent company, Oxford University Press, have long records of publishing scientific texts with these qualities, while at the same time managing to make the books accessible by keeping prices considerably lower than other publishers. We are delighted to continue our association with SA and OUP. But books aren't produced by a faceless company pressing buttons; special thanks are due to the skilled professionals who have worked so hard to make it happen. Jason Noe was our Aquisitions Editor, and a key player on our team for this book. We are especially grateful to Kathaleen Emerson, the supervising editor who provided a quiet and skillful hand to steering the project to completion; to Chandra Linnell, our skillful, patient, and driven production editor; and Lou Doucette, our copyeditor, who sometimes knew what we meant (or what we should have meant) even better than we did and who asked rather penetrating questions when she wasn't sure (because we hadn't made it clear). Jan Troutt's scientific illustrations and art grace this edition of the textbook and enhance both the science and the esthetics; it has been a delight to work with her. Mark Siddall's keen eye was invaluable in obtaining and choosing many of the photographs for the book, as well as overseeing the many issues about the photography used to illustrate the science and provide context for the words. Many friends, colleagues, and strangers generously shared their photographs for publication, sharing their passion for the organisms and landscapes they photographed. Michele Beckta meticulously oversaw the crucial job of obtaining permissions for figures and illustrations. Grant Hackett composed the very professional and useful index. The book was designed by project leader Meg Britton Clark, who along with Michele Ruschhaupt created the stunning page layouts. The cover was designed by Donna DiCarlo, and Joan Gemme provided exceptional support and overall project management. And we must mention that the final production of the book was carried out in the course of a global pandemic that had all of us working from home. Thank you all!

We began with a question from Andy Sinauer, and we want to thank Andy not only for his patience and wise advice about the first two editions of this book, but also for his larger contribution to our field. By publishing high-quality books in ecology and evolution for decades and by his encouragement and support for their authors (including the three of us), Andy provided an enormously valuable contribution for the sciences of ecology and evolutionary biology. We wish him well in his retirement.

Our spouses, children, and colleagues tolerated us while we were writing and revising this book rather than doing all of the things we were supposed to be doing or that they wished we were doing. We appreciate their forbearance. To all the students who use this book, we hope that you enjoy the book and learn a lot from it, and that some of you will go on to make scientific contributions of your own.

JESSICA GUREVITCH

SAMUEL M. SCHEINER

GORDON A. FOX

April, 2020

**Go to oup.com/he/gurevitch3e to access the following resources for *The Ecology of Plants* 3e:**

**Instructor Resources**

**Textbook Figures and Tables –** provided in both JPEG and PowerPoint formats.

**Student Resources**

**Literature Cited –** now online as a PDF for reference.

©Omikron/Science Source

1

# The Science of Plant Ecology

The biological science of **ecology** is the study of the relationships between living organisms and their environments, the interactions of organisms with one another, and the patterns and causes of the abundance and distribution of organisms in nature. In this book, we consider ecology from the perspective of terrestrial plants. Plant ecology is both a subset of the discipline of ecology and a mirror for the entire field. In *The Ecology of Plants*, we cover some of the same topics that you might find in a general ecology textbook, while concentrating on the interactions between plants and their environments over a range of scales. We also include subjects that are unique to plants, such as photosynthesis and the ecology of plant-soil interactions, and others that have unique aspects in the case of plants, such as the acquisition of resources and mates. While we focus largely on terrestrial plants, we include freshwater and wetland plants in some discussions. Our emphasis is on the seed plants, particularly eudicots and monocots because they constitute much of the diversity in terrestrial environments, but we also discuss gymnosperms, which are dominant plants in some environments.

Above: The HMS *Beagle* sailed from England December 27, 1831, on a 5-year mission to chart the oceans and collect biological information from around the world.

## 1.1 Ecology Is a Science

Ecologists study the function of organisms in nature and the systems they are part of. Applied ecologists and conservation biologists are particularly concerned with the use of ecological principles to solve environmental problems, while fundamental ecology is concerned with basic knowledge of ecological principles, processes, and patterns. Sometimes the distinction between fundamental and applied ecology becomes blurred, as when the solution to a particular applied problem reveals underlying understanding about ecological systems. In both fundamental and applied ecology, the rules and protocols of science must be rigorously followed.

Ecology is *not* environmental advocacy or political activism, although ecologists are sometimes environmental activists in their personal lives, and environmental activists may rely on ecological research. Ecology is not about one's feelings about nature, although ecologists may have strong feelings about what they study. Ecological systems are complex, with a great many parts, each of which contributes to the whole in different ways. But ecology is indeed a science, and it works like other scientific disciplines.

Here it is important for us to call your attention to a major point. Much of the content of this chapter concerns the nature of science and the scientific method. Many students, at this point, may yawn and conclude that they do not need to pay much attention because they already know about the scientific method, and some students may feel that such discussions are dull and pointless. You might be surprised to know that the scientific method and the nature of science itself have always been the subjects of heated intellectual debate. In recent years it has even led to political controversy and a great deal of confusion among the general public about what is and what is not science and what value science has. The nature of science and the scientific method is the essence of how scientists add to and confirm scientific knowledge, and doing science, as well as learning science, requires a nuanced and thoughtful approach.

How do we know whether something is true? Science is one way of knowing about the world—not the only way, but a spectacularly successful one. In contrast to some of the other ways of knowing that are part of our lives, the legitimacy of science is not based on authority, or opinion, or democratic principles, but on the weight of credible, repeatable evidence.

Why is this characteristic of science so important? Consider the contrast between a scientific approach to an environmental issue—say, the consequences of fragmentation for the persistence of tropical rainforests—and an aesthetic approach. Addressing this issue from a scientific perspective might involve asking questions about how changes in the relative amount of forest edge will affect the physiology of some of the tree species, how these physiological changes translate into effects on population growth, and how dispersal between remaining fragments will affect these populations as a whole. By contrast, an aesthetic approach—often seen in popular literature on conservation—might emphasize the beauty of the intact forest. There is nothing wrong with this approach—indeed, many ecologists speak quite freely about such aesthetic values. But these values are not science; it is not meaningful to debate whether intact forests or fragmented forests are more beautiful, because there is no evidence that one could bring to bear that would settle the issue.

We could make a similar argument if we compared the scientific approach with moral, religious, or artistic approaches: the conclusions one might reach with nonscientific approaches do not depend on testing empirical evidence. This is not to say that only science is worthwhile; indeed, these other ways of interpreting the world play a large and critical role in our individual lives and in human societies. But they are fundamentally different from science.

### Where does scientific knowledge come from?

Throughout this book, we examine how ecologists have come to their current knowledge and understanding of organisms and systems in nature. Ecology has both a strong and a rich theoretical basis and has developed from a foundation based on an enormous collective storehouse of information about nature.

Ecology, like all of science, is built on a tripod of pattern, process, and theory. **Patterns** consist of the relationships between elements or entities of the natural world. Processes are the causes of those patterns. **Theories** are the explanations of those causes. When ecologists carry out original scientific research, they seek to document patterns, understand processes, test and validate their understanding of those patterns and processes, and ultimately put together theories that explain what they have learned.

There is a distinction between the kind of research a scientist does and the kind of research done for a term paper, or by any member of the public trying to gather information about a topic using textbooks (such as this one), library books, or material posted on the internet. Although there are exceptions, the kind of research carried out by students or the general public is usually **secondary research**: gathering and summarizing facts that are already known. This kind of research is not only useful, but essential: every scientific study must begin by assessing what is already known. But the heart of what research scientists do is **primary research**: gathering information that no one has ever known before, confirming or refuting patterns and explanations from other scientific studies, or coming up with new, testable ideas about how nature

works. These experiences of discovery are what make doing science so incredibly exciting and fun.

Scientists gain knowledge by using the **scientific method**. They carry out a series of steps, although not always in a fixed order (Figure 1.1). In ecology, these steps can be summarized as follows: observation, description, quantification, posing hypotheses, testing those hypotheses using experiments (in a broad sense of the word, as discussed below), and verification, rejection, or revision of the hypotheses, followed by retesting of the new or modified hypotheses. Throughout this process, ecologists gather various kinds of information, look for patterns or regularities in their data, and propose processes that might be responsible for those patterns. They often put together some kind of model to help in advancing their understanding. They construct theories, using assumptions, data, models, and the results of many tests of hypotheses, among other things. The building of comprehensive scientific theories proceeds simultaneously from multiple directions and involves numerous people, sometimes working in synchrony and sometimes at cross-purposes. Science in operation can be a messy and chaotic process, but out of this chaos comes our understanding of nature.

The construction of scientific theories is central to the scientific method. The word *theory* has a very different meaning in science than it does in common usage. A scientific theory is a broad, comprehensive explanation of a large body of information that, over time, must be supported and ultimately confirmed (or rejected) by the accumulation of a wide range of different kinds of evidence (Table 1.1). In popular usage, the word *theory* usually refers to a limited, specific conjecture or supposition, or even a guess or hunch. Equating the meaning of a scientific theory with "a guess" has caused no end of mischief in the popular press and in public debates on politically charged issues. A well-known example is the theory of evolution by natural selection: While sometimes portrayed as "just a theory" by creationists and advocates of "intelligent design," it is actually a comprehensive and rigorously tested explanation of an enormous amount of evidence from experiments and documentation of patterns in nature. In fact, it is one of the best-tested theories in biology.

When a theory is buttressed over many years by the accumulation of strong evidence, with new findings consistently supporting and amplifying the theory while producing no serious contradictory evidence, it becomes an accepted framework or pattern of scientific thought from which new speculation can spring. This is what occurred with Einstein's theory of relativity and Darwin's theory of evolution. Scientists use such overarching theories to organize their thinking and derive additional predictions about nature.

The ultimate goal is to produce a **unified theory**, consisting of a few, general propositions that characterize a wide domain of phenomena and from which can be derived an array of models. The best example in biology is the unification of Darwin's theory of natural selection with Mendel's theory of particulate inheritance. This unification—largely complete by the 1940s—allowed biologists to derive many specific models and testable predictions and to amass a large and coherent body of information and knowledge about the natural world, including many discoveries, both practical

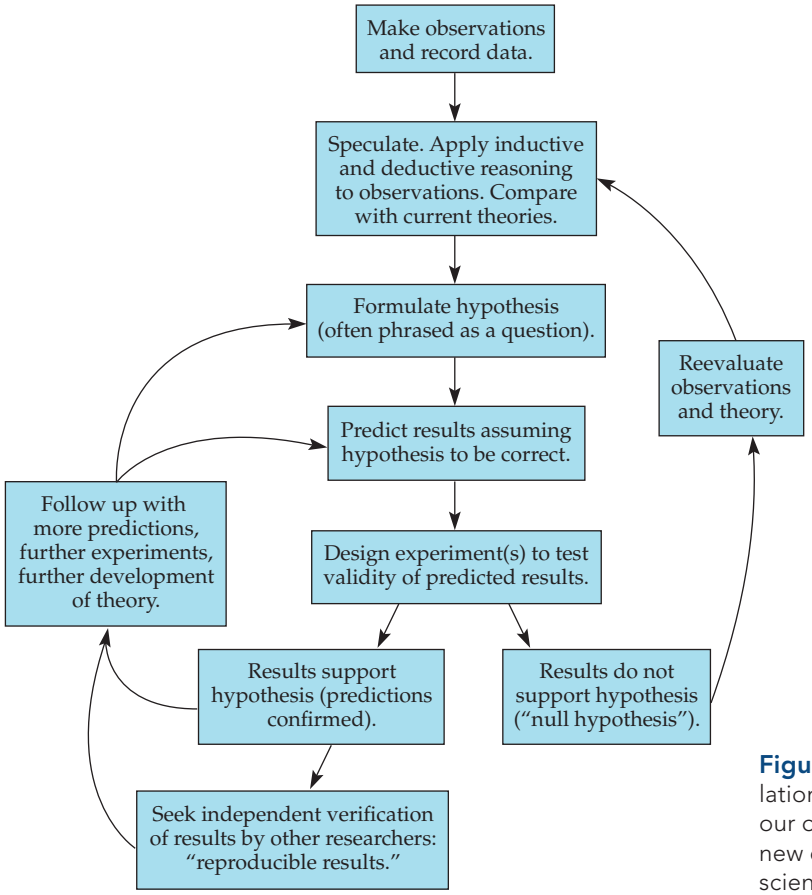



**Figure 1.1** The scientific method. The cycle of speculation, hypothesis, and experimentation is a spiral, with our overall understanding of the world increasing as new questions constantly emerge from the answers scientists obtain.

**TABLE 1.1** The components of a scientific theory

| Component | Description |
|---|---|
| Assumptions | Conditions or structures needed to build a theory or model |
| Concepts | Labeled regularities in phenomena |
| Confirmed generalizations | Condensations and abstractions from a body of facts that have been tested |
| Definitions | Conventions and prescriptions necessary for a theory or model to work with clarity |
| Domain | The scope in space, time, and phenomena addressed by a theory or model |
| Facts | Confirmable records of phenomena |
| Framework | Nested causal or logical structure of a theory or model |
| Fundamental principle | A concept or confirmed generalization that is a component of a general theory |
| Hypotheses | Testable statements derived from or representing various components of the theory or model |
| Laws | Conditional statements of relationship or causation, or statements of process that hold within a universe of discourse |
| Model | Conceptual construct that represents or simplifies the natural world |
| Translation modes | Procedures and concepts needed to move from the abstractions of a theory to the specifics of model, application, or test |

*Source*: After S. T. A. Pickett et al 1994. *Ecological Understanding*. Academic Press. San Diego, CA.

(of benefit to humanity) and fundamental (increasing understanding of living organisms).

A scientific **hypothesis** is a possible explanation for a particular observation or set of observations. A hypothesis is smaller in scope than a fully developed theory. Hypotheses must be testable by containing a prediction or statement that can be verified or rejected using scientific evidence. Experiments are the heart of science, and we discuss their design and use in more detail later in this chapter. A crucial characteristic of science is the need to revise or reject a hypothesis if the evidence does not support it. In science, hypotheses are not accepted based on belief. A scientist should not say, "I believe in human-caused climate change," but rather, "I am convinced by the accumulation of abundant evidence for human-caused climate change."

Some of the most important tools in the scientist's toolkit are models. A **model** is an abstraction and simplification that expresses structures or relationships. Models are a way in which the human mind attempts to understand complex structures, whether in science or in everyday life. Building a model airplane from a kit can tell you a lot about the basic form of an airplane; similarly, civil engineers often build small models of structures such as bridges or buildings (earlier, as physical models and now as three-dimensional images on a computer) before construction begins. You have no doubt seen models of DNA and of chemical reactions, and you may have heard about global climate models, which we discuss at length in Chapter 16.

Models can be abstract or tangible, made of words or plastic. They can be diagrams on paper, sets of equations, or complex computer programs. In science, models are used to define patterns, summarize processes, and generate hypotheses. One of the most valuable uses of models is to make predictions. Ecologists deal almost exclusively with abstract models that can range from a simple verbal argument to a set of mathematical equations. One reason their models so often rely on mathematics is that ecologists are often concerned with the numbers of things. (Is a species' population size so small that it is becoming endangered? How rapidly is an invasive species spreading? How many species can coexist in a community, and how does this number change as conditions change?) Mathematical models offer well-defined methods for addressing questions in both qualitative and quantitative terms, and they require that many assumptions be made explicit. Some ecological models are verbal, some rely entirely on complex computer simulations, and others use relational diagrams (graphs).

All models are necessarily based on simplifications and rest on sets of assumptions. Those simplifications and assumptions (both implicit and explicit) are critical to recognize, because they can alert you to the limitations of the model and because faulty assumptions and unjustified simplifications can sink even the most widely accepted or elegant model. It is often more clear what assumptions are being made in a mathematical or a simulation model than in other model types, but since models are just representations of more complex things, no model ever can state every assumption it requires—any more than a sentence can do so.

## Scientific research involves objectivity, subjectivity, choice, and chance

When you read a typical scientific paper, it may at first seem obscure and difficult to penetrate. The format follows a rigid protocol, designed for efficiently conveying essential information to other scientists. Ideas are tightly packaged, with a clear logical line running from start to finish. It may seem as if the researchers knew exactly what they would find even before they began. We will let you in on an open secret: that is not how much of real science works. The results may not be what was anticipated at the start of the study. The justifications for the research presented in a paper's introduction may have been thought up or discovered long after the research project began, or even after the work was finished. Serendipitous discoveries, surprising natural occurrences, or other unplanned happenstance may modify the original course of a research project. However, this misdirection is now starting to change. Increasingly, the goals and protocols of a project are posted before it is initiated, especially in medical studies. If modifications are necessary, the reasons are made clear when the results are published. And those justifications that were previously added to the paper's introduction should more properly be placed in the discussion section at the end of the paper and considered as new hypotheses to be tested in subsequent studies.

Ideas in science, especially in ecology, come from a variety of sources. While everyone knows that science is held to the standards of being objective and rational, that is only half the story. In order to reach a genuinely new understanding, subjectivity and creativity must also come into play. What one chooses to study is a subjective decision. Do I pay attention to the entire forest or the individual trees? Which forest, and what am I asking about it? Given those choices, there is usually a range of possible places to look for answers—another subjective decision. Do I travel to the Arctic or Amazonia, or study urban forests close to home? Such choices depend on the questions one wishes to ask, but the system one chooses to study also shapes the questions. While determining the answers must be objective, choosing what questions to ask, and how to ask them, is largely subjective.

Many scientific endeavors are highly creative as well. Coming up with a good experiment, looking at a seemingly intractable problem from a new perspective, switching gears after a disastrous laboratory failure, and pulling a large number of disparate facts together to build a comprehensive theory are all highly creative activities. Tests and confirmation must be objective and rational to be science. Starting from the known and leaping across to the unknown requires creative, synthetic, and sometimes other-than-rational thought processes, as in the famous example of Kekule's dreamy vision of snakes swallowing their own tails leading to his discovery of ring structures in organic chemistry.

Many scientific discoveries start with casual observations, as with Newton's proverbial apple. Or an idea may arise as a what-if thought: What if the world works in a particular way? Or a previous experiment may have raised new questions. Sometimes we ask questions about what is not present, or what does not exist, rather than noticing what is present. What makes a scientist most successful is the ability to recognize the worth of these casual observations, what-if thoughts, and new questions. From these sources, an ecologist constructs hypotheses and designs rigorous, objective experiments to test them.

## Observational studies detect and quantify patterns

If we didn't know what patterns exist, there would not be anything to try to explain. Since the earliest humans, observations of nature and attempts to recognize patterns of all sorts have been central to human survival. Early scientists recognized and documented patterns in nature, and this work continues to the present. The first part of finding patterns is to observe what exists and does not exist, and to attempt to generalize those observations. The next step is to quantify observations. Pattern detection and quantification included much of the work of gradient analyses and ordinations (see Chapter 15). Modern observational studies rely on analysis of remotely sensed images, large databases of plant traits and other variables (see Chapter 10), and spatial distribution data. The goal of such studies is documenting and quantifying patterns, rather than hypothesis testing, but the results are often critical to hypothesis generation and future tests.

## Experiments are central to research

A cornerstone of the scientific process is the **experiment**. We use the term *experiment* here in its broadest sense: a test of an idea. Ecological experiments can be classified into three broad types: manipulative, natural, and observational. **Manipulative experiments** are what most of us think of as experiments: a person alters a system in some way and looks for a pattern in the response. For example, an ecologist might be interested in the effects of nutrients on the growth of a particular plant species. One can grow plants under different nutrient treatments, replicating the plants exposed to the different treatments, measure such things as the height at flowering, and ask whether plants under one treatment are taller at flowering than under another. If the treatment groups differ, you have an answer!

This procedure sounds simple, but planning the experiment raises a number of questions. A central question is whether you can perform the experiment

while making sure that the only things that vary are the parameter(s) of interest, such as the amounts of nutrients received by the plants. Classical scientific experiments—first laid out by Francis Bacon in the seventeenth century—vary only a single factor, and you may have learned that this is how experiments are properly done. Can you do this for a plant growth experiment? You might conduct the experiment in a growth chamber or greenhouse. You might try to rigorously control all of the sources of variation in your experiment, but experiments on living things invariably incorporate heterogeneity. Even controlled-environment growth chambers turn out to have environmental variation (e.g., some spots are warmer or cooler than others). You might unintentionally water plants at the rear less than others, because they are harder to reach. You might choose to use seeds that are highly inbred and do not vary genetically. But attempts to control variation have their own problems—the results may not be replicable if a greenhouse experiment is conducted at a different season, when the sun is at a different angle and daylight is longer. Even more problematic, they may not be replicable by other researchers, whose seeds and growth chambers differ. There may be no way to easily generalize your results.

Perhaps worse is this problem: plants grown in pots in artificial environments differ in a number of important respects from those grown in soil outside, so your results might not really be realistic. In sum, this sort of experiment can be useful, but it is also fraught with difficulties. It is easy to fool yourself into thinking that you have controlled all variation except in the factor that you are studying, and even if you have reduced that variation greatly, your results might not be generalizable beyond the conditions of the experiment. What to do?

Garden experiments are more realistic ecologically, with some factors controlled but many uncontrolled, and field experiments in nature may be the most realistic but with only the tested factors controlled and many other factors varying in an uncontrolled fashion. In a field experiment in a natural community, an ecologist might control one or a few factors—reducing herbivory and adding water, for instance, but factors such as soil, competing plants, and pathogens are uncontrolled and varying. One major approach to such experiments is, instead of attempting to control all variation, to randomize the variation due to factors other than the experimental ones among replicates, and base conclusions on the use of statistical inference. For a nutrient experiment in the field, you would need to take into account the fact that the soil probably varies in space, and you might need to think carefully about how to administer treatments so that plants with the same treatment receive the same doses of nutrients at the same time. The major tool used to design and analyze this kind of experiment is randomization. For example, you might assign replicated treatments to different areas (often called *blocks* in statistics). When you later analyzed the data, you would use standard techniques that allow you to account for the possibility that one block is, perhaps, wetter than others. Randomized experiments were first developed by Ronald A. Fisher (one of the founders of both modern statistics and population genetics) in the early twentieth century, and they are a mainstay of ecology and evolutionary biology. Their results are more generalizable than experiments that attempt to rigorously control all variation, because heterogeneity in responses is taken into account in the design and in the analysis of the results, instead of your having to try to eliminate it. Randomized experiments typically require larger samples than those in which you attempt to rigidly control variation. Where along this continuum of control versus realism ecologists carry out their experiments depends on practical considerations and on their scientific goals.

Courtesy of the University of Adelaide, Rare Books and Manuscripts

Ronald A. Fisher

These kinds of experiments (controlled environment, garden experiments, and experiments in natural communities) are called manipulative experiments, and they are powerful tools for two major reasons. First, the scientist can control which aspects of the natural world will be altered. Second, the experiment can separate factors that typically occur together so they can be tested individually. But there are also difficulties with manipulations. One problem is that sometimes they cause artifacts—outcomes that are side effects of the manipulation itself, rather than being responses to the experimental treatment being tested. For example, an ecologist interested in comparing seed production in self-pollinated versus open-pollinated flowers might place netting over some flowers to exclude pollinators. Seeds from those treated flowers would all be self-pollinated, but the flowers would also have experienced reductions in air flow and light, and this could conceivably affect seed production. A thoughtful experimenter might put netting on the "control" flowers but leave the netting open to pollinators as a way to get around this artifact, but it is often impossible in a biological system to really change only one thing at a time. Good experiments avoid or reduce artifacts, or they include ways to take them into account when the results are evaluated. As you read about experiments, consider what artifacts might be present that might explain some of the results.

There are also scales on which we cannot do experiments. Ecology is often concerned with learning about patterns and processes that occur across large extents of space and time—for example, finding why there are differences in the numbers of species on different

continents, or predicting the responses of populations to climate change over the next two centuries. We cannot do manipulative experiments at these great extents of time and space, and in many cases, no true replicates (of continents, for example) could exist. Ecologists are increasingly making use of long-term and large-scale manipulative experiments (Figure 1.2; see Box 5B and Box 12D). Even so, there are often limits to the range of possible treatments. Prescribed fire must often be limited to particular seasons, for example, which may or may

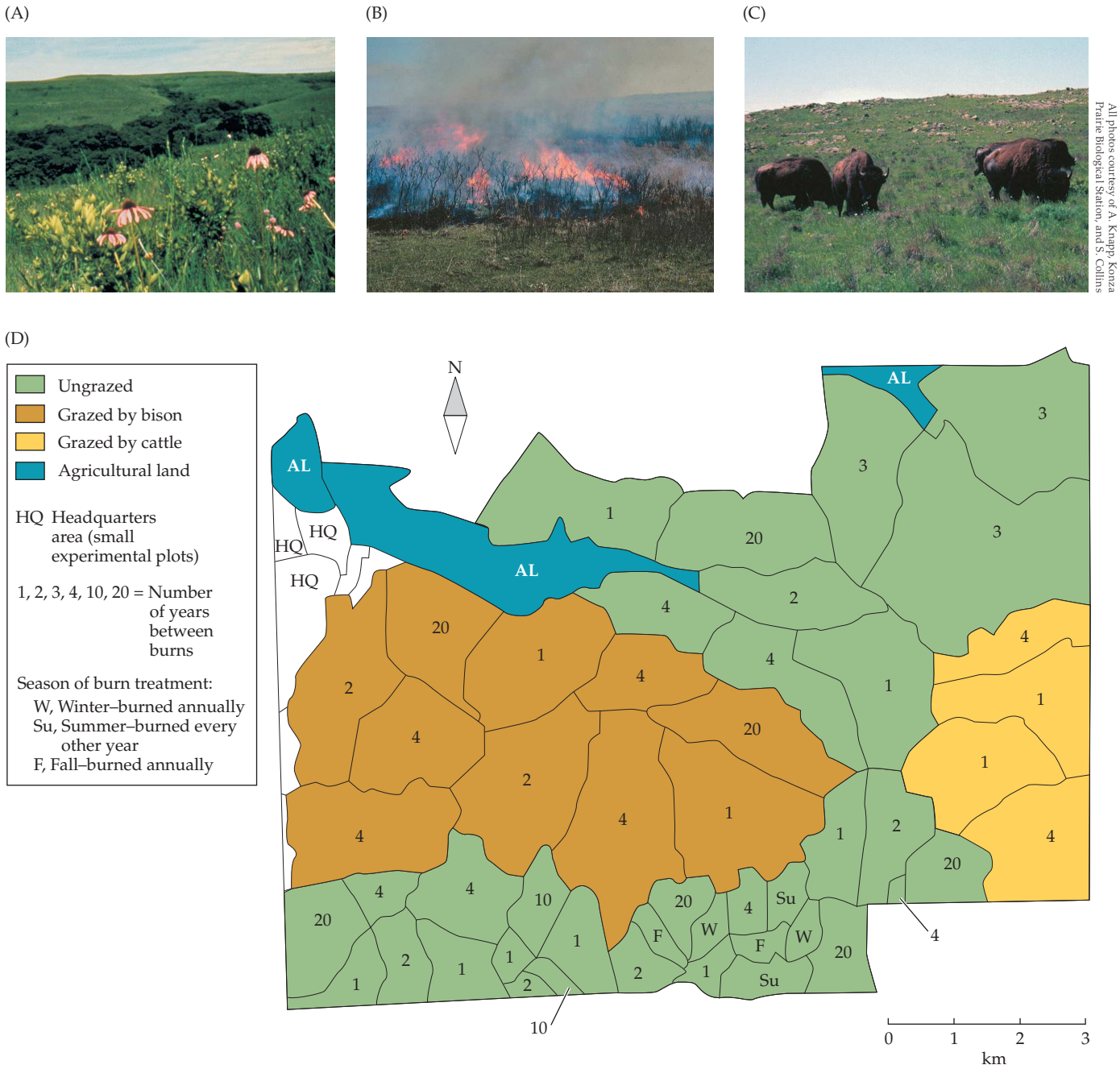



**Figure 1.2** Large-scale manipulative experiments are being carried out at the Konza Prairie Research Natural Area in Kansas (A). Controlled burns (B) are done at various intervals to investigate the effects of fire and fire frequency on prairie communities. In addition, areas grazed by bison (C) are studied and compared with ungrazed areas and with plots subjected to cattle grazing. The experimental patches (D), which are watershed units, vary in size from approximately 3 to 200 ha. In this map, each patch is designated by a code indicating the burn treatment. Patches with the same code are replicates. All burns occur in spring, except for the seasonal burn treatments. (After A. K. Knapp et al. 1998. *Grassland Dynamics: Long-Term Ecological Research in Tallgrass Prairie*. Oxford University Press: New York.)

not be the seasons in which fire occurred naturally in the past. A more subtle problem of scale can occur when different parts of the system respond to the manipulation differently. For example, an ecologist might want to ask how much plant mortality is caused by drought in a desert plant and might design an experiment in which some plots get water added but others do not. Unfortunately, creating small patches of growing vegetation during a drought might well attract large numbers of herbivores, leading to more mortality among the watered plants. Because only experimental plots, and not the entire region, would receive more water, the treatment expected to reduce mortality might well increase it by attracting another source. Screening might exclude the herbivores, but it would also shade the plants and reduce wind on them, causing other responses.

Some types of experiments would be unethical to carry out. For example, we would not cause the extinction of a species just to study the effects of such an event. In such cases, ecologists must rely on two other types of studies. These are natural and observational studies, which may be thought of as different kinds of experiments.

A **natural experiment** is a "manipulation" caused by some natural occurrence. For example, a wildfire may occur in an area. Volcanic eruptions, hurricanes, and accidental introductions of pathogens are all examples of natural experiments whose effects ecologists have studied. Natural and manipulative experiments represent a trade-off between realism and precision, similar to the trade-off between laboratory and field experiments. Just as with a manipulative experiment, the ecologist compares the altered system either with the same system before the change or with a similar, unchanged system.

The major limitation of natural experiments is that there is never just a single difference before and after a change or between systems being compared. For example, if we are comparing sites burned in a wildfire with others that were not, the unburned sites might have been wetter, might have had different vegetation before the fire, or might be different in area. Natural experiments are essentially unreplicated. Therefore, it can be difficult to determine the true causes of any changes we might measure.

The best natural experiments are ones that repeat themselves in space or time. If we find similar changes each time, then we gain confidence about the causes of those changes. Another approach is to combine natural experiments with manipulative experiments. For example, the patches subjected to experimentally manipulated grazing and fire treatments at Konza Prairie (see Figure 1.2) are being compared with patches elsewhere, some of which are also experiencing grazing and fire but are not subjected to experimental manipulation.

**Observational experiments** consist of the systematic tests of hypotheses attempting to explain natural variation. Such observations are experiments if an ecologist starts with one or more hypotheses (predictions) to test. For example, one could measure patterns of species diversity across a continent to test hypotheses about the relationship between the number of plant species and productivity (see Chapter 19). A major limitation of this type of experiment is the potential for multiple factors to vary together. For example, if the number of herbivores is observed to increase as the number and productivity of plant species increases, the ecologist cannot be sure whether the increase in herbivores is a result of increased plant numbers and productivity, or whether the increased productivity is a result of increased herbivory.

As with natural experiments, observational experiments repeated in space or time add confidence to our conclusions (Figure 1.3). Other sciences, notably geology, climate science, and astronomy, rely strongly on observational experiments because of the spatial or temporal scales of their studies, or because direct manipulation is impossible. One way around this limitation is to run an "experiment" using a complex computer model. Variables in the model can be manipulated, and then the output from the model can be compared with empirical observations.

Ecological knowledge comes from combining information gained from many different sources and many different kinds of experiments. The ecologist's use of this complex variety of information makes ecology a challenging and exciting science.

### In ecology, "controls" are what you are using for baseline comparisons

All experiments involve comparisons. For example, an ecologist might compare how much leaf tissue is removed by insects when plants are raised in an environment with either enhanced $CO_2$ or ambient $CO_2$. Without the comparison, it would be difficult or impossible to interpret the cause(s) for the amount of herbivory in the enhanced $CO_2$ environment. Explanations of the scientific method often state that all experiments require a "control" treatment. A classic example is the typical medical experiment: some patients are given a pill that contains a drug, and others are given a placebo, a pill without the active ingredient, to control for psychologically caused effects of taking a pill, which can be substantial. This is an example of a null control treatment, one completely missing the studied factor. Null treatments can be useful but are not needed or even meaningful in all settings. In an experiment studying the effect of moisture availability on plant growth, it would not usually be meaningful to include a "no water" treatment as a control if all of the plants would simply die from no water. Instead, "control" treatments should be comparisons chosen to account for some possible cause, for example, comparing a treatment that just receives natural rainfall with one that includes additional watering.

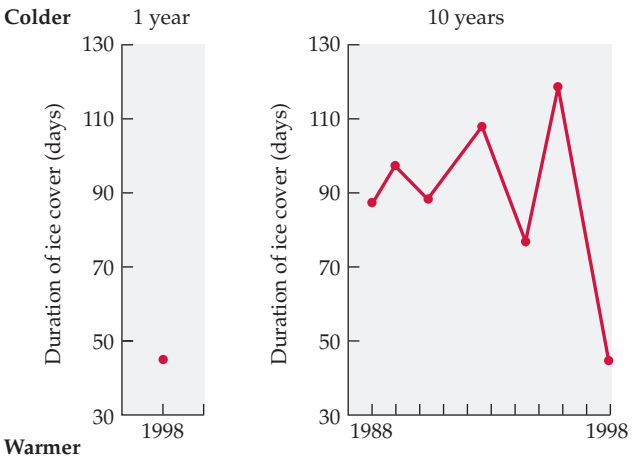
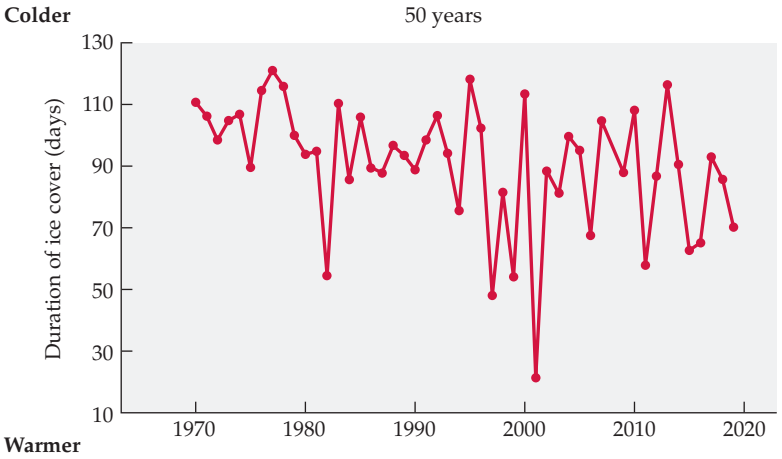
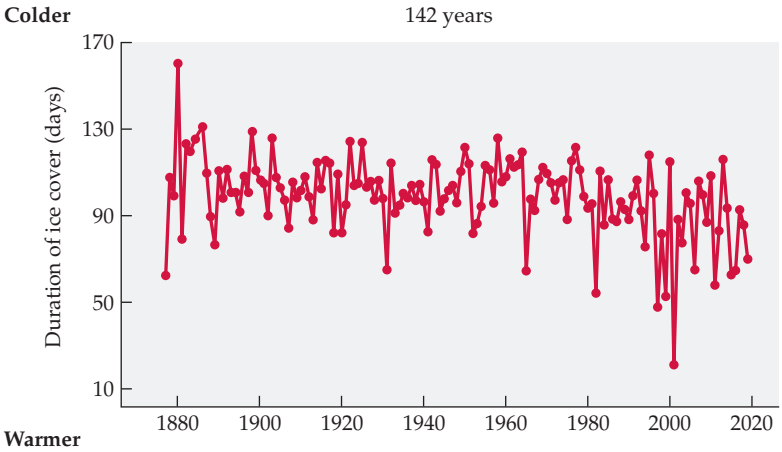



**Figure 1.3** Repeated observations over space or time can reveal information that is not apparent from one or a few observations. As an example, records of the duration of ice cover on Lake Mendota, Wisconsin, have been kept for more than 158 years. The information for a single year is fairly meaningless, but expanding the context with increasing numbers of observations over time shows that there is a cycle of warmer winters recurring every few years (now known to be the result of the El Niño Southern Oscillation; see Chapter 16); and overall, there is a strong trend for winters in Wisconsin to be warmer now than they were in the 1850s. (Data from B. Benson et al. 2000, updated 2013. *Global Lake and River Ice Phenology Database*, Version 1. [Lake Mendota duration]. Boulder, Colorado U.S.A. NSIDC: National Snow and Ice Data Center. Doi: https://doi.org/10.7265/N5W66HP8); Additional updated data from E. Hopkins. 2020. Wisconsin Climatology Office. http://www.aos.wisc.edu/~sco/lakes/mendota-dur.gif.)




All treatments in a well-designed experiment should be chosen to make useful comparisons, and not for any other reason. Similarly, in randomized experiments, what one randomizes actually matters. We randomize to reduce possible biases—for example, we randomly assign individual plants to different treatments, to reduce the chance that plants receiving one treatment differ in some other way than just the treatment. Strict randomization can reduce all sorts of unintended biases, for example, to unconsciously choose the largest plants first for one particular treatment, or to put all of the plants for a particular treatment in a spot that happens to have the most moisture. Randomization, in other words, is a technique that should make our comparisons more meaningful. To then account for those randomized effects, we need statistics.

Statistical analysis of data is an essential tool in ecology and in science more generally. Ecologists use statistics for at least three reasons. First, we use statistics to describe data and search for patterns. In the case of prescribed fire, for example, an ecologist might find that in burned plots the average density of newly germinating individuals of *Pinus ponderosa* (Ponderosa pine, Pinaceae) was more than three times the density in unburned plots, but also that the variation in densities among plots was much greater for those that had burned. Averages and measures of variation are basic statistical descriptions of data. They might allow the ecologist to make statements about the relationship between fire and pine regeneration. Making such statements would involve the second reason ecologists use statistics: to assess hypotheses. In this case, ecologists might hypothesize that germination in this species depends heavily on fire. The third reason ecologists use statistics is to estimate quantities. For example, how much more germination occurs in burned sites, and how much does it vary? We often need these numbers either to evaluate the importance of particular processes, or to use in models (say, models of population

growth or forest cover; see Chapter 8). While ecological statistics is much too large a subject to treat in this book (see Shipley 2000; Scheiner and Gurevitch 2001; Gotelli and Ellison 2004; Lindsey 2004; Fox et al. 2015), notice that almost every figure or table about real data includes statistics—for example, estimates of means, standard errors, and confidence intervals. As you read this book, consider what these quantities tell you.

## How do we test theories?

The testing of scientific theories, especially ecological ones, is a more subtle, nuanced, and complicated endeavor than nonscientists or even students of science often realize. The popular image of the scientific method portrays it as a process of falsifying hypotheses. This approach was codified by the Austrian-born philosopher of science Karl Popper (1959). In this framework, we are taught that we can never prove a scientific hypothesis or theory. Rather, we propose a hypothesis and test it; the outcome of the test either falsifies or fails to falsify the hypothesis. While hypothesis testing and falsification is an important part of theory testing, it is not the whole story, for two reasons.

First, the falsification approach fails to recognize knowledge accumulation. In a strict Popperian framework, all theories are held to be potentially false. We never *prove* anything to be true; we merely *disprove* ideas that are false. This assumption goes against our own experience and the history of the accumulation of scientific understanding. Today we know that the Earth revolves around the sun, even though this was once just a hypothesis. We know that the universe is approximately 15 billion years old (give or take a few billion) and began with the Big Bang, even if we still do not know the details of that event. We know that life on Earth assumed its present shape through the process of evolution. We know that many diseases are caused by viral infections, not by "humours," and that hereditary traits are conveyed by DNA (or in a few viruses, by RNA), not by blood. While we may acknowledge that all of this knowledge has not, in a strictly philosophical sense, been proved to be true but has only failed thus far to be falsified, we also recognize that some knowledge is so firmly established and supported by so many facts—by the accumulation of evidence—that the chance that we are wrong is infinitesimally small (Mayo 1996).

Second, and more important, is that the Popperian framework fails to account for a second type of question that we very commonly ask in ecology. Often the issue is not one of falsifying a hypothesis. Rather, we ask about the relative importance of different processes. When we examine the structure of a plant community, we do not ask, "Is it true or false that competition is occurring?" Instead, we ask, "How much, and in what ways, do the processes of competition and herbivory each contribute to shaping this community?" So, when we are building our theories about plant community structure, our activities are more akin to estimating the necessary quantities and assembling a complex model than to falsifying a set of propositions.

Falsification does play a role in science, but a more limited one than Popper envisaged. Theory construction is like assembling a jigsaw puzzle from a pile of pieces from more than one box. We can ask whether a particular piece belongs in this spot—yes or no—by erecting a hypothesis and falsifying it. We may even conclude that this particular piece does not belong in this puzzle. Less often are we attempting to completely throw away the piece, saying that it does not belong in any puzzle.

Controversy also plays an important part in ecology, as it does in all scientific fields. During the process of amassing evidence regarding the validity of a theory, different interpretations of experimental data, and different weights given to different pieces of evidence, will lead different scientists to differing opinions. These opinions may be passionately held and argued forcefully; discussion can sometimes become heated. As the evidence supporting a theory accumulates, some scientists will be willing to accept it sooner, while others will wait until the bulk of the evidence is greater (see Box 13A).

If the issue under debate has political or economic implications, nonscientists will also contribute to the debate and may be able to offer valuable insight, judgment, and perspective to the discussion. But when the evidence in favor of a scientific theory becomes overwhelming, and the vast majority of scientists knowledgeable in that field are convinced of its validity, then the matter becomes settled (unless startling new evidence or a new, broader theory forces a reevaluation). When a scientific consensus has been reached on a scientific theory, it is unreasonable to consider that theory to be just another guess or opinion and to hold that everyone's opinion is equally valid. That may work for a democratic process, but it is not how science works. Opinions not supported by evidence are not the same as those supported by the weight of a great deal of evidence; giving them equal weight would be contrary to the way science works. The controversy over teaching creationism or "intelligent design" in science classes in American public schools is interesting in this light: Some have argued that since many Americans are persuaded by one of these viewpoints, they should be taught in science classes. Along with nearly all scientists, we argue instead that these ideas are not scientific ideas (because it is impossible to prove or disprove the existence and function of a deity, and no evidence can refute a faith) and that their only potential place in science classes is to illustrate the difference between science and religion.

The fact that scientists are the judges of science should not be interpreted as meaning that scientists should decide issues of public policy. For example, if scientists are in strong agreement about something—say, that if more than 50% of its remaining habitat is lost, then plant species X has a 90% chance of extinction within the next 20 years—that does not necessarily dictate any particular public policy. Policy decisions depend on how important people think it is to save species X and on what costs they are willing to pay to do so. While we personally hope this would never happen, we recognize that people who wanted (for whatever reason) to exterminate a plant species could use the scientific conclusions for their own ends, just as we could use those same scientific conclusions to promote conservation.

Ecologists also have a responsibility to carry out their research in an ethical manner. We have already mentioned one imaginary pertinent example: deliberately causing a species to go extinct to study the effects of that extinction. In that case, the ethical position is clear—unless, of course, we are considering the extinction of a deadly pathogen, such as the malaria parasite *Plasmodium falciparum.* Other situations are more complex, however, involving instances in which one must weigh the ethical values on either side of an issue. For example, how much of a protected area is it acceptable to disturb in order to study the processes that affect it? Is it acceptable to dig up some plants or cut down multiple trees? What if the study increases the chances that an endangered species might go extinct, even if only by a very small amount? A new field of ecological ethics is being developed that focuses on establishing ethical principles for ecological research and procedures for resolving ethical dilemmas (Minteer and Collins 2005).

We also have a responsibility to be ethical with respect to the science itself. Other scientists need to know how we gathered our data and how we analyzed it. In principle, the steps in every study should be understandable by any scientist, and therefore replicable. Many fields of science—including ecology—are experiencing what has been called a "replicability crisis" because single results are too often accepted without further replication. The movement toward "open science" emphasizes replicability, transparency, and other aspects of how we conduct science. We discuss several examples in this book in which ideas became widely accepted with remarkably little empirical support. Insofar as the scientific literature is not clear about what was done in a study, this cannot be resolved. These and other problems ecology faces in terms of academic transparency are summarized by Timothy Parker and his colleagues (2016).

Replication can be especially difficult in ecology. As Shinichi Nakagawa and Timothy Parker (2015) discuss, part of the problem is that different settings for studies are different. Can a study done in a tropical forest in Peru in 2015 be precisely replicated if we try to do so in a tropical forest in Brazil in 2020? Not always! But as Nakagawa and Parker point out, there are different sorts of replication. Replication of studies can be done to fairly account for measurement error, or it can be done to ask how general some results are. Many results can be expected to vary quantitatively if studies are done in different locations at different times, but if the results can be qualitatively replicated, they may have some generality. For example, the removal of herbivores in two different experiments might increase the total number of species in a plot, even if the exact numbers of species and their identities differ. While traditionally academic success at all levels has depended heavily on developing new results in new systems, the replicability crisis tells us that some of this is illusory: we need more replicative studies to actually assess our ideas.

Beyond being scrupulously honest in your experimental design, data collection, analysis, and writing, what can an individual scientist do? Deliberate dishonesty makes the replicability crisis worse, but it is only a very minor cause of the replicability problem. A number of things can help. Data were once regarded by many scientists as private property, to be guarded jealously. Today many journals and funding agencies require that the entire data set (not just the published portion) from a study be deposited in a freely available online archive, along with sufficient information (metadata) that a reader can interpret each column and row in the data file. Many journals and agencies also require that the computer code used to analyze the data be freely available online. Using these two measures, any scientist should be able to see how results were obtained and also reanalyze the data as desired. We strongly support these measures, and we urge future scientists to make their data and computer codes available even when not required to do so.

Another measure advocated by open science initiatives includes preregistration of studies and plans of data analysis. This is simplest to think about in the case of a drug study: if company A preregisters a study of what it hopes will be a promising new drug, and the study does not find that the drug is very helpful, the results cannot simply be hidden. Similarly, if you want to study the effect of prescribed fire on soil nitrogen, a preregistered study makes it clear what you plan to do and why. You might still change your plans, but then you would need to make clear that you did so, and why. Preregistration helps to reduce the problem of selectively publishing only studies that favor a desired outcome, sometimes called cherry picking. Cherry picking can also occur within a study: measuring 20 variables and then only publishing results about the one that you find interesting. But if data and computer code are public, including the parts of the data

that we choose not to pursue, cherry picking becomes obvious. Highlighting unexpected findings is okay, but you need to treat those as new hypotheses to be further tested, rather than as a test of a hypothesis that was actually built only after seeing the data.

Open science presents both opportunities and challenges for ecology and for science as a whole. Not only can it help make our science more robust, it can foster interactions among scientists and between scientists and interested individuals outside of science. Learn more about it, and we hope that you too will embrace it.

### Studies can lead to specific results but contribute to general understanding

Because ecologists work at such a variety of scales and on such a diversity of organisms and systems, the question arises about how far one can extend the conclusions of a particular study to other organisms or places. In the fields of chemistry and physics, the results of an experiment are considered to be absolutely true for all times and places: an atom of helium is made up of two protons and two neutrons, which in turn are made up of quarks, with no qualifications needed. This is the popular image of scientific theories.

Ecology is different. Do the results of a field experiment on competition between two plant species extend to other seasons or locations, or to other pairs of species within the same families or functional groups? Experiments involving helium deal with a universal entity, the helium atom. In contrast, in experiments on plant competition, the exact composition of the entities changes (e.g., the individual plants used each time are not genetically identical), and the surroundings change as well (e.g., the weather this year is different from last year). For this reason, extremely cautious scientists take the position that no conclusion can be extended beyond the particular conditions that existed when the experiment was conducted. These ecologists have argued that ecology is a collection of case studies. We do not agree. If this were so, ecology would not be a science, and there would be little value in doing any experiments, because anything they would tell us would be of such limited scope as to be virtually meaningless.

The truth is somewhere between the extremes of universal truth and a collection of disparate cases, creating a constant and dynamic tension in ecology. One approach to resolving this tension is to see how the outcome of a particular experiment fits into the workings of existing models, and whether it supports or rejects the predictions of those models. Another approach is to use methods for the quantitative synthesis of the results of independent experiments. These methods, known collectively as **meta-analysis** (see Box 10B), can be used to evaluate where the outcome of a particular experiment fits in with—or differs from—the results of other similar experiments conducted on different organisms at different places and times. This approach has been used to evaluate the broad body of experimental evidence for many important ecological questions (Gurevitch et al. 2001).

### Science is ultimately consistent, but getting to consistency is a challenge

Science demands internal and external consistency. Ultimately, theories must be consistent with one another, and data must be consistent with theories, although contradictory data and theories can coexist for long periods of time before they are reconciled. Other ways of interpreting the world do not share this characteristic. Works of art can be self-contradictory. Systems of morality or religions may or may not include obvious contradictions, but none demand consistency with data, in any sense of the term.

It is important not to take this too far and conclude that only science is useful. Science is useful for addressing scientific questions (such as whether wildfires increase or decrease the species diversity of a forest), but not questions that cannot be addressed scientifically. Science cannot tell you how to behave, whether a novel is good, or what color clothing you should wear.

Making science internally and externally consistent is a constant effort. Theories—even successful ones—can contradict one another in places. Some experimental results seem to contradict theory at times. Well-designed studies can contradict one another. This is the stuff that allows our knowledge to continue to grow. The fact that we find contradictions simply means that we are still learning. Resolving those contradictions can be some of the most exciting areas of research.

## 1.2 Ecological Phenomena Are Heterogeneous in Many Ways

All electrons are the same. All $^{12}C$ atoms are the same. But we cannot make statements like this about ecological phenomena or processes. Individual plants within a clone are different from one another, and different salt marshes are certainly quite different from one another. Similarly, processes like herbivory or carbon cycling vary over time and space.

Most things ecologists study are heterogeneous, and we often need to account for that heterogeneity. A glaringly obvious example concerns weather: "average weather" is not actually what occurs anywhere, ever. If a study site experienced the average precipitation, there would be constant rainfall at a location—but of course, the site really experiences stormy periods and periods without much precipitation. Thinking about averages can be useful (it really is true, and meaningful, that on

average, London is rainier than Tucson, Arizona) but for many purposes it is a mistake to think of averages as the only things that matter. Within a plant population, individuals experience different microhabitats, and (usually) they differ somewhat genetically; that heterogeneity can be important!

A great deal of recent interest in ecology has been generated by consideration of how ecological patterns and processes vary as a function of the scale at which they operate and are studied (see Figure 15.2). The same phenomenon can be seen very differently when studied within a small local area and across a landscape or region—that is, at different spatial extents. Likewise, one's perspective can change dramatically when studying an ecological process over a single growing season of a few months or over a period of decades or centuries (see Figure 1.3). Different kinds of things may be going on over different spatial extents, and expanding one's focus to more than one extent can be richly rewarding. In a study of a local community, for example, we might see that competitive interactions keep individual plants of a particular species at a distance from one another. At a larger extent, we might notice that the plants are grouped together across the landscape, because individuals that are too far apart from any others never become pollinated and fail to leave descendants or because the seeds have limited ability to disperse. At a regional scale, herbivory might be important in determining plant densities and occurrence, while at a continental extent, the plants may exist in several large but separated enclaves, determined by patterns of glaciation and species migration thousands of years in the past.

We often refer to these scale changes in terms of a hierarchy, and one can move up and down many different kinds of hierarchies in ecology. For instance, one can move from the level of molecules to tissues to organs to entire organisms. A different kind of hierarchy could expand from individual organisms to populations to communities to ecosystems and up to entire biomes; an alternative hierarchy might move from things that occur at the level of organisms to those that function at the level of habitats, landscapes, **watersheds**, regions, and so on up to global phenomena. These different kinds of levels are not necessarily congruent. One might, for instance, study the individual adaptations of plants over a range of different environments across an entire landscape or even a region, or consider how population interactions at local extents contribute to the global range limitations of a species. Likewise, one's interpretation of data collected over a short time period may be completely upended when the same data are examined for trends over longer periods of time.

One of the reasons scale is now recognized as being so central to ecology is that the world is a very heterogeneous place. Even over very small distances, conditions can change in ways that may be important to living organisms. Environmental conditions are a particular concern in plant ecology because plants cannot move—or, at least, mature terrestrial plants generally are firmly rooted in place, although their offspring may be dispersed some distance away. So, the environment immediately surrounding an individual plant is overwhelmingly important to its survival, growth, and reproduction.The **habitat** of a population or species is the kind of environment it generally inhabits, and it includes the set of **biotic** (living) and **abiotic** (nonliving) factors that influence it in the places one usually finds it. But the conditions in the immediate surroundings of an individual plant—its **microhabitat**—may differ considerably from the average conditions in the general habitat (see Figure 15.3). Factors operating to distinguish a microhabitat from others around it include the composition of the soil; the microclimate of the immediate area; the presence, size, and identity of neighboring plants; and other organisms in the immediate surroundings (e.g., grazers, pollinators, seed eaters or dispersers, and mutualistic or pathogenic fungi or bacteria).

Similarly, the environment varies from moment to moment. There are no specific ecological terms for the components of temporal heterogeneity, but time also exists at many extents, and that has major consequences for plants. Variations in conditions from day to night; from summer to winter; across periods of wet years, cold years, or snowy years; and at a longer extent as climate changes over thousands of years all have important influences on plants. Depending on the ecological process being studied and the organisms involved, it may be the small-scale, moment-to-moment variation that matters most (such as fluctuations in light levels in a small forest gap on a partially cloudy day), or it may be long-term average conditions (such as $CO_2$ concentration in the atmosphere), or it may be the interplay between processes occurring at different durations (such as $CO_2$ flux in a forest canopy over the course of a day or a season).

Groups of organisms, such as populations and species, sometimes average these microenvironmental influences over larger areas and over generations of organisms' lives. This averaging acts to counter the effects of heterogeneity, particularly over evolutionary time. At even larger scales, heterogeneity again becomes critical. As continents are carried apart on tectonic plates and climates are altered, organisms must respond to changing conditions by evolving or changing their distributions, or else become extinct. However, there are many situations in which the heterogeneity, and not the average, is what matters. A major theme in much current ecological research is understanding the interplay between heterogeneity (whether temporal or spatial) and long-term or large-scale averages.

## 1.3 Plant Ecology Has Developed through the Interaction of Observation, Measurement, Analysis, Technology, and Theory

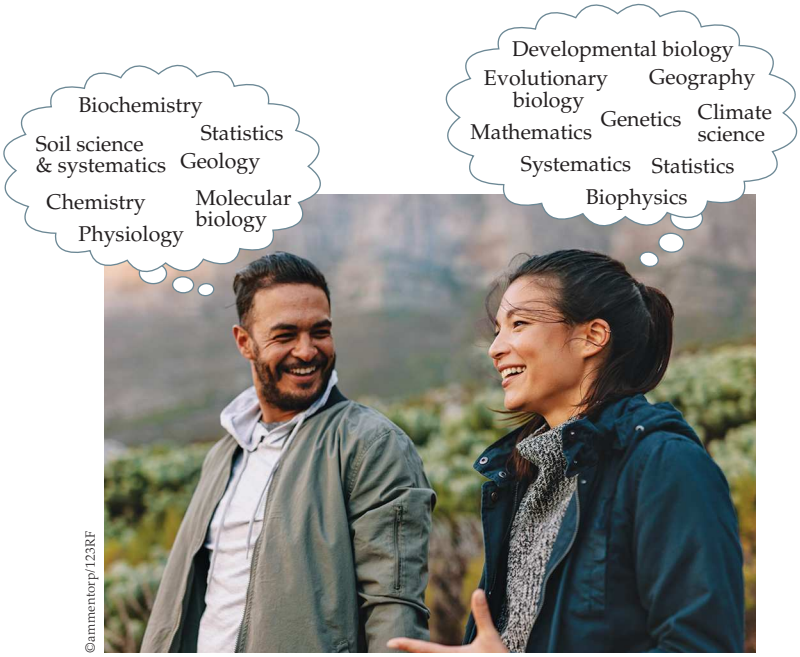
Ecology is a synthetic subject. By that we do not mean that it is unnatural or artificial, but that it brings together a very wide range of other fields of science (Figure 1.4). Some of the fields that ecology encompasses or overlaps with include geology, geography, climatology, soil science, anthropology, sociology, evolutionary biology, genetics, statistics and other branches of mathematics, systematics, behavior, physiology, developmental biology, molecular biology, and biochemistry. We touch on many of these fields throughout this book, showing you how they fit into the toolkit of an ecologist and how familiarity with them affects the ways in which ecologists think about and study organisms in nature.

This is not the place to present a detailed and definitive history of plant ecology. Instead, we sketch some of its major milestones, with an admitted bias toward the English-speaking scientific community. Other historical details are scattered throughout the book as we discuss particular topics and subfields. While no single definitive history of plant ecology exists, several books and papers describe parts of its history (McIntosh 1985; Westman and Peet 1985; Nicholson 1990; Allen et al. 1993).

Plant ecology began with simple observations, because even in prehistoric times, people's health and survival depended on their abilities to understand many aspects of the ecology of plants. Ecology as a science began with the Greeks, most notably Aristotle, in the fourth and fifth centuries BCE. He and his students wrote about relationships among some plants and animals, but much of this knowledge was descriptive in nature.

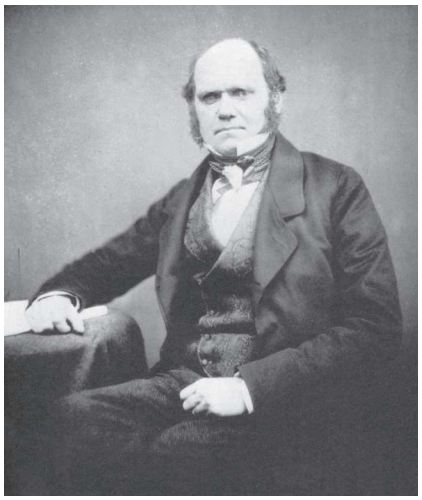
The development of worldwide transportation was essential to the modern science of plant ecology. Begun as what was called natural history in the eighteenth and nineteenth centuries, ecological science was conducted at first by professional and amateur naturalists in Europe and North America and in their travels throughout the world. Once they could travel readily, it was possible to discern many of the patterns that are now well known to ecologists and to begin to consider how those patterns may have developed. The travels of the early nineteenth-century Prussian naturalist Alexander von Humboldt (see Figure 18.10)—at one time deservingly one of the world's most renowned scientists—led him to systematize information on the effects of altitude and air pressure on patterns of temperature and precipitation. Von Humboldt was also the first to codify our understanding of how coastal climates differ from those inland. From von Humboldt's research comes our current understanding of the major causes of climate in the world, as well as their connection with the major causes of patterns in vegetation. While earlier sailors had certainly noticed that the plants in, say, Brazil were different from those in England, von Humboldt was the first to generalize descriptions of these patterns, discovering and writing about what we now call biomes, as well as proposing how these patterns were related to variation in climate. A fine (and very readable) introduction to his life is the book by Andrea Wulf (2015).

Charles Darwin was another traveler whose work is one of the foundations of ecology. As the ship's naturalist on the British ship HMS *Beagle*, a vessel commissioned to provide geological information to the British navy, Darwin acquired the basis for his later work on the theory of evolution by natural selection. The story of his travels from 1831 to 1836 became a widely read book after its publication in 1839 (Figure 1.5). The patterns he described (in geology, but also patterns of the sorts of animals and plants in Brazil, Argentina, Chile, Ecuador,




©ammentorp/123RF

**Figure 1.4** Ecology requires information from many fields of science and mathematics. No ecologist is conversant with all of these, but all ecologists need to know something about some of these other disciplines. There are good reasons your ecology instructors expect you to have some background in other areas!



©Omikron/Science Source

**Figure 1.5** The HMS *Beagle* sailed from England December 27, 1831, on a 5-year mission to chart the oceans and collect biological information from around the world. Charles Darwin sailed with the *Beagle* as ship's naturalist; he is pictured here at the age of 27, shortly after completing the voyage. Darwin collected vast numbers of plant and animal specimens and recorded copious scientific observations that were instrumental in the creation of his most famous work, *On the Origin of Species*.

Australia, and elsewhere and the fossil animals of Argentina) were new to the world and, in many cases, quite astonishing. Darwin later published a key theory explaining much of the world's biota in his work, *On the Origin of Species* (Darwin 1859). While this book is famous for its ideas and their continuing influence, it was not the only book that reflected and attempted to synthesize the new information that was available from the travels of European ships; indeed the period from the late eighteenth century through the nineteenth century saw many important discoveries and syntheses of the natural world, especially in biology and geology. The first map of mean monthly world temperatures was published in 1848. Two decades later, in 1866, the first world vegetation map was produced. Over the next several decades, a number of naturalists developed formal classifications of plant communities, noting the relationships between different types of communities (such as forests and grasslands) and the climate at different latitudes and altitudes.

What was different about what Darwin and von Humboldt contributed was that beyond their observations, they proposed theories to explain many of those observations, as well as verbal models that were applications of those theories to more particular problems. Like all science, ecology has depended not only on observation and on technology that permits new sorts of observation, but also on theory that makes sense of the observations and suggests new testable hypotheses and new ways to look at the world.

Ecology as a recognized discipline coalesced in the latter half of the nineteenth century. The German biologist Ernst Haeckel, a major voice in support of Darwin in central Europe and influential morphologist of marine animals, coined the term *oecology* in 1866. Among the first to write specifically about the topic of plant ecology was the Danish scientist J. Eugenius Warming, considered by many to be the founder of plant ecology as a distinct field. During the period from the 1870s to the end of the nineteenth century, Warming developed an evolutionary, adaptation-based perspective, and he created the concept of plant communities. During the same period, the German scientist Andreas Schimper created the first map of plant distributions, information that was critical to early progress in plant ecology. By the early twentieth century, the Ecological Society of America and the British Ecological Society had formed.

Developments in many other fields have made modern ecology possible. Understanding of weather and climate also depended on the development of travel and, more recently, on fast communication and data storage. Prior to the mid-1700s, no one had realized that weather moves in predictable ways across the globe—news traveled far more slowly than weather (McIlveen 1992). On October 21, 1743, Benjamin Franklin attempted to observe a lunar eclipse in Philadelphia but was prevented from seeing it by a storm. Later, he was surprised to learn that the eclipse had been visible in Boston and that the storm had arrived there the following day. By contacting people living between the two cities, he was able to reconstruct the movement of the storm. It was the development of communication technology, however, that really changed the sciences of meteorology and climatology. When the telegraph became available after 1844, it became possible to organize large numbers of people to observe and forecast the movement of storms. Modern weather data wasn't

available, however, until well into the twentieth century. Accumulation and analysis of large quantities of data required the development of modern computers, which has also greatly affected ecology.

Two other fields deserve special mention: genetics and statistics. Many topics in ecology could not be addressed without genetic data and an understanding of the processes underlying changes in gene (and phenotype) frequencies. And no data would be interpretable without modern statistics, including graphs, which were invented in the late eighteenth century by William Playfair (1786), a Scottish engineer and economist. Both of these are areas that continue to develop and continue to permit new ecological insight. It is also now clear that much evolutionary change occurs more rapidly than was previously thought, so ecological and evolutionary dynamics can play out together, influencing each other. The older view was that the two types of processes happened on very different time scales and that ecology provided the "theatre" in which an evolutionary "play" was performed (Hutchinson 1965). Instead, the two sorts of processes are closely interlinked, and often one cannot be studied without considering the other.

## Plant ecology is situated in the more general theoretical framework of ecology

There is a general theoretical framework of ecology, and plant ecology fits within it. Ecology concerns the spatial and temporal patterns of the distribution and abundance of organisms, including causes and consequences. By distribution and abundance, we mean that ecological studies focus on numbers of individuals (or related properties such as biomass or size) and on numbers of species and how those individuals and species vary across space and time. Although the examination of causes makes up the bulk of the science of ecology, consequences also are an important component. For example, climate change is driven in large part by the accumulation of $CO_2$ in the atmosphere, and that accumulation in turn is driven in part by how plants cycle $CO_2$ through the biosphere (see Chapter 2 and Chapter 16).

What are the basic principles of ecology? They are actually rather simple (Table 1.2). The complexity of ecology lies in thinking out what their consequences are, how they interact with one another, and when each of them is important (Scheiner and Willig 2011a):

1. Organisms are heterogeneously distributed. Large-scale heterogeneity is discussed in Chapter 15, Chapter 17, Chapter 18, and Chapter 19, while heterogeneity at the scales of individuals underlies much of Chapter 2, Chapter 3, and Chapter 4.
2. Plants interact with one another, as well as with animals, fungi, and other kinds of organisms. Some of these interactions cause heterogeneity in space or time, some are consequences of that heterogeneity, and some are both causes and consequences. Part I of this book is primarily concerned with abiotic interactions; Parts II and III are primarily concerned with biotic interactions.
3. Contingencies ("accidents of history") affect the distributions of organisms and their interactions. This idea has grown in importance in ecological theory over the past 50 years. A seed lands in once place but not another, and a particular species originates on a particular continent, setting the stage for later events. Contingencies play a particularly important role in Part II of this book.
4. Individual organisms vary in their characteristics. In turn, this variation creates variation in ecological patterns and processes. The individual oak trees in a forest vary in their photosynthetic rates for a number of reasons, and this affects many characteristics of the

**TABLE 1.2** The general principles of the theory of ecology

1. Organisms are distributed heterogeneously in space and time.
2. Organisms interact with their abiotic and biotic environments.
3. The distributions of organisms and their interactions depend on contingencies.
4. Variation in the characteristics of organisms results in variation of ecological patterns and processes.
5. Environmental conditions as perceived by organisms are heterogeneous in space and time.
6. Resources as perceived by organisms are finite and heterogeneous in space and time.
7. Birth rates and death rates are a consequence of interactions with the abiotic and biotic environment.
8. The ecological properties of species are the result of evolution.

*Source*: S. Scheiner. 2010. *Q Rev Biol* 85: 293–318.

population and forest. Individual plants within a population vary in their chances of birth and death, and this affects population growth (see Chapter 8).

5. The environment varies in space and time. But most important, the variation is of a type to affect plants, animals, and other species. It is warmer in one spot than another; one location has sandier soil than another nearby. This variation is important for the processes in Part I, and plant growth is affected by this heterogeneity (see Chapter 6).
6. Resources are finite and their availability varies in both time and space and on many scales. Water is available to some plants (depending on their characteristics—see principle 4) at some times of year but not others. This resource heterogeneity is again important for the processes in Part I, and it strongly affects plant-plant interactions (see Chapter 10).
7. All living organisms were born at some point, and their death is inevitable. The rates at which births and deaths happen are results of how organisms interact with both the abiotic and biotic environment (see Chapter 6, Chapter 7, and Chapter 8).
8. Organisms, and therefore their ecological properties, are consequences of evolution. Evolution plays an important part in every chapter, and the processes are discussed in detail in Chapter 9.

We have deliberately retained the word *organisms* in this list because these principles apply to all organisms. But if you substitute the word *plants*, you will find a basic list of statements that provide much of the structure of plant ecology. This list is a bare sketch of the theoretical underpinnings of ecology; more are found throughout this book. You also can glean much more from the book edited by Samuel Scheiner and Michael Willig (2011b), which is both about ecology and its theory in general, and about particular theories within ecology. Ecology is a subject with rich theoretical underpinnings that are constantly being extended.

## Ecology has a range of subdisciplines

Plant ecology as a discipline is made up of a number of different subdisciplines, some of which have quite distinct traditions and histories. Some early plant ecologists and botanists focused on whole communities, while others focused on single species and the properties of individuals. The older (now largely archaic) terms for these two subfields are **synecology** and **autecology**. Plant community ecologists, in particular, were active in the origins of ecology as a discipline in the last part of the nineteenth century and dominated plant ecology during the last two-thirds of the twentieth century. A more detailed discussion of the history of plant community ecology and some of the key figures in that history is given in Chapter 12.

Early studies in plant autecology were especially concerned with understanding unique plant adaptations to extreme environments, such as deserts, and a number of famous studies were concerned with plant performance in the field. Although some major insights were gained, technological limitations severely hampered the development of the field. As instrumentation and methodology became more sophisticated, plant physiologists began to carry out most of their research in controlled laboratory environments.

Considered then as part of autecology, as far back as the nineteenth century, individuals in many countries around the world were carrying out studies that today we would call plant physiological ecology or plant population ecology. Around the middle of the twentieth century, autecology began to divide into subfields that focused on single individuals and on populations. Studies of individuals were enhanced by further advances in technology that made it possible for physiological studies to come out of the greenhouse and into nature, creating the fields of plant physiological ecology and functional ecology. Plant population ecology as a recognizable subdiscipline had its origins in Great Britain in the 1960s, particularly with John Harper and his students. It then spread to North America in the 1970s.

For the most part, during the first three-quarters of the twentieth century, plant ecology developed independent of animal ecology. Animal community ecology has a long history parallel to that of plant community ecology (Mitman 1992). Substantial work in animal population ecology extends back to at least the 1920s (to the work of G. F. Gause, Raymond Pearl, Alfred James Lotka, and others). Plant population ecology drew on these ideas and theories as it was developing, as well as on other ideas that originated among plant ecologists. Eventually, new theories were needed as discoveries about the unique nature of plants made it obvious that they could no longer be shoehorned into many of the theories constructed for animals.

Conversely, physiological ecology advanced earlier and more rapidly among plant ecologists than it did among animal ecologists. Undoubtedly this was because the characteristics of plants are much easier to measure, and their environments easier to characterize, than those of animals (for most purposes, one does not have to catch plants!). On the other hand, in the 1980s,

animal physiological ecology joined with evolutionary biology to create the field of evolutionary physiology (Feder et al. 1987), a move that plant biologists have not yet clearly made.

The gap between the fields of plant and animal ecology was bridged in the 1970s, although distinct subfields continue to this day. Two related developments were responsible. The first was the rise of studies of plant-animal interactions, especially pollination (see Chapter 6) and herbivory (see Chapter 11). The second was the burgeoning interest in the evolutionary aspects of ecology in the 1970s and 1980s, which transcended the traditional separation of the studies of plants and animals.

Recent changes in the field of plant ecology include the rise of landscape ecology and conservation ecology as recognized disciplines in the late 1980s. Landscape ecologists came to the discipline from various different directions, including fields as diverse as plant community ecology and remote sensing. Conservation ecologists likewise created their field from backgrounds in mathematical modeling and population, community, and ecosystem ecology. The 1990s saw the creation of the discipline of urban ecology, of which plant ecology is an important component, and the general recognition that nearly all parts of the globe have been affected by humans to at least some extent. At around the same time, the fields of biogeography and biogeochemistry emerged, building in part on areas of plant ecology and on new technological developments and interactions with remote sensing and climatology. Other fields within plant ecology have seen major shifts in emphasis. Plant community ecology has seen a large shift from questions about whole-community patterns and processes to a major focus on questions about interactions within and among species.

A major trend in contemporary ecology, including plant ecology, is toward larger, more integrated research projects that involve many collaborators and examine phenomena across large extents of space and time or across levels of organization. Except in the subdiscipline of ecosystem ecology, which was undertaking projects with large teams of scientists in the 1960s and 1970s, such multi-investigator studies were very rare in ecology until recently. Current ecological research, including that in plant ecology, is almost always the work of collaborations among scientists rather than the work of single individuals. These collaborative groups are often international, facilitated by the ubiquity of electronic communication. Contemporary studies in plant ecology range from molecular genetics up through ecosystems and social systems, and they are erasing many of the traditional boundaries among subdisciplines. Plant ecology is experiencing exciting times, and we hope you will sense and share that excitement in this book.

## Science is a human endeavor

Science is a way of understanding; it is not a list of disembodied facts that somehow were handed to us anonymously. Of course you know this, but it is easy to lose sight of the fact that scientists are real people who do scientific research. Throughout this book we try to make it a bit easier to keep this in mind. We often refer to researchers by their full names. We have included photos of many scientists (living and dead) whose contributions to plant ecology are important. They vary quite a bit: they are women and men of many countries and nationalities, as well as different times. What they have (or had) in common is a fascination with the natural world, especially of plants, and a drive to understand it better. Some of you will become scientists as well; in any case, we hope that you will learn to appreciate the contributions these people have made.

We began this chapter by saying that ecology is a science and that it is distinct from environmentalism. We stand by those statements, but we think it is important to add something: ecology is a *useful* science. Most people become ecologists because they are interested in the natural environment and want to help protect it. When we authors were graduate students, this was also true, but there has been a big change since those days: in academia few ecologists then worked on applied problems, like conservation, although ecologists employed within federal and state agencies and nongovernmental organizations certainly did. Most academics worked instead on what they considered fundamental science, in systems that they thought were undisturbed.

Today we recognize that no place on Earth is free of human influence and that we may often be able to advance the science of ecology by furthering its application. Despite the enormous environmental problems the world faces, there is a great deal of beauty that remains and is worth saving—at all scales, from continents and biomes down to individual plants and their organs. It is also quite wonderful to learn about. We have been studying it for many years and still find it gratifying and exciting to learn more. We hope you will too!

# PART I

# Individuals and Their Environments

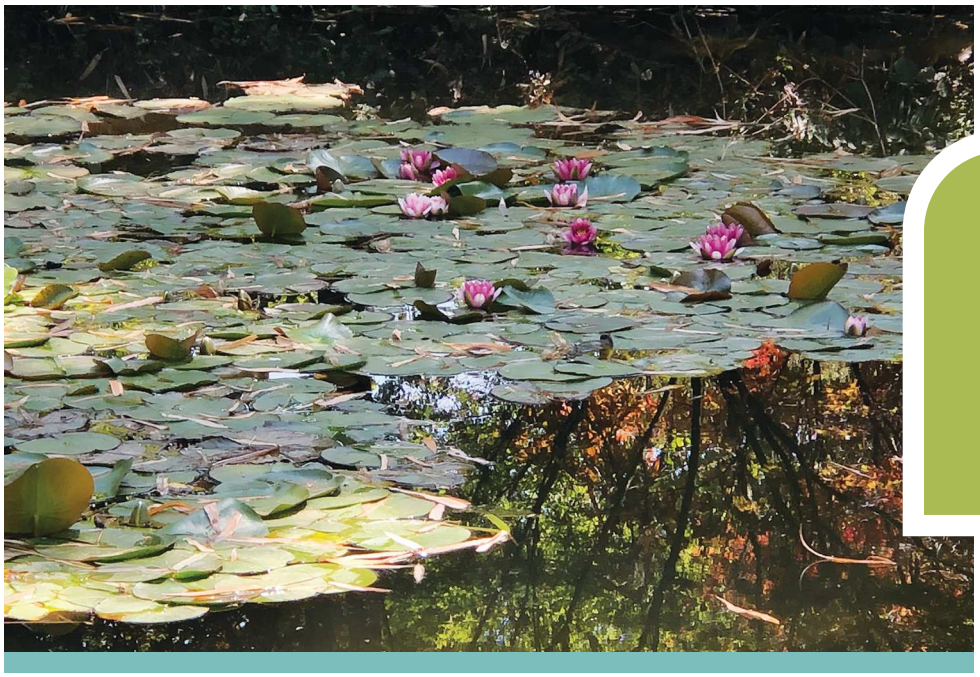
Courtesy of J. Gurevitch


**2** Photosynthesis and Light 21

**3** Water Relations and Thermal Energy Balance 53

**4** Soil and Terrestrial Plant Life 83

**5** Ecosystem Processes 111



Courtesy of J. Gurevitch

2

# Photosynthesis and Light

How does a plant detect and respond to the environment that surrounds it? Why can some plants survive extremes of temperature and drought while others cannot? What enables certain plants to thrive in the deep shade of the understory of a tropical rainforest, and others to succeed only in the sunniest habitats?

The **functional ecology** of plants is concerned with how the molecular biology, genomics, biochemistry, and physiology of individual plants determine their responses to their environments within the structural context of their anatomy and morphology. **Physiological ecology** is concerned specifically with the physiological mechanisms underlying whole-plant responses to the environment. Plant functional ecology is shaped by evolution, and it is fundamental for what happens at population, community, and higher ecological levels. Part I of this book begins with plant functional ecology and then continues with how plant function scales up to ecosystem processes.

Plants must acquire energy and materials for growth, maintenance, and reproduction. They must also limit their losses; for example, if a plant loses too much water, it will wilt and eventually die. Plants must also allocate resources in ways that maximize their chances of contributing offspring to the next generation while simultaneously maximizing their chances of surviving to reproduce. In this chapter and the next three, we examine how plants capture the energy of sunlight and incorporate carbon from the atmosphere in photosynthesis, their adaptations to the light environment, their water relations, and the mineral nutrients they get from the soil. These processes in turn can play out at the larger scope of ecosystem processes. We also will take a look at the structures in which some of these processes take place and some of the biochemistry involved. There is a tremendous amount of work currently on the genetic and hormonal control of photosynthesis that is beyond the scope of this book; interested readers should see reviews by Sujith Puthiyaveetil and John

Above: Water lilies in Yunnan, China.

Allen (2009), Paul Jarvis and Enrique Lopez-Juez (2013), and Norman Hüner and colleagues (2016).

While we focus first on processes occurring at the small scale of a cell, a leaf, or an individual plant, it is important not to lose sight of the forest: Plants have evolved and live in an ecological context. Photosynthesis is usually carried out in natural environments, not in a laboratory. The photosynthetic machinery, and the leaf in which it is housed, are both adapted by natural selection and adjust to the moment-to-moment and longer-term environment in which the individual plant is living. The temperature and the amount of available light, water, and nutrients in the environment determine when and how rapidly a leaf can photosynthesize and the rate at which the plant grows and is likely to survive.

The physical conditions that a plant experiences are determined not only by the physical features of the environment, but also by other living organisms in that habitat. The amount of light available for photosynthesis may be limited by other plants competing for that light. Pathogens and pollutants may reduce the plant's ability to photosynthesize. A plant's ability to capture carbon and energy may be diminished by herbivores eating its leaf tissue. Individual plant processes translate to how vegetation responds to and affects regional and global environments. A plant responds to its environment as an integrated unit, although in textbooks (like this one) we arbitrarily separate the plant's responses into categories for convenience, treating them in different chapters. We begin looking at plants' interactions with their environment by considering the process by which they acquire energy and carbon: photosynthesis.

## 2.1 Photosynthesis Is the Engine of Life on Earth

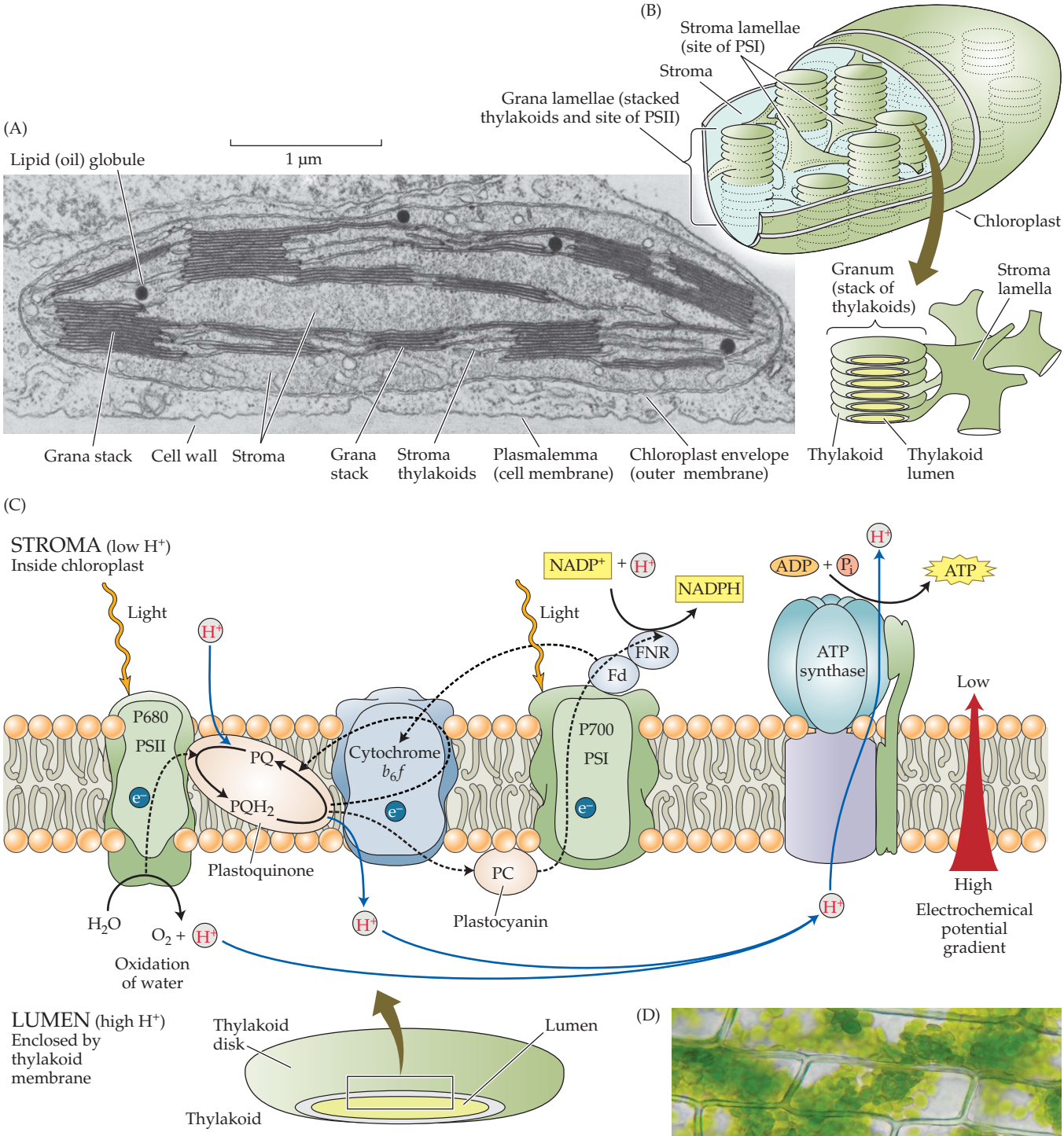
**Photosynthesis** is the set of biochemical processes by which plants acquire energy from sunlight and incorporate it with carbon from the atmosphere into organic compounds. Photosynthesis consists of two distinct parts that take place in different parts of the chloroplast. The first part, the **light reactions**, captures light energy from sunlight and temporarily stores it in high-energy chemical bonds. The second part is photosynthetic **carbon reduction** (called "fixation" in older literature), where that energy together with carbon from $CO_2$ is incorporated into organic compounds. The organic molecules formed in photosynthesis are used by the plant to create new tissues, regulate the plant's metabolic processes, supply energy to those metabolic processes and for many other processes such as reproduction and plant defense. The energy and carbon captured in photosynthesis are the foundation for almost all terrestrial and aquatic food webs, and this is where animals ultimately obtain their energy and the backbone of the molecules in their bodies.

The light reactions of photosynthesis occur on the membranes that make up **thylakoid disks** in the interior of the chloroplasts. The thylakoid disks are piled up in the form of **grana stacks** alternating with sheets of interconnecting membranes called **stroma lamellae** (also known as stroma thylakoids; Figure 2.1A,B). There might be anywhere from fewer than a dozen to about 100 grana stacks in a chloroplast. The successful capture of light energy depends on the precise spatial arrangement of these photochemical reactions on the membranes on which they occur (Figure 2.1C). The architecture of the thylakoid membranes is complex and very specific, and plant scientists have made great progress in understanding their structure and components. The machinery of light capture in photosynthesis takes place inside chloroplasts, which are commonly inside the cells of leaves (Figure 2.1D).

Light absorption in photosynthesis depends on pigments, which are organic molecules that absorb specific wavelengths of light energy. The pigment molecules responsible for the capture of light energy form two distinct molecular complexes in multicellular plants, **photosystem I** and **photosystem II** (Figure 2.1C and Figure 2.2). Unicellular eukaryotic algae, such as those in the Chlorophyta, and prokaryotic cyanobacteria also have photosystems I and II, while other photosynthetic bacteria have only photosystem I. Photosystem II is located mostly on the grana stacks, and photosystem I is mainly on the stroma lamellae.

Photosynthetic rates depend on the light wavelengths—called "light quality"—and not only on the total amount of light (Figure 2.3A,B). Most of the energy in sunlight is in the visible part of the spectrum (Figure 2.3C). Blue and red wavelengths are preferentially captured by the light reactions. Paradoxically, given our image of the beautiful green world, green wavelengths are particularly *ineffective* for photosynthesis. We see nature as green because green light is reflected by plants—"discarded" rather than used. The wavelengths of light that can be used in photosynthesis are termed **photosynthetically active radiation**, or **PAR**. The amount of usable light energy impinging on a leaf per unit time is called the **photosynthetic photon flux density** (**PPFD**).

Each photosystem consists of hundreds of pigment molecules, including several forms of chlorophyll plus accessory pigments, which form an "antenna" for capturing certain wavelengths of light. In terrestrial plants, the accessory pigments are primarily the orange- and yellow-colored carotenoids known as carotenes and xanthophylls. Each antenna has up to 400 molecules of accessory pigments. The accessory pigments can absorb wavelengths of light that chlorophyll molecules are poor at capturing (Figure 2.3D). They then transfer the energy from the wavelengths of light they captured




Juan Carlos Fonseca Mata/CC BY-SA 4.0

**Figure 2.1** Chloroplasts and photosynthetic membranes. (A) Chloroplast from a leaf of *Nicotiana tabacum* (tobacco, Solanaceae), showing grana stacks and stroma thylakoids. (B) A model of grana stacks and stroma thylakoids within a chloroplast. (C) The arrangement of the four main components of the light reactions on the thylakoid membranes of the grana stacks: photosystems I and II (PSI and PSII), cytochrome $b_6f$, and ATP synthase. (D) Chloroplasts (green) in the living cells of a leaf; notice that the leaves are mostly clear—the green color of leaves is due to the reflection of green light by chlorophyll in the chloroplasts. (A from K. Esau. 1977. *Anatomy of Seed Plants*. John Wiley: New York, NY; B after L. Taiz et al. 2018 *Fundamentals of Plant Physiology*, 1st ed. Oxford University Press/Sinauer: Sunderland, MA; C from L. Taiz and E. Zeiger. 1998. *Plant Physiology*, 4th ed. Oxford University Press/Sinauer: Sunderland: MA.)

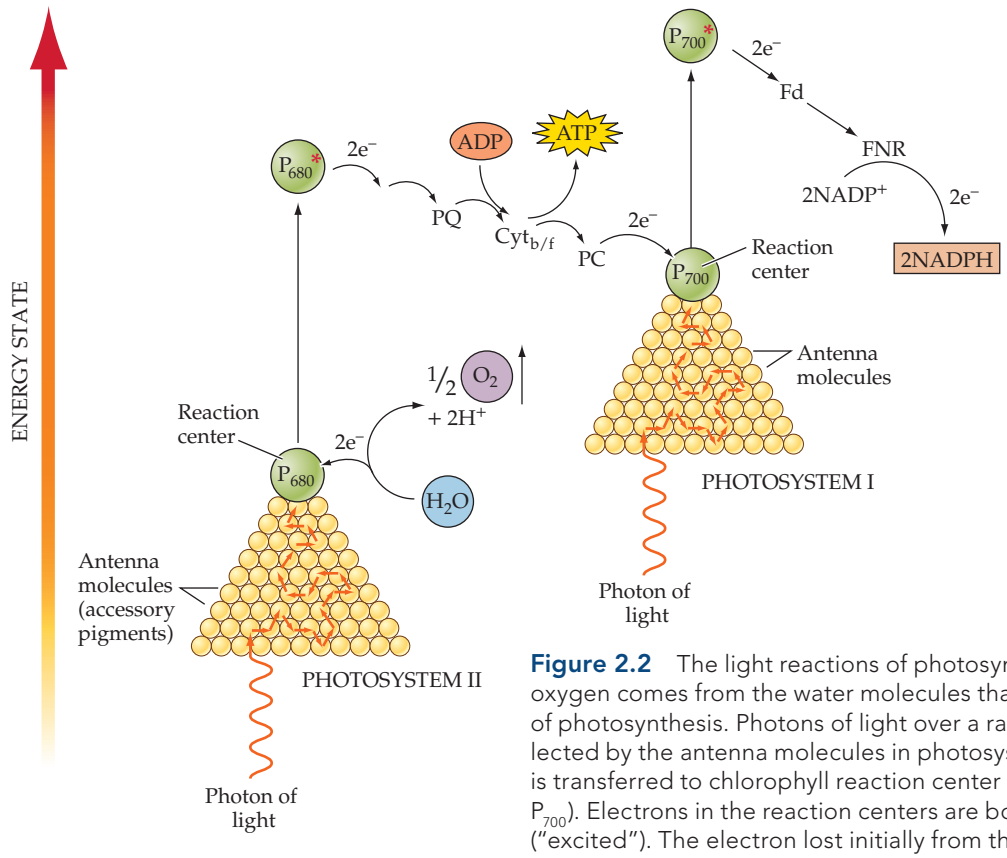



**Figure 2.2** The light reactions of photosynthesis in plants. Atmospheric oxygen comes from the water molecules that are split in the light reactions of photosynthesis. Photons of light over a range of wavelengths are collected by the antenna molecules in photosystems I and II, and the energy is transferred to chlorophyll reaction center molecules (chlorophyll $P_{680}$ and $P_{700}$). Electrons in the reaction centers are boosted to higher energy states ("excited"). The electron lost initially from the reaction center of photosystem II is replaced with an electron removed from a water molecule, which is split in the process. Oxygen ions from two split water molecules are joined to form an oxygen molecule, which is released to the atmosphere, plus hydrogen ions. The reactions in the two photosystems occur essentially simultaneously. (PQ, plastoquinone; $Cyt_{b/f}$, cytochrome $b_6f$; PC, plastocyanin; Fd, ferrodoxin; FNR, ferredoxin–$NADP^+$ reductase.)

to chlorophyll molecules, using a process called resonance transfer. The accessory pigments can also protect leaves from sun damage by absorbing excess light energy when photosynthesis cannot keep up with the amount of light received. Beta-carotene and lycopene are some of the photosynthetic carotenoids that are also important nutrients for people. (Carotenoids are also found in other parts of plants; for instance, they are what make carrots, marigolds, some autumn leaves, and pumpkins orange or yellow.) Eukaryotic algae and photosynthetic bacteria also use other accessory pigments. For example, phycobilins are important photosynthetic pigments in red algae, enabling these eukaryotic organisms to absorb red, orange, yellow, and green wavelengths of light. Those red algae that photosynthesize deeper in the ocean than any other organisms have green-absorbing phycobilins that enable them to capture the wavelengths of light that penetrate deepest in water (to about 100 m).

When a photon of light is captured by the complex of tightly packed antenna molecules on the **thylakoid membranes**, it is passed from one molecule to another by the process of resonance transfer, until it is finally trapped by the chlorophyll molecule at the reaction center. The chlorophyll molecule then becomes "excited"—the molecule is at a higher energy state, and its electrons jump to higher energy orbitals. The excited chlorophyll molecule at the reaction center then passes a light-excited, high-energy electron to an electron acceptor, which passes it to other electron acceptors (see Figure 2.2). The energy in this high-energy electron is ultimately captured in high-energy bonds in ATP and NADPH. The photochemical reactions occur incredibly quickly—the whole process is complete within picoseconds (trillionths, $10^{-12}$, of a second).

These electrons ultimately come from water molecules, which are split when their electrons are donated to photosystem II. The oxygen that we breathe was released into the atmosphere from water molecules in the light reactions of photosynthesis that were split to replace the electrons in photosystem II. Oxygen from photosynthetic light reactions that occurred in photosynthetic bacteria was first released into the atmosphere beginning about 2 billion years ago.

The light reactions provide the energy for incorporating $CO_2$ into organic molecules. That energy is stored as high-energy bonds in ATP and NADPH. These molecules move from the thylakoid membranes into the stroma

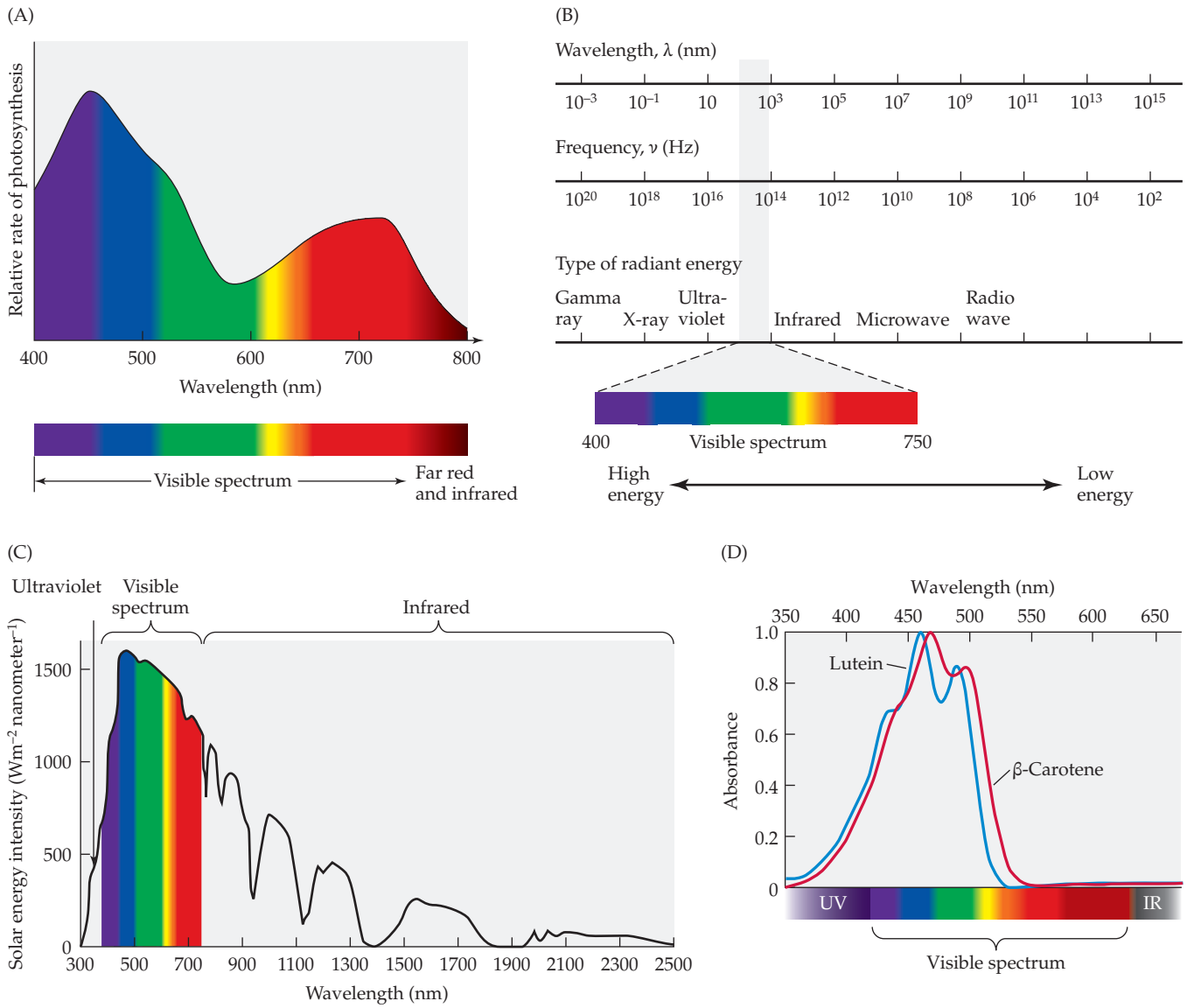



**Figure 2.3** (A) All wavelengths aren't equal: the action spectrum identifies the wavelengths that are actually absorbed and used in photosynthesis. Green wavelengths are reflected and not used in photosynthesis. (B) Visible wavelengths are only a small part of the electromagnetic spectrum. (C) Most of the energy in sunlight reaching Earth's surface is in the visible part of the electromagnetic spectrum. (D) Absorption spectra for two of the most important accessory pigments, lutein and beta-carotene. (B after L. Taiz et al. 2018. *Fundamentals of Plant Physiology*, 1st ed. Oxford University Press/Sinauer: Sunderland, MA; C after ASTM G173-03. U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy, National Renewable Energy Laboratory. https://www.nrel.gov/grid/solar-resource/spectra-am1.5.html; D after A. M. Collins et al. 2011. *PLOS ONE* 6: e24302. CC BY 4.0. https://doi.org/10.1371/journal.pone.0024302).

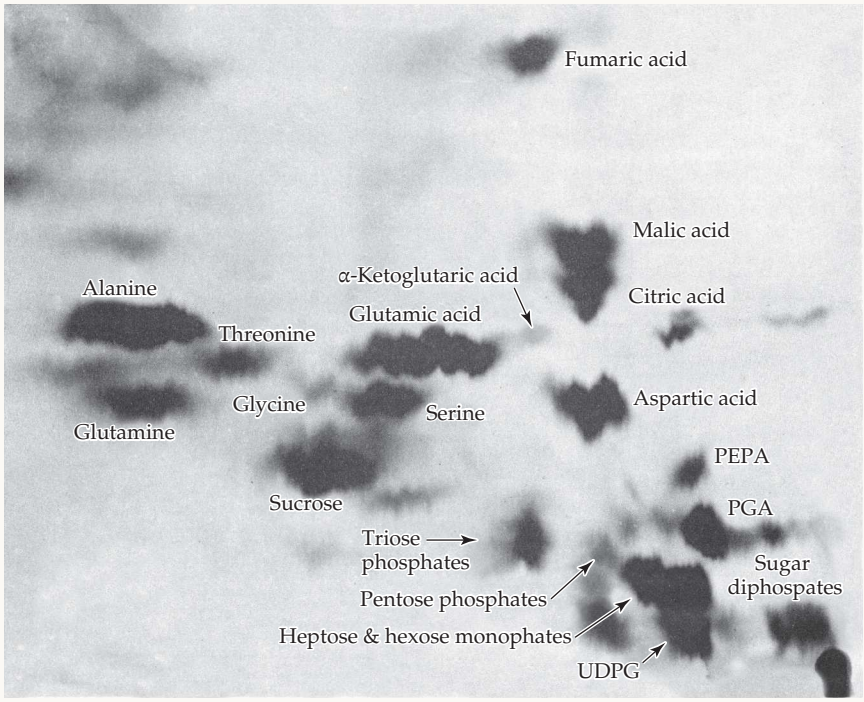
of the chloroplast, where they fuel carbon reduction. Carbon reduction occurs in the biochemical reactions of the **Calvin-Benson cycle** (Figure 2.4; Box 2A), in which $CO_2$ is taken up from the atmosphere and the carbon is incorporated into organic compounds, along with the energy captured in the light reactions. These reactions take place in the **stroma**, the watery matrix that fills the chloroplast. In $C_3$ plants (plants with the most common type of photosynthetic pathway, discussed later in this chapter), large amounts of the enzymes that catalyze the reactions of the Calvin-Benson cycle (sometimes called the Calvin cycle) are dissolved in the chloroplast stroma. The **quantum yield** (or quantum efficiency) of photosynthesis is the number of moles of $CO_2$ fixed per mole of photons absorbed; it depends on the wavelength of the light absorbed as well as other factors.

To sum up photosynthesis: the reduction of carbon is powered by the light reactions, and the light energy

## BOX 2A The Discovery and Elucidation of Photosynthetic Carbon Reduction

The Lawrence Berkeley Laboratory at the University of California, Berkeley, was a physics facility that was central to the development of the atomic bomb and radar during World War II, as well as fundamental research in physics. Because it had the ability to generate and work with radioactive isotopes, after the war the "rad lab" also became the site of research on photosynthesis using radioactivity to decipher the mechanisms by which inorganic carbon was reduced to become organic molecules. Chemists Sam Ruben and Martin Kamen were the first to synthesize radioactive $^{14}C$ in 1940; Sam Ruben was killed in 1943 at age 30 in a horrific lab accident while working on this project (Benson 2002b). Following World War II, Andrew Benson, as a young scientist working in the lab of Melvin Calvin, used radioactive $^{14}C$ to be able to trace the pathway of $CO_2$ incorporation into organic compounds by using $^{14}CO_2$ to see which compounds the radioactive $^{14}C$ was incorporated into. In these experiments, algal cells were exposed to light and air, then the air was flushed with pure $N_2$, and a solution of $^{14}CO_2$ dissolved in water (making labeled bicarbonate) was injected into a flask with the algal cells. After a short time, the cells were dropped into boiling alcohol to kill them, and paper chromatography was used to separate the compounds that had taken up the $^{14}C$. The paper chromatogram was used to expose X-ray film, and the "spots" were subsequently analyzed to identify them (Bassham 2003). Benson discovered that RuBP and $CO_2$ were joined to create the first product of photosynthesis, 3PGA, catalyzed by the enzyme rubisco, in 1949 and 1950 (Benson 2002a; Bassham 2003; Govindjee and Krogmann 2004; Nonomura et al. 2016). He was subsequently dismissed by Calvin from his position at the lab, apparently for persuing this line of research when Calvin had a competing theory he was working on (Benson 2010; Sharkey and Weise 2015). At first there were several hotly competing theories for the mechanisms of photosynthesis, including one promoted by Calvin himself (later disproved), and it was unclear whether the light reactions and carbon reduction occurred together in a single set of reactions or as two distinct but intertwined parts of photosynthesis. James Bassham, starting as a graduate student in the lab, worked first with Benson and later with Calvin to more fully elucidate the pathway of photosynthetic carbon reduction. Calvin received a Nobel Prize for this work in 1961. Andrew Benson eventually obtained a position as a research professor at the University of California at San Diego, where he had a long and distinguished career.

$C_4$ photosynthesis was discovered and uncovered over a period of time, as the biochemistry, anatomy, and other features were slowly pieced




Two minutes after exposure of cells of the green alga *Chlorella pyrenoidosa* (Chlorellaceae) to radioactive $CO_2$ ($^{14}CO_2$), many of the products of photosynthesis show up on this radioautograph. (From J. A. Bassham and M. Kirk. 1960. *Dynamics of the Photosynthesis of Carbon Compounds 1. Carboxylation reactions. UCLR-9033.* Lawrence Radiation Laboratory, University of California: Berkeley, CA.)

From L. Orr and Govindjee. 2013. *Photosynth Res* 115:179

Melvin Calvin and Andrew Benson

**BOX 2A** *(continued)*

together (recounted by Hatch 2002). Hugo Kortschak, a plant physiologist working in Hawaii on photosynthesis in sugarcane (*Saccharum officinarum*, Poaceae, a tropical grass) and also using labeled $^{14}CO_2$, was the first to recognize, in 1954, that there was a novel photosynthetic type operating in this species, but he did not publish his findings in a scientific journal until 1965. A young Russian scientist, Yuri Karpilov, also worked on the biochemistry of $C_4$ photosynthesis in the late 1950s but only published in Russian; that and his early death in a bicycle accident prevented wide recognition of his work. The pathway was finally fully worked out largely by two Australian plant biochemists, Marshall Hatch and Roger Slack, culminating in several major publications in the late 1960s and early 1970s and general recognition that an entirely new type of photosynthesis had been discovered, 20 years after the main mechanisms of $C_3$ photosynthesis were understood. Aspects of CAM were identified slowly over a long period starting in the 1800s, and even the name was developed over time by several different scientists (Black and Osmond 2003), but it began to be more definitively understood (and named CAM) by a Welsh scientist, Meirion Thomas, in 1946. Research on $C_4$ photosynthesis spurred a large group of researchers to plunge into research on CAM, and from the late 1960s through the entire 1970s the anatomy and biochemistry of CAM were finally worked out. While we can only introduce a few of the major names here, in fact many different scientists at different labs all over the world contributed to working to uncover the mechanisms by which plants use sunlight and $CO_2$ from the air to create the compounds that fuel essentially all ecosystems on Earth.

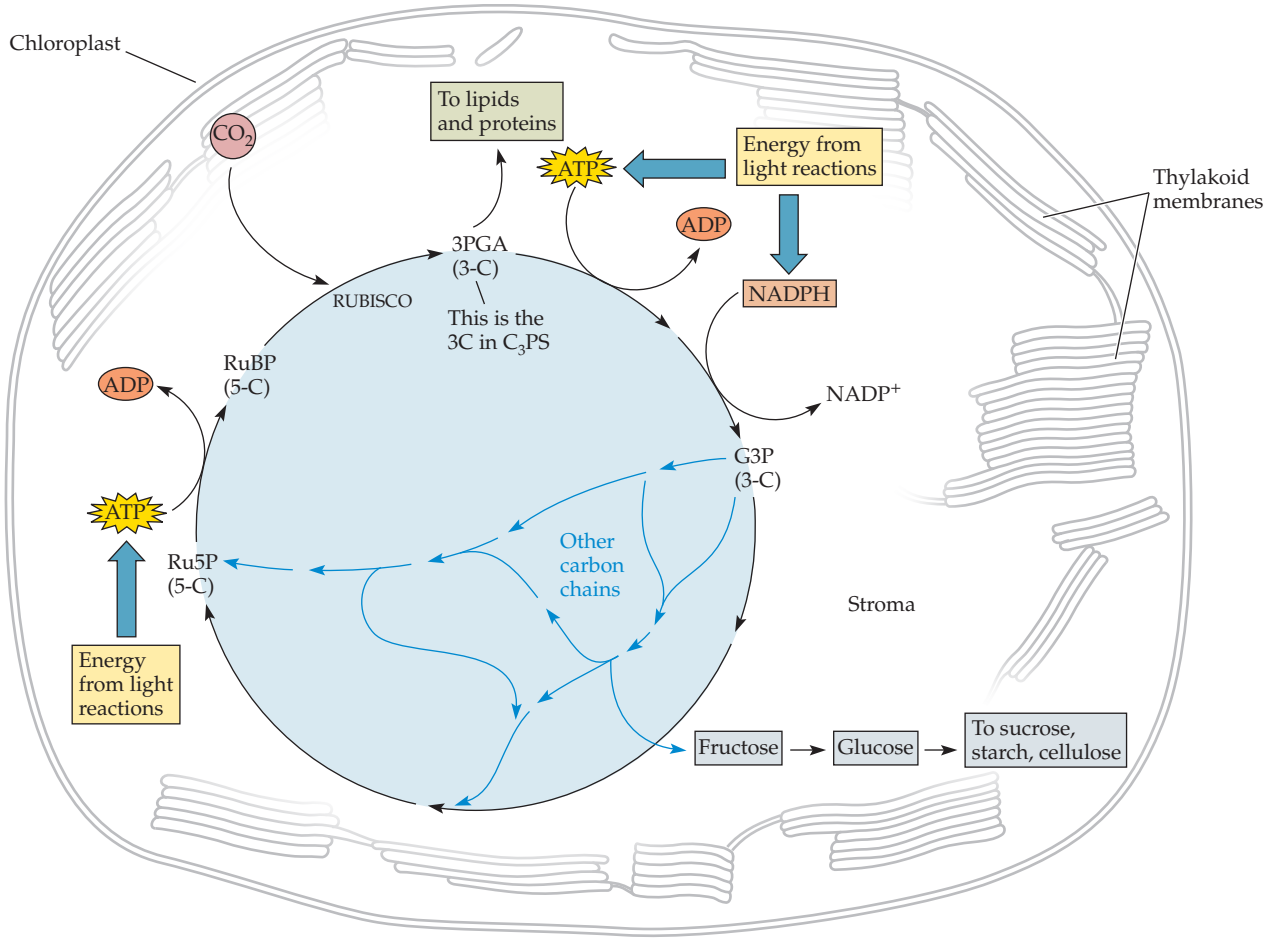



**Figure 2.4** The Calvin-Benson cycle of carbon reduction in plants. $CO_2$ enters the stomata from the air surrounding the leaf. In a reaction catalyzed by the enzyme rubisco, $CO_2$ is joined with the five-carbon molecule RuBP to form two three-carbon molecules (3PGA). Eventually, simple sugars such as fructose and glucose are formed, to be ultimately transformed into many other organic (carbon-based) molecules.

captured in photosynthesis is ultimately stored in the chemical bonds of carbohydrates and other organic molecules.

## 2.2 Photosynthesis Is Affected by the Environment and by Plant Adaptations

The rate at which a leaf can capture light energy and fix carbon is determined by several factors. Plants, like other aerobic organisms, use oxygen and release $CO_2$ in the process of **cellular respiration**, by which organic compounds are broken down to release energy. **Gross photosynthesis**, or the total amount of carbon captured, is reduced by the plant's respiratory release of $CO_2$. Photosynthetic uptake of $CO_2$ by plants is far greater on average than respiratory losses, however, resulting in a net gain of carbon by plants.

### The amount of light available limits photosynthesis

The most basic factor limiting photosynthesis is the total amount of light energy that reaches the thylakoid membranes. In darkness, cellular respiration results in a net loss of carbon and energy from the plant, as there is no photosynthetic capture of either light or carbon (for a partial exception, see the discussion of CAM photosynthesis below). As the light level increases, plants begin to take up $CO_2$. At the **light compensation point**, photosynthetic gains exactly match respiratory losses (in other words, net $CO_2$ exchange is zero) (Figure 2.5). Beyond that point, the more light that is available to be captured, the greater the photosynthetic rate, up to a maximum, at which the rate plateaus in most plants. Too much light can damage the tissues, and the accessory pigments and photorespiration (below) can be important in protecting the leaf from this destructive excess energy.

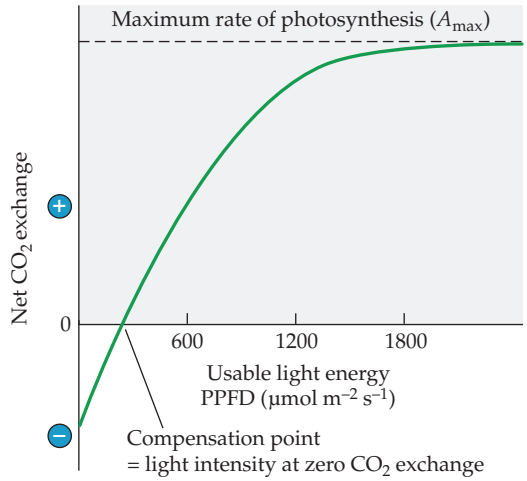



**Figure 2.5** Net $CO_2$ exchange (per unit leaf area) for a typical $C_3$ leaf as a function of increasing light levels, showing the light compensation point and a plateau at a maximum rate of photosynthetic carbon assimilation, $A_{max}$. (After A. H. Fitter and R. K. M. Hay. 1981. *Environmental Physiology of Plants*, 3rd ed. Academic Press: London, U.K.)

The light compensation point can differ among plant species living in different regions or even within a single habitat or within individual plants, depending on the structure and biochemical constituents of the leaves and on the light environment and season. David Rothstein and Donald Zak (2001) contrasted the photosynthetic characteristics of three forest floor herb species within a northern hardwood deciduous forest. Light levels are high in the understory under trees in early spring before the trees produce their leaves, low in midsummer, and higher again in autumn as leaves start to fall. A **spring ephemeral** (active above ground only for a short time in spring), *Allium tricoccum* (wild leek, Liliaceae), had a constant light compensation point (Table 2.1) but was photosynthetically active only during a short period in spring when light levels were high. In

**TABLE 2.1** Maximum photosynthetic rates ($A_{max}$), light compensation points (LCP), and rubisco levels for three forest understory herbs

| | Spring | | | Summer | | Autumn |
|---|---|---|---|---|---|---|
| Parameter | *Allium* | *Viola* | *Tiarella* | *Viola* | *Tiarella* | *Tiarella* |
| $A_{max}$ | 15.4 ± 0.9 | 12.1 ± 0.7 | 6.8 ± 0.7 | 5.6 ± 0.5 | 3.9 ± 0.5 | 5.4 ± 0.3 |
| LCP | 21.6 ± 1.4 | 8.4 ± 1.3 | 9.0 ± 1.0 | 4.1 ± 0.9 | 3.2 ± 0.5 | 6.5 ± 0.8 |
| Rubisco | 2.83 ± 0.21 | 1.84 ± 0.25 | 1.47 ± 0.12 | 0.93 ± 0.07 | 0.50 ± 0.17 | 0.78 ± 0.11 |

*Source*: D. E. Rothstein and D. R. Zak. 2001. *Func Ecol* 15: 722–731.

*Note 1*: Values are expressed on a per unit leaf area basis. $A_{max}$ is given in μmol $CO_2$/m$^2$/s; LCP is given as the PPFD at which net $CO_2$ assimilation is zero, in μmol/m$^2$/s; and rubisco levels are in g/m$^2$. Values are means ± 1 standard error, with $n$ = 5 plants per measurement.

*Note 2*: The duration during which each species had green leaves above ground was: *Allium tricoccum*, about 75 days; *Viola pubescens*, about 150 days; and *Tiarella cordifolia*, about 185 days.

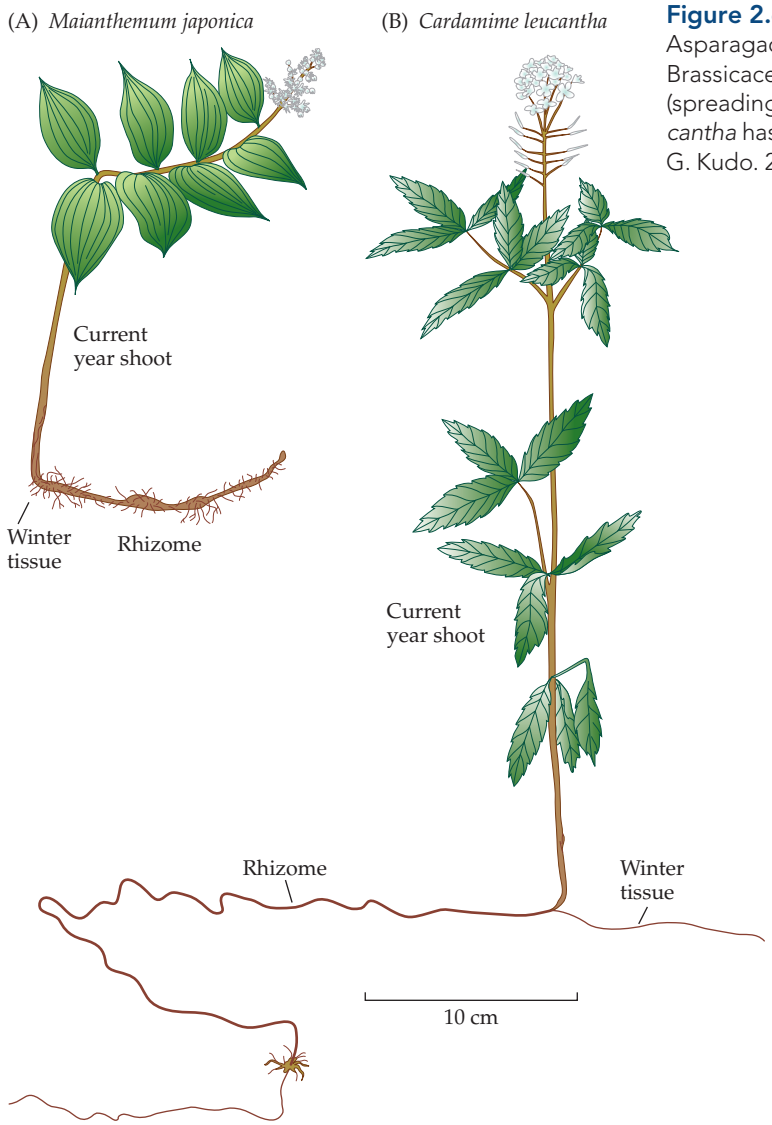



**Figure 2.6** (A) *Maianthemem japonicum* (false Solomon's seal, Asparagaceae) and (B) *Cardamine leucantha* (Korean bittercress, Brassicaceae). *M. japonicum* has thick underground rhizomes (spreading underground stems) that store nutrients, while *C. leucantha* has little underground storage. (After T. Y. Ida and G. Kudo. 2009. *J Plant Res* 122: 171–181.)

contrast, a summer-green plant, *Viola pubescens* (downy yellow violet, Violaceae), shifted its light compensation point downward from spring to midsummer, while a semievergreen species, *Tiarella cordifolia* (foamflower, Saxifragaceae), also shifted its light compensation point downward over that period, but it shifted upward again in autumn. The spring ephemeral appears to be adapted to optimize its photosynthetic uptake in the high-light environment it experiences in spring, while the other two species are both better adapted for photosynthesis under shady conditions, at least in part due to their ability to shift the light compensation point.

Similar kinds of adaptations may occur in similar but geographically distant environments. Takashi Ida and Gaku Kudo (2009) studied photosynthetic rates and carbon allocation in two perennial herbaceous plants, *Cardamine leucantha* (bittercress, Brassicaceae) and *Maianthemum japonicum* (false Solomon's seal, Asparagaceae; called *Smilacina japonica* in the original publication), living in the understory of a deciduous forest in Hokkaido, northern Japan (Figure 2.6). In spring, light is high in the forest understory before the canopy trees leaf out, and both species had similar maximum photosynthetic rates. In summer, when the canopy trees were in full leaf, the maximum photosynthetic rates of both species were approximately half of their spring photosynthetic rates. However, in light gaps in the tree canopy in summer, the two species were very different, with *C. leucantha* having almost the same high rates as in spring, while *M. japonicum* had the same low photosynthetic rates in the canopy shade and in light gaps. This difference is linked to a difference in the turnover of leaves in the two species. *C. leucantha* has rapid leaf turnover, so new leaves with higher photosynthetic rates were continually being produced, while *M. japonicum* has limited leaf production in summer, and photosynthetic rates decreased as the leaves got older. Instead of investing in new leaves, *M. japonicum* allocates much of the carbon gained in photosynthesis to storage tissue in the rhizome, which is important for its longer-term survival.

The quantity of light reaching the thylakoid membranes of a chloroplast can be limited by a number of factors. The location of the chloroplast within the leaf can affect the light reaching the thylakoid membranes, as can the angle at which sunlight hits the leaf. In a typical $C_3$ leaf, photosynthesis takes place in the spongy and palisade parenchyma cells that make up the **mesophyll** (the photosynthetic tissue between the upper and lower epidermis of a leaf) (Figure 2.7). There are many chloroplasts in each photosynthetic cell. On a larger scale, self-shading by other leaves on the same plant, or shading by competitors, can also limit the amount of light available to be captured. We will examine some of these factors in more detail in other chapters.

Robin Chazdon (1985) studied the efficiency of light capture in two understory dwarf palms in the rainforests of Costa Rica, *Asterogyne martiana* and *Geonoma cuneata* (both in the Arecaceae). Both species have narrow, spirally arranged leaves that minimize self-shading.

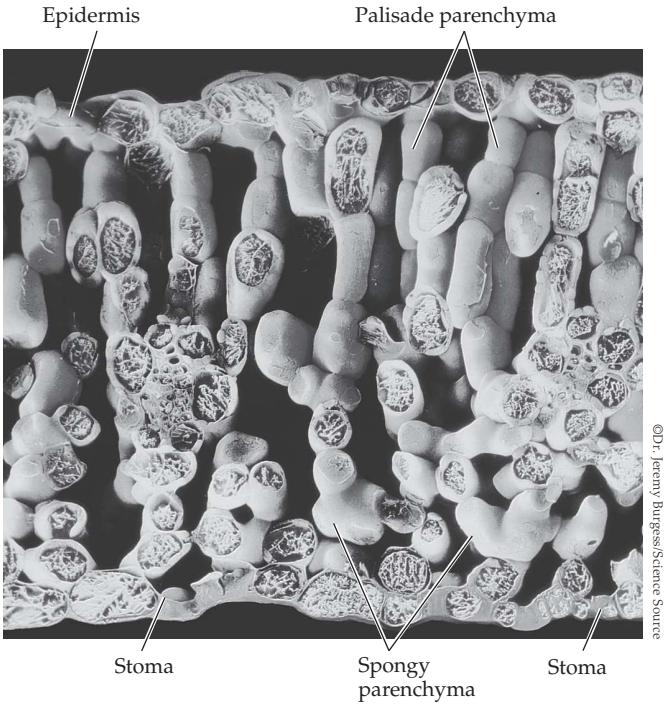



**Figure 2.7** Scanning electron micrograph of a cross section of a leaf of *Brassica septiceps* (turnip, Brassicaceae), showing the palisade parenchyma and spongy parenchyma cells inside which most chloroplasts are found and in which most of the plant's photosynthesis takes place. Many of the cells in this micrograph have been broken open to expose their internal structure. The upper epidermis is visible, as are several stomata on the underside of the leaf, along with the substomatal cavities they open into on the inside of the leaf.

*A. martiana* was found in locations with somewhat higher light levels, and had a greater number of leaves and a greater total leaf area, than *G. cuneata*. As a result, *G. cuneata* had greater efficiency of light interception (the proportion of incident light intercepted by the plant canopy, which depends on leaf arrangement and display angle), but *A. martiana* had a greater total capacity to capture light (where the light interception capacity, or effective leaf area, is the product of total leaf area and light interception efficiency). Akio Takenaka and associates (2001) analyzed the effects of leaf display on light capture efficiency in another understory palm, *Licuala arbuscula* (Arecaceae), which grows in lowland rainforests in Southeast Asia. This species has compound, fan-shaped leaves with long petioles. The authors found that the angle at which the petioles are held changed as the number of leaves increased. As plants grew from juveniles with few leaves to mature plants with many leaves, this shift reduced self-shading to a minimal level and optimized light capture for individuals of very different forms and total leaf areas.

Leaves are not the only plant organs that can contain chloroplasts and carry out photosynthesis, although most photosynthesis most of the time in most terrestrial plants takes place in leaves. Bark, stems, the skin of ripe or unripe fruit, and the sepals covering flower buds that have not yet opened may be important for photosynthesis under some conditions and for providing energy for developing structures (Figure 2.8).

Differences in species' adaptations to different light levels was the focus of an experimental study in Hawaii Volcanos National Park by Jennifer Funk and Sierra McDaniel (2010), where invasive, non-native grasses had become dominant. Experimental shading was used to mimic a canopy understory and to investigate shading as a restoration technique for the native woody species. Seedlings of native woody shrubs and trees had lower rates of photosynthesis in both sun and shade compared with the grass seedlings. In full sun, the biomass of the invasive grasses was much higher than that of the woody species, but the grass biomass was reduced much more than that of the native woody plants in the shade. The invasive grasses also had greater reductions in survival in experimental shade than the native woody species did. The results suggested that creating shade might help with restoration of native species in these degraded Hawaiian landscapes.

The light environment also varies on a global scale. Day and night are close to being equally long at tropical latitudes, and this pattern is the same all year, while at polar latitudes it is continuously light at midsummer and continuously dark at midwinter (see Chapter 16). Maximum daily PPFD is greater in the tropics than in polar regions, and greater at high altitudes than at sea level (see Figure 16.5). Certain other parts of the solar spectrum vary across Earth's surface to a much greater degree than does PPFD. In particular, ultraviolet B (UVB) radiation, which is damaging to plants and dangerous as well to other organisms, including people, is up to ten times as great at high elevations in the tropics than at low elevations in Arctic environments. The ozone layer in the stratosphere absorbs UV radiation (see Box 16C). UV radiation passes through much more ozone before reaching the ground in the Arctic than in equatorial regions, because the path of solar energy through the atmosphere to Earth's surface is much longer in the Arctic than in the tropics (see Chapter 16). However, several types of man-made chemicals began to destroy the stratospheric ozone layer that protects organisms at the surface of the Earth from UV radiation, creating an "ozone hole" that was particularly severe in the Southern Hemisphere. Global cooperation in reducing these chemicals has been

effective in considerably reducing this danger. While these problems still exist and this requires continued efforts and commitment, as of the time of writing this, the ozone hole is in the process of healing.

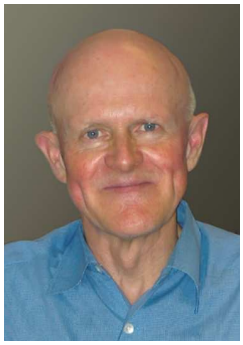


Photo by Donna Chambers. Courtesy of Martyn Caldwell

Martyn Caldwell

Plants have numerous biochemical adaptations to high UVB radiation (Searles et al. 2001). To avoid UVB damage, plants increase the leaf concentration of compounds that absorb UVB radiation, primarily flavonoids. They also limit damage by increasing concentrations of antioxidant enzymes and DNA repair enzymes. Martyn Caldwell (1968) found that high concentrations of these compounds were particularly common in plants that grow in high alpine environments.

## Carbon uptake is limited by the ways plants respond to their environments

Plants take up $CO_2$ from the atmosphere as air moves through the stomata and into the intercellular spaces surrounding the photosynthetic cells within a leaf. Carbon uptake is driven by a concentration gradient of $CO_2$, set up by the biochemical reactions in the chloroplasts that remove $CO_2$ from the intercellular spaces. The uptake of $CO_2$ is regulated by conductance to $CO_2$ diffusion on the pathway from the surrounding air into the leaf and into the chloroplast. The concentration of $CO_2$ in the intercellular spaces depends on how rapidly $CO_2$ is removed by being fixed in organic compounds, and on how readily $CO_2$ comes into the leaf to replace that $CO_2$.

(A)

Grendelkhan/CC BY-SA 3.0

Courtesy of J. Gurevitch

(B)

William Herron/CC BY-SA 2.0

(C)

Courtesy of J. Gurevitch

**Figure 2.8** (A) Green tissues in structures like fruits (unripe oranges, chiles), (B) bark (a *palo verde* tree with a bird's nest in the Sonoran Desert), and (C) sepals (surrounding the bud of a peony flower) may also contain chloroplasts and carry out photosynthesis.

The **leaf conductance** to $CO_2$ is the rate at which $CO_2$ flows into the leaf at a given concentration difference between ambient and intercellular $CO_2$. The inverse of conductance is resistance. Low conductance or high resistance at a particular point in the pathway of $CO_2$ will limit its movement along that pathway. If the overall leaf conductance to $CO_2$ is high and $CO_2$ concentrations in the intercellular spaces are being continually drawn down by the rapid reduction of carbon, then $CO_2$ influx from the air surrounding the leaf will be high.

The rate of $CO_2$ uptake can be modeled with a flux equation. **Flux equations** are used to model flow rates and are of the general form

$$\text{flux} = (\text{conductance}) \times (\text{driving force})$$

For $CO_2$ uptake, the driving force is a difference in $CO_2$ concentration, and the flux equation can be stated as

$$CO_2 \text{ uptake rate} = (\text{leaf conductance to } CO_2 \text{ diffusion}) \times (\text{difference in } CO_2 \text{ concentration between air and chloroplast})$$

or, using conventional symbols,

$$A = g_{leaf} \times (C_a - C_i)$$

The term $A$ is the **assimilation rate** (in $\mu mol/m^2/s$); this is the rate at which $CO_2$ is taken up by the leaf. The terms $C_a$ and $C_i$ are the ambient and intercellular concentrations of $CO_2$, respectively, that is, its concentrations in the surrounding air and at the surface of the photosynthetic cell. The term $g_{leaf}$ is the total conductance of the leaf to $CO_2$.

We can separate leaf conductance into its two major components, $g_a$ and $g_s$—the conductances to $CO_2$ through the boundary layer of air surrounding the leaf and through the stomata, respectively—and then

$$\frac{1}{g_{leaf}} = \frac{1}{g_a} + \frac{1}{g_s}$$

Generally, $g_a$ is large, since $CO_2$ readily passes through the boundary layer and so does not contribute much to regulating $CO_2$ flux. The conductance to $CO_2$ through the stomata ($g_s$), however, is highly variable and is under the control of the plant. Stomatal conductance regulates leaf $CO_2$ flux under most conditions. Thus, plants do not act as merely passive recipients of $CO_2$, but regulate its uptake closely. This regulation occurs over short time scales (seconds to minutes), as stomata are opened or closed, and over longer time scales (days to months), as leaf morphology and chemistry are altered. Over even longer time scales (years to millennia), natural selection acts to alter the capacity of plant populations in different environments to take up carbon under different conditions as morphology, physiology, and other plant characters evolve (see Chapter 9).

Why would plants ever restrict their uptake of $CO_2$? We examine this question in more detail in Chapter 3, but briefly, mostly it is because $CO_2$ gain is linked inextricably with loss of water through the same stomatal openings in the leaf where $CO_2$ is taken up.

A different formulation for photosynthetic rate is sometimes employed to describe net photosynthesis at **light saturation**, $A_{sat}$, the light level at which the maximum photosynthetic rate is reached, when stomata are wide open and $CO_2$ uptake is not limited by stomatal conductance:

$$A_{sat} = g_m \times (C_i - C_c)$$

where $C_c$ is the compensation point for $CO_2$ (the $CO_2$ concentration at which net photosynthesis is zero), $C_i$ is as defined above, and $g_m$ is the **mesophyll conductance** or **intracellular conductance**, the conductance to $CO_2$ through the leaf mesophyll cells and cell walls.

An enormous amount of air must be processed by the leaf in the course of photosynthesis. To make a single gram of the carbohydrate glucose, a plant needs 1.47 g of $CO_2$, which is the amount in about 2500 L of air. Looked at another way, the air needed to fill a structure the size of the current largest sports stadium in the world, the North Korean Rungrado 1st of May Stadium located in Pyongyang, could supply enough $CO_2$ to synthesize about 5000 kg of glucose (if there were no fans present breathing out $CO_2$). (In contrast, the Astrodome, the world's first superdome, had enough air to synthesize 600 kg of glucose.)

When the stomata of a leaf are fully open, its conductance to $CO_2$ is generally high. The exact value depends on the number and size of the stomata, and it varies among species, individual plants, and even leaves on the same plant. (We will return to this issue in Chapter 3.) When the stomata are closed, leaf conductance to $CO_2$ approaches zero, although sometimes small amounts of $CO_2$ may "leak" through the cuticle.

Stomata are often very dynamic. The guard cells that determine the degree of stomatal opening are continually in motion, widening and narrowing the stomatal pores to regulate $CO_2$ entering the leaf and water leaving it. Some of the stomata may begin to close while others remain open (Figure 2.9). Such patchy stomatal closure may be more common when plants are experiencing stress (Beyschlag and Eckstein 2001). The guard cells are under a complex set of controls that respond to both internal and external factors.

### Photosynthetic rates can vary among species in different habitats

Plant physiological ecologists have been able to study photosynthetic gas exchange and other physiological processes in natural environments, leading to considerable progress in our understanding of how these

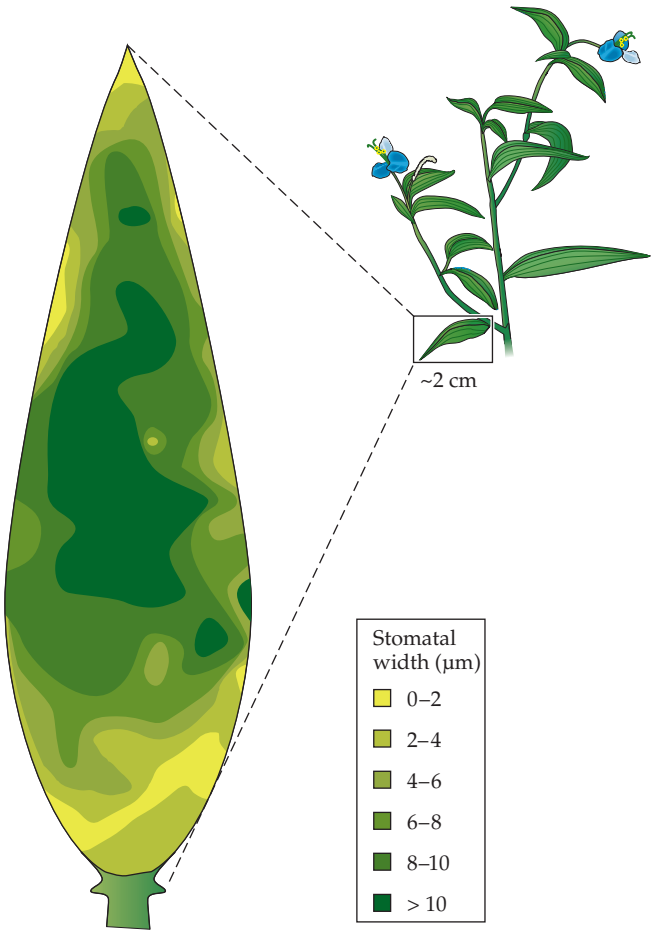



**Figure 2.9** Stomatal widths in different parts of a leaf of *Commelina communis* (dayflower, Commelinaceae) at midday. Some stomata are wide open, while others are partially open or fully closed. (After W. Larcher. 1995. *Physiological Plant Ecology*, 3rd ed. Springer: Berlin; S. Smith et al. 1989. *Plant Cell Environ* 12: 653–659.)

processes function in nature (see Box 3A). Photosynthetic rates sometimes vary among plants within a habitat, and across habitats, in ways that seem to make sense because they are correlated with species composition, habitat preferences, or growth rates. In other cases, photosynthetic rates may have little role in determining population processes or species distributions. Even growth rates may be minimally related to photosynthetic rates. The total carbon accumulated by a plant depends not only on the rate of photosynthesis on a leaf area basis, but also on the total leaf area of the plant, as well as on other factors, such as the length of time the leaves are maintained and are photosynthetically active.

In the study of northern forest understory species discussed above (Rothstein and Zak 2001), maximum photosynthetic rates were correlated with the growth environments of the three species studied (see Table 2.1). The spring ephemeral, *Allium tricoccum* (wild leek, Amaryllidaceae), which grew only in the highest-light period, had the highest maximum photosynthetic rates overall. During spring, the summer-green *Viola pubescens* (downy yellow violet, Violaceae) had intermediate photosynthetic rates, and the evergreen *Tiarella cordifolia* (foamflower, Saxifragaceae) had the lowest rates. In midsummer, when light levels were lowest, photosynthetic rates declined substantially for both *V. pubescens* and *T. cordifolia*, but *T. cordifolia* still had the lower rate of photosynthesis. In autumn, only *T. cordifolia* was photosynthetically active, and its maximum photosynthetic rate increased again in the higher-light environment. These differences among species and seasons were positively correlated with leaf levels of rubisco (the enzyme that catalyzes the initial capture of $CO_2$; see below), as well as patterns of plant growth and the duration of time that each plant was photosynthetic. The spring ephemeral, *A. tricoccum*, gained all of its biomass during the spring high-light period and lost biomass after that time. The other two species also gained biomass rapidly during spring but continued to increase in biomass during summer. *V. pubescens* sharply declined in biomass from late summer through winter, while *T. cordifolia* continued to accumulate biomass through early winter. Only 25% of the biomass gain for *T. cordifolia* occurred during the low-light period of summer.

## 2.3 There Are Three Photosynthetic Pathways: $C_3$, $C_4$, and CAM

Plants fix carbon using one of three different photosynthetic pathways: $C_3$, $C_4$, or CAM (crassulacean acid metabolism; see below). $C_3$ photosynthesis and $C_4$ photosynthesis are named for the three-carbon and four-carbon molecules that are the first stable products of photosynthesis in these pathways, while CAM is named after the plant family Crassulaceae (the stonecrops), in which it was first discovered. The vast majority of plants use $C_3$ photosynthesis, and $C_3$ plants are found everywhere that plants exist. $C_3$ photosynthesis was the first pathway to evolve and the first to be understood by scientists. $C_4$ and CAM photosynthesis are modifications of $C_3$ photosynthesis and evolved from it.

### $C_3$ photosynthesis is the most common and original type of photosynthesis

$C_3$ photosynthesis is found in the largest number of plant species, and $C_3$ plants dominate many parts of the Earth, from the oceans' phytoplankton to the vast northern coniferous forests and tropical rainforests. In the Calvin-Benson cycle of **$C_3$ photosynthesis** (see Figure 2.4), $CO_2$ is joined with a five-carbon molecule, RuBP (ribulose bisphosphate), to form a six-carbon compound

that instantly separates into two three-carbon molecules (phosphoglycerate; 3PGA). In $C_3$ photosynthesis, therefore, the first stable product of carbon reduction is a three-carbon chain.

The initial step in which $CO_2$ is captured—the **carboxylation** of RuBP—is catalyzed by the enzyme RuBP carboxylase/oxygenase, which is mercifully nicknamed **rubisco**. Rubisco is probably the most abundant protein on Earth but is curiously inept at capturing $CO_2$. This is particularly strange considering how important this task is for primary productivity on Earth—one might have expected a more efficient process to have evolved and replaced it in plants long ago. Not only does rubisco have a relatively low affinity for $CO_2$, it also has an alternative function that competes with its role in capturing $CO_2$. Besides catalyzing the initial step of photosynthesis, rubisco can also catalyze a process called **photorespiration**, in which oxygen is taken up instead of carbon dioxide (see Box 2B). At higher temperatures, rubisco increasingly favors the oxygenation reaction over carboxylation, or photorespiration over photosynthesis. Likewise, the higher the concentration of $O_2$ and the lower the concentration of $CO_2$ reaching the chloroplast, the more $O_2$ is taken up in preference to $CO_2$. These properties of rubisco limit photosynthetic $CO_2$ uptake.

The limitations of rubisco are not especially important for plants whose leaves are shaded, because in their case photosynthesis is limited mainly by light levels, rather than by the efficiency of $CO_2$ uptake. However, for plants growing in warm, bright environments, the limitations posed by the properties of rubisco can have major ramifications for photosynthetic rates, and ultimately for growth. Even under the best conditions, $C_3$ plants must maintain large quantities of rubisco to support adequate rates of photosynthesis. Rubisco, like all enzymes, contains a substantial amount of nitrogen. Between 10% and 30% of the total nitrogen in the leaves of $C_3$ plants is in rubisco.

Because of the limitations of rubisco, photosynthetic rates are also limited by the concentration of $CO_2$ in the atmosphere. Consequently, at elevated $CO_2$ concentrations, $C_3$ plants can achieve higher photosynthetic rates, all else being equal. Plant growers sometimes make use of this response by growing plants in greenhouses with artificially high concentrations of $CO_2$ in the air. Plants evolved under atmospheric $CO_2$ levels very different from those of today, as we will see shortly. The current rapid increases in atmospheric concentrations of $CO_2$ caused by human activities may have long-term consequences for $CO_2$ uptake by plants. We return to this issue in Chapter 16.

## BOX 2B Photorespiration

Plants have mitochondria, of course, which carry out cellular respiration much as those of animals do, consuming $O_2$ and releasing energy to be used by the cells. Plants also carry out another kind of respiration, called photorespiration. Like ordinary cellular respiration, photorespiration consumes $O_2$ and releases $CO_2$, but unlike cellular respiration, it depends on light. It takes place in cells that contain chloroplasts, but it involves two additional organelles: mitochondria and peroxisomes.

The enzyme rubisco catalyzes the initial capture of $CO_2$ in the Calvin-Benson cycle, but rubisco also has another, competing function. Rubisco also catalyzes the binding of $O_2$ to ribulose bisphosphate (RuBP) in the process of photorespiration. Photorespiration competes with photosynthesis not only for rubisco, but also for RuBP, the substrate of both reactions. Photorespiration results in large losses of previously captured $CO_2$ to the atmosphere, making photosynthesis much less efficient (without releasing usable energy as ordinary cellular respiration does). Conditions that favor photorespiration in place of photosynthesis are low $CO_2$ concentrations, high partial pressures of oxygen, and warm temperatures. $C_4$ photosynthesis is highly efficient because it overcomes all of these factors by concentrating $CO_2$ and separating the Calvin-Benson cycle in the bundle sheath cells away from atmospheric $O_2$. Another mechanism partially reduces photorespiratory carbon loss: $C_2$ photosynthesis. $C_2$ photosynthesis has been detected so far in only a small number of plant species. In $C_2$ photosynthesis, $CO_2$ that has gone through photorespiration is trapped and refixed in the Calvin-Benson cycle instead of diffusing out to the atmosphere. This is accomplished by spatially separating the oxygenation of rubisco (in chloroplasts in the mesophyll cells) from glycine decarboxylation (in mitochondria in the bundle sheath cells). The $CO_2$ released in these internal bundle sheath cells builds up to high levels, allowing recapture in the Calvin-Benson cycle in chloroplasts in the bundle sheath cells. In a few cases (such as in the genus *Flaveria*), $C_2$ photosynthesis appears to be a precursor to the evolution of $C_4$ photosynthesis.

Although photorespiration is often considered to be disadvantageous because it competes with photosynthesis, it may have a protective function. Photorespiration may "soak up" excessive electron flow in bright light, thereby protecting photosystem II from damage when the leaf's carboxylation capacity is not capable of keeping up with the energy captured in the light reactions (e.g., when drought forces stomata to close partially or fully, limiting or cutting off the supply of $CO_2$).

There is an evolutionary "solution" to the dilemma posed by photorespiration and the limitations of rubisco as a catalyst for $CO_2$ uptake. That solution is $C_4$ photosynthesis.

## $C_4$ photosynthesis is a specialized adaptation for rapid carbon uptake in warm, bright environments

$C_4$ photosynthesis and CAM are specializations that evolved from $C_3$ ancestors. Like $C_3$ photosynthesis, **$C_4$ photosynthesis** ultimately depends on the Calvin-Benson cycle to convert $CO_2$ into carbohydrates. However, $C_4$ photosynthesis contains an additional step that is used for the initial capture of $CO_2$ from the atmosphere (Figure 2.10). In this additional step, a three-carbon molecule called PEP (phosphoenolpyruvate) is joined with $CO_2$ to form a four-carbon acid, OAA (oxaloacetate). The first product of carbon reduction in $C_4$ photosynthesis is a molecule with four carbons. The initial capture of $CO_2$ is catalyzed by the enzyme **PEP carboxylase**, which functions only to fix $CO_2$. It has a much higher affinity for $CO_2$ than does rubisco. Because it does not also catalyze photorespiration, PEP carboxylase can maintain high rates of $CO_2$ uptake even at warm temperatures as long as there is enough sunlight energy for carbon capture.

After its formation, the four-carbon molecule is decarboxylated (the $CO_2$ is removed), and the $CO_2$ is then incorporated into organic molecules via the Calvin-Benson cycle. Rubisco functions to fix this internally liberated $CO_2$ molecule in $C_4$ plants, just as it acts to fix $CO_2$ coming in from the external atmosphere in $C_3$ plants. There are three different subtypes of $C_4$ photosynthesis, each with its own enzyme for decarboxylation: NADP-ME, which uses NADP-malic enzyme; NAD-ME, which uses NAD-malic enzyme; and PEPCK, which depends on PEP carboxykinase.

$C_4$ photosynthesis depends on specialized leaf anatomy (Figure 2.11). In the typical Kranz (German for "wreath") anatomy found in $C_4$ plants, there is a spatial separation of the $C_4$ and $C_3$ reactions. The initial capture of $CO_2$ from the atmosphere takes place in the mesophyll cells just under the epidermis and adjacent to the substomatal cavities, while the incorporation of $CO_2$ into carbohydrates via the Calvin-Benson cycle takes place deep inside the leaf in the bundle sheath cells. In $C_3$ plants, the concentration of oxygen in the chloroplasts is typically about 1000 times greater than the concentration of carbon dioxide, resulting in substantial rates of photorespiration. In $C_4$ plants, rubisco is located (along with the other enzymes of the Calvin-Benson cycle) in the bundle sheath cells, which are not exposed directly

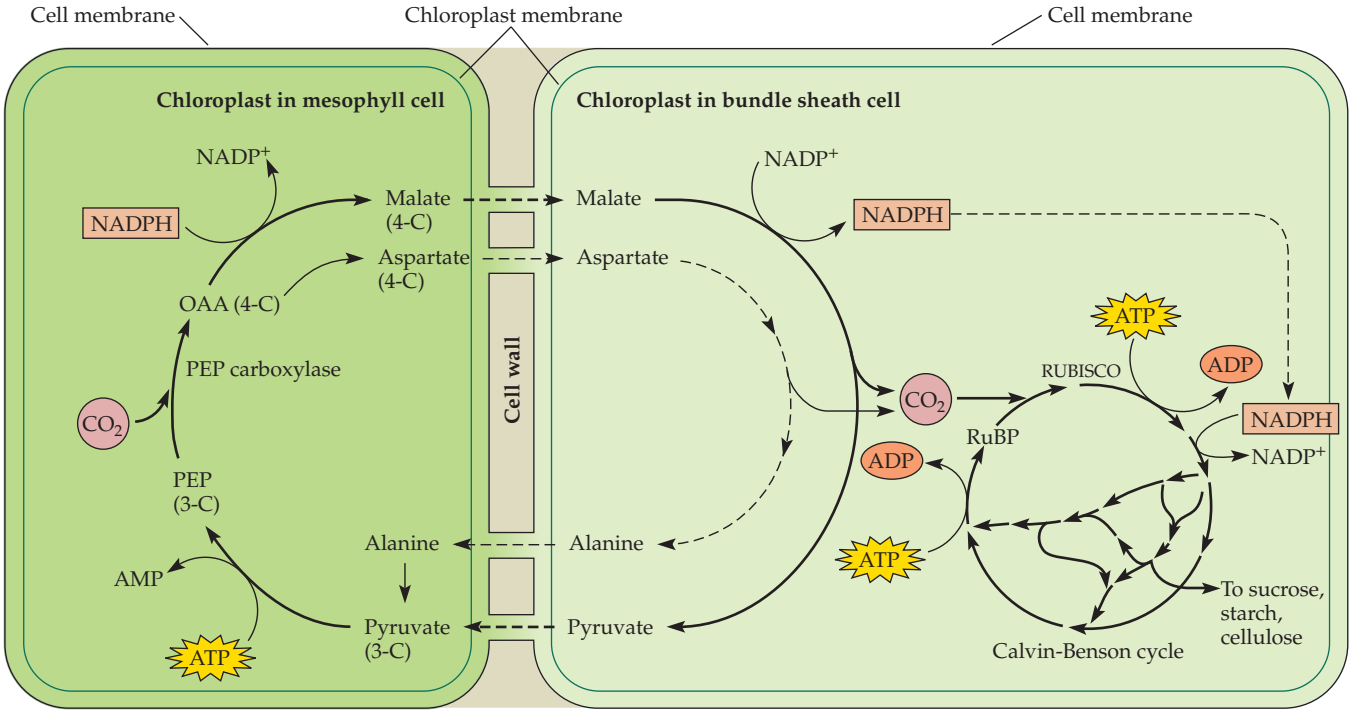



**Figure 2.10** The $C_4$ photosynthetic pathway. The biochemical steps that take place in the chloroplasts of the mesophyll cells are shown on the left, and those that take place in the chloroplasts of the bundle sheath cells are shown on the right. The Calvin-Benson cycle functions in the bundle sheath cells in the interior of the leaves, where oxygen concentrations are low.

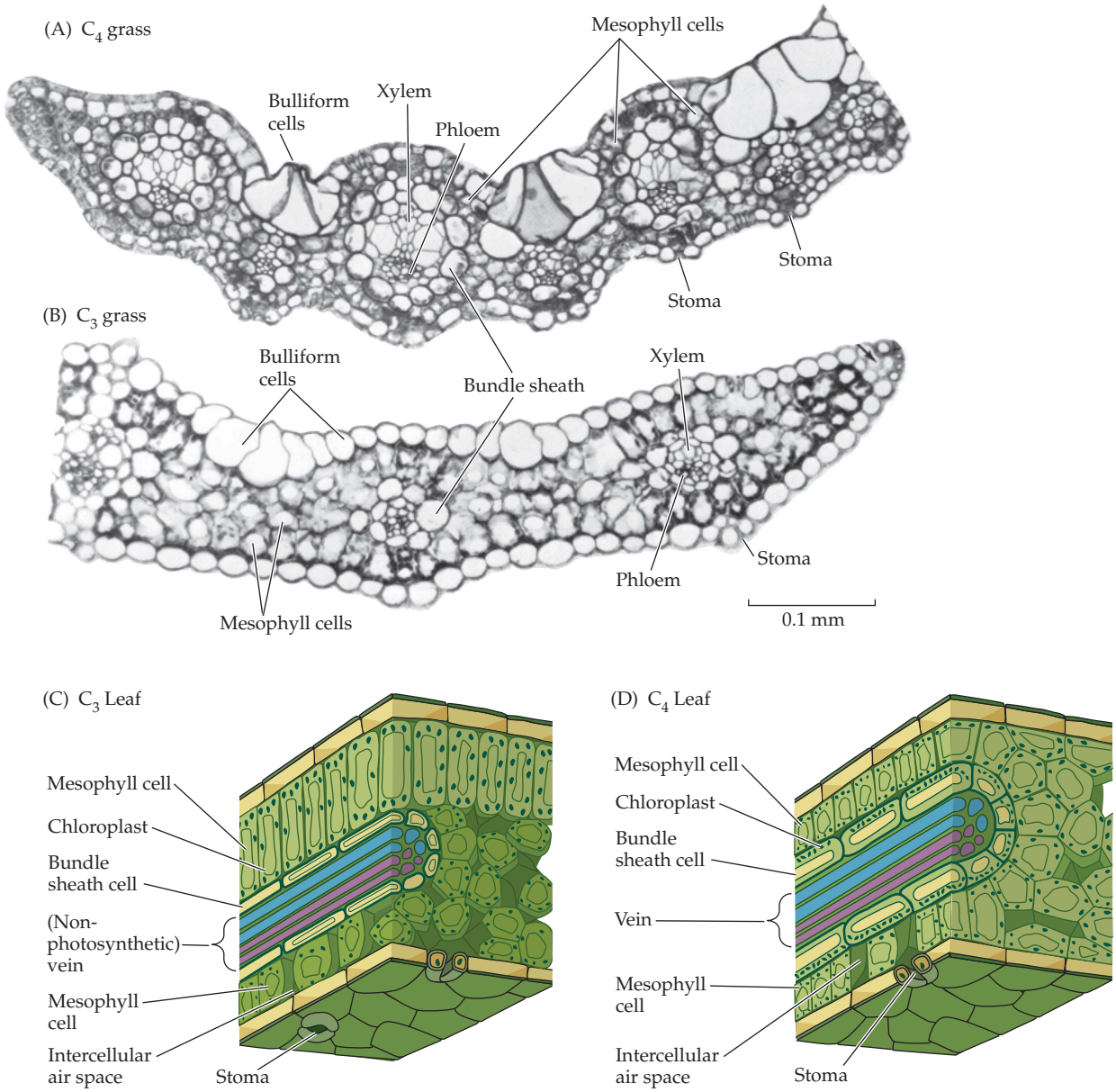



**Figure 2.11** Anatomy of (A) a leaf of *Saccharum officinarum* (sugarcane, Poaceae), a $C_4$ grass, and (B) a leaf of *Avena* sp. (oats, Poaceae), a $C_3$ grass, both in cross section, showing the differences in their architectures. Note the tight packing of the mesophyll cells in the outer ring surrounding the bundle sheath cells of the $C_4$ leaf, which themselves tightly encircle the vascular bundle (xylem and phloem), in contrast to the more loosely packed photosynthetic cells in the $C_3$ leaf. There are large numbers of chloroplasts in the bundle sheath cells (clustered at the outer edge within each cell) as well as in the mesophyll cells of the $C_4$ leaf, in contrast with the absence of chloroplasts in the bundle sheath cells surrounding the vascular bundle in the $C_3$ leaf. The bulliform cells act as hinges to allow the leaf to roll up during drought (see Chapter 3). (C, D) Diagrams of $C_3$ and $C_4$ leaf anatomy, showing arrangement of photosynthetic cells, vascular bundles with xylem and phloem, and other structures. (A, B from K. Esau. 1977. *Anatomy of Seed Plants*. John Wiley: New York, NY © John Wiley & Sons, Inc.; generously contributed to K. Esau by J. Sass; C, D after T. J. Mabry, unpublished.)

to the external atmosphere and so are shielded from high oxygen levels. The four-carbon acid OAA travels directly through thin strands of living cells, called **plasmodesmata**, from mesophyll cells to bundle sheath cells, where it is decarboxylated. The concentration of $CO_2$ in the bundle sheath cells of $C_4$ plants is an order of magnitude higher than its concentration in the photosynthetic cells of $C_3$ plants, and the ratio of $O_2$ to $CO_2$ is greatly reduced. Rubisco is "fed" a concentrated stream of $CO_2$ molecules and kept away from high $O_2$, resulting in the effective elimination of photorespiration in $C_4$ plants.

The consequences for plants of overcoming the limitations of rubisco are enormous (Sage and Monson 1998). (See Box 2C for a less obvious consequence—an isotope signature that can tell whether living, dead, or fossil herbivores ate $C_3$ or $C_4$ plants.) $C_4$ plants generally have higher

## BOX 2C Stable Isotopes and Photosynthesis

Chemically identical carbon compounds fixed via $C_4$ photosynthesis differ from those fixed via the $C_3$ pathway in a subtle way: in their ratios of the two stable isotopes of carbon, $^{12}C$ and $^{13}C$. Most molecules of $CO_2$ in the air contain $^{12}C$, the stable carbon isotope with an atomic weight of 12. Approximately 11 out of every 1000 $CO_2$ molecules instead have the slightly heavier stable isotope $^{13}C$, with an atomic weight of 13. ($^{14}C$, which is unstable and thus radioactive, is much rarer still.)

When the initial capture of $CO_2$ is made by rubisco, $^{13}CO_2$ molecules, being slightly heavier, are disproportionately left behind, unfixed, in the atmosphere (that is, the process discriminates against $^{13}C$). PEP carboxylase is more effective at "mopping up" all of the $CO_2$ molecules it encounters ($^{13}CO_2$ as well as $^{12}CO_2$) and thus discriminates less against $^{13}C$. The relative proportion of the two stable carbon isotopes in carbon compounds is expressed in reference to a standard (dolomite from the Pee Dee formation):

$$\delta^{13}C = \frac{\left[\left(^{13}C/^{12}C\right)_{sample} - \left(^{13}C/^{12}C\right)_{standard}\right] \times 1000}{\left(^{13}C/^{12}C\right)_{standard}}$$

For terrestrial $C_3$ plants, the mean $^{13}C$ value is –27‰ (range –36‰ to –23‰), while $C_4$ plants have a mean of –13‰ (range –18‰ to –10‰). (CAM plants have a range of $^{13}C$ values, depending on the relative proportion of carbon fixed initially by rubisco and by PEP carboxylase.) Because the ranges of the $^{13}C$ values for $C_3$ and $C_4$ plants are nonoverlapping, the carbon isotope ratio provides a convenient way to distinguish between $C_3$ and $C_4$ plants and plant tissues. These carbon isotope ratios also persist in the food chain. The fossilized carbon of herbivores that ate largely $C_4$ plants, for example, can be distinguished from that of those that ate $C_3$ plants. This property has been used by paleontologists and archaeologists to distinguish food sources; for example, it has been used to determine when grazers first began to depend heavily on $C_4$ grasses.

maximum rates of photosynthesis than do $C_3$ plants. The temperature optimum for photosynthesis is usually much higher for $C_4$ than for $C_3$ species (Figure 2.12). $C_3$ species typically become light saturated at levels well below full sunlight, whereas $C_4$ species often do not become light saturated even in full sunlight (Figure 2.13) because $CO_2$ uptake is not limited by the high oxygenase activity of rubisco. $C_4$ plants contain only one-third to one-sixth the amount of rubisco that $C_3$ plants have, and yet they are able to maintain the same or greater photosynthetic rates, resulting in higher **nitrogen use efficiency** (maximum photosynthetic rate per gram of nitrogen in the leaf; see Chapter 4). $C_4$ plants also have high water use efficiency—the ratio of $CO_2$ fixed to water lost (see below and Chapter 3), even for $C_4$ and a $C_3$ leaf at equal stomatal conductances, because the $C_4$ pathway makes $CO_2$ available to the Calvin-Benson cycle with great efficiency.

The cost of the numerous advantages of $C_4$ photosynthesis is also substantial: It takes additional energy from ATP to run the $C_4$ pathway. When light levels are high and conditions are optimal, the energy invested is more than compensated by the additional photosynthetic gains. At lower light levels, however, this expensive machinery must continue to be paid for, resulting in potential disadvantages for plants that possess it. A great deal of ecophysiological research in recent decades has been conducted with plants growing in the field to discover how environmental factors affect photosynthetic rates, including comparisons of $C_3$ and $C_4$ plants (see Box 3A). Insights into the genomic and transcriptomic basis of $C_4$ photosynthesis are being revealed by current research, including work to harness the mechanisms of the $C_4$ pathway to increase productivity in rice and other $C_3$ crop species (Huang and Brutnell 2016).

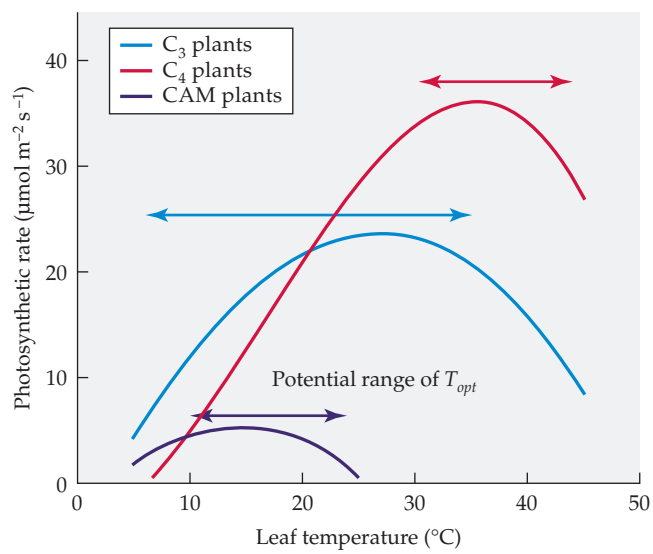



**Figure 2.12** Responses of photosynthetic rate to temperature in typical $C_3$, $C_4$, and CAM plants. $C_4$ plants typically have the highest maximum photosynthetic rates and peak at the highest temperatures, $C_3$ plants have lower rates on average and peak at lower temperatures, and CAM plants have the lowest maximum rates of photosynthesis. (From W. Yamori et al. 2014. *Photosynth Res* 119: 101–117.)

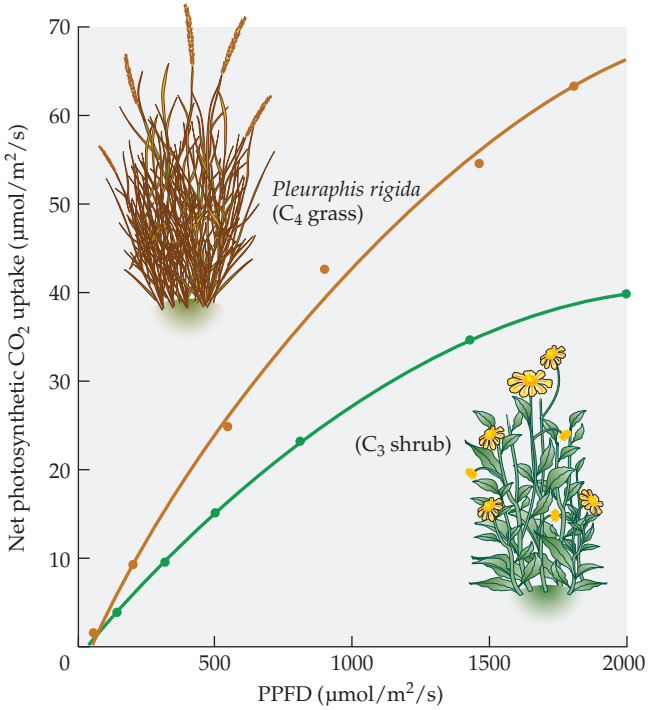



**Figure 2.13** Photosynthetic response to light intensity (PPFD, photosynthetic photon flux density) in a $C_4$ desert grass, *Pleuraphis rigida* (Poaceae), and in a $C_3$ desert shrub, *Encelia farinosa* (brittlebush, Asteraceae), from the western United States. At full sunlight (approximately 2000 µmol m2/s), the leaves of the $C_3$ shrub are light saturated, while those of the $C_4$ grass are not. (After P. Nobel. 1980. *Ecology* 61: 252–258; J. Ehleringer et al. 1976. *Science* 192: 376-377.)

## Crassulacean acid metabolism (CAM photosynthesis) is a specialized adaptation for minimizing water loss but at the cost of reduced photosynthesis and slow growth

**CAM photosynthesis**, or crassulacean acid metabolism, uses essentially the same biochemistry as $C_4$ photosynthesis to overcome the limitations of rubisco and eliminate photorespiration. However, CAM plants accomplish this in a very different way than $C_4$ plants (Figure 2.14). In $C_4$ plants, rubisco is found only in the bundle sheath cells, which are spatially segregated from external air. In CAM plants, rubisco is found in all photosynthetic cells. Instead of using spatial separation, CAM plants temporally separate the capture of light energy and the uptake of $CO_2$.

In $C_3$ and $C_4$ plants, stomata open during the day, when light capture and carbon uptake and reduction occur. CAM plants open their stomata at night. During the night, $CO_2$ is captured by PEP carboxylase, and large pools of organic acids are accumulated in the vacuoles of the photosynthetic cells. In daylight, the stomata close,

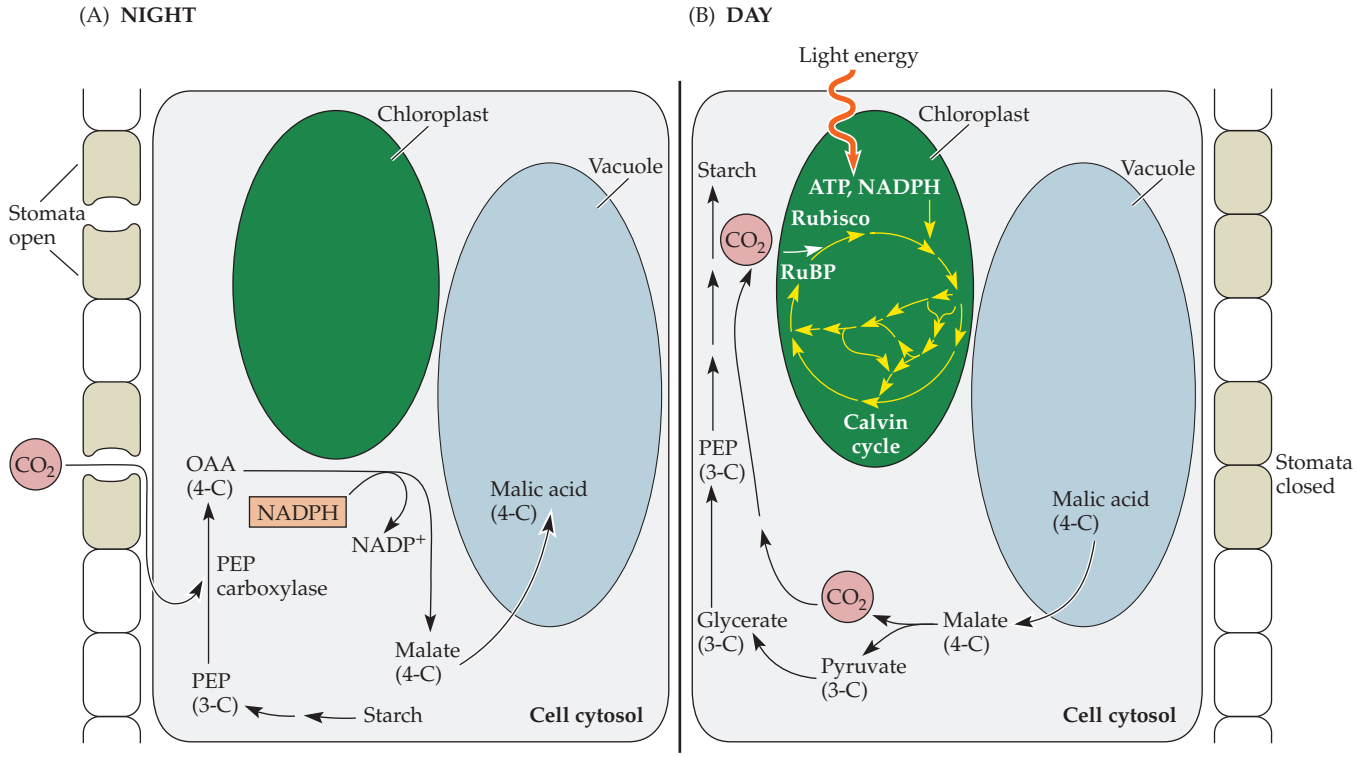



**Figure 2.14** Biochemical reactions involved in CAM photosynthesis. (A) At night, $CO_2$ enters the open stomata, diffuses into the cell, and is captured by the four-carbon organic acid phosphoenolpyruvate (PEP). Oxaloacetate (OAA) is formed, and then malate is produced and stored in a vacuole (a nonliving, membrane-bound storage organelle). (B) In the same photosynthetic cell, in the daylight, the stomata close, malate is transferred out of the vacuole, PEP is regenerated, and $CO_2$ is removed and diffuses to the chloroplasts. Once in the chloroplast, the $CO_2$ is recaptured in the Calvin-Benson cycle; the energy from the light reactions in the chloroplast provides the energy as it does in $C_3$ and $C_4$ photosynthesis.

and the light reactions occur just as they do in other plants. How do they do that? The organic acids that accumulated overnight are decarboxylated, providing $CO_2$ to the Calvin-Benson cycle. Because the stomata are closed, the Calvin-Benson cycle is not exposed to oxygen from the external air. Like $C_4$ photosynthesis, CAM photosynthesis allows rubisco to function in a high-$CO_2$ and low-$O_2$ environment, in which photosynthesis is favored over photorespiration.

CAM plants must have thick, succulent photosynthetic tissues with sufficient physical capacity to accumulate large amounts of organic acids overnight. The amount of photosynthate (carbon compounds produced in photosynthesis) that can be accumulated by CAM photosynthesis over a 24-hour period is limited by the amount of space available in the vacuoles, the large, nonliving storage sacs in the photosynthetic cells. Thick, succulent photosynthetic tissues accommodate more CAM carbon uptake (Figure 2.15). The temporal separation of the $C_4$ and $C_3$ reactions also slows maximum photosynthetic rates. CAM plants generally cannot accumulate carbon as rapidly as can either $C_4$ or $C_3$ plants.

CAM photosynthesis excels in one area, and that is **water use efficiency**, the grams of carbon fixed in photosynthesis per gram of water lost in transpiration. Because their stomata are open only at night, when temperatures are cooler and the air is more humid, water use efficiency is far higher in CAM plants than in $C_3$ or $C_4$ plants because water loss is much lower in CAM plants. Unlike $C_3$ and $C_4$ plants, in which the photosynthetic pathway is obligate (cannot be turned on or off), some CAM plants utilize both nighttime $CO_2$ uptake, with carbon reduction via CAM, and daytime $CO_2$ uptake via the $C_3$ pathway (see Figure 2.15). Some species behave in a facultative manner, using CAM photosynthesis when water is most limiting and also incorporating $CO_2$ using $C_3$ photosynthesis during the day, when conditions are more favorable, so as to achieve higher photosynthetic rates.

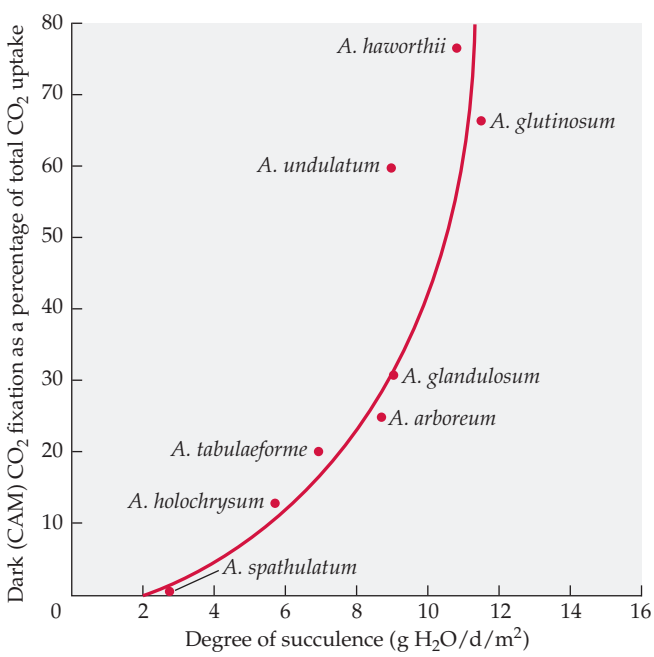



**Figure 2.15** Relationship between leaf succulence (grams of water in fresh leaves per unit leaf area) and the percentage of $CO_2$ uptake that occurs via CAM among species in the genus *Aeonium* (Crassulaceae). Plants with thicker, more succulent leaves exhibit a greater proportion of their total $CO_2$ uptake at night, using CAM. (After W. Larcher, 1995. *Physiological Plant Ecology*, 3rd ed. Springer: Berlin.)

## 2.4 $C_3$ Photosynthesis Is the Foundation for the Evolution of $C_4$ and CAM

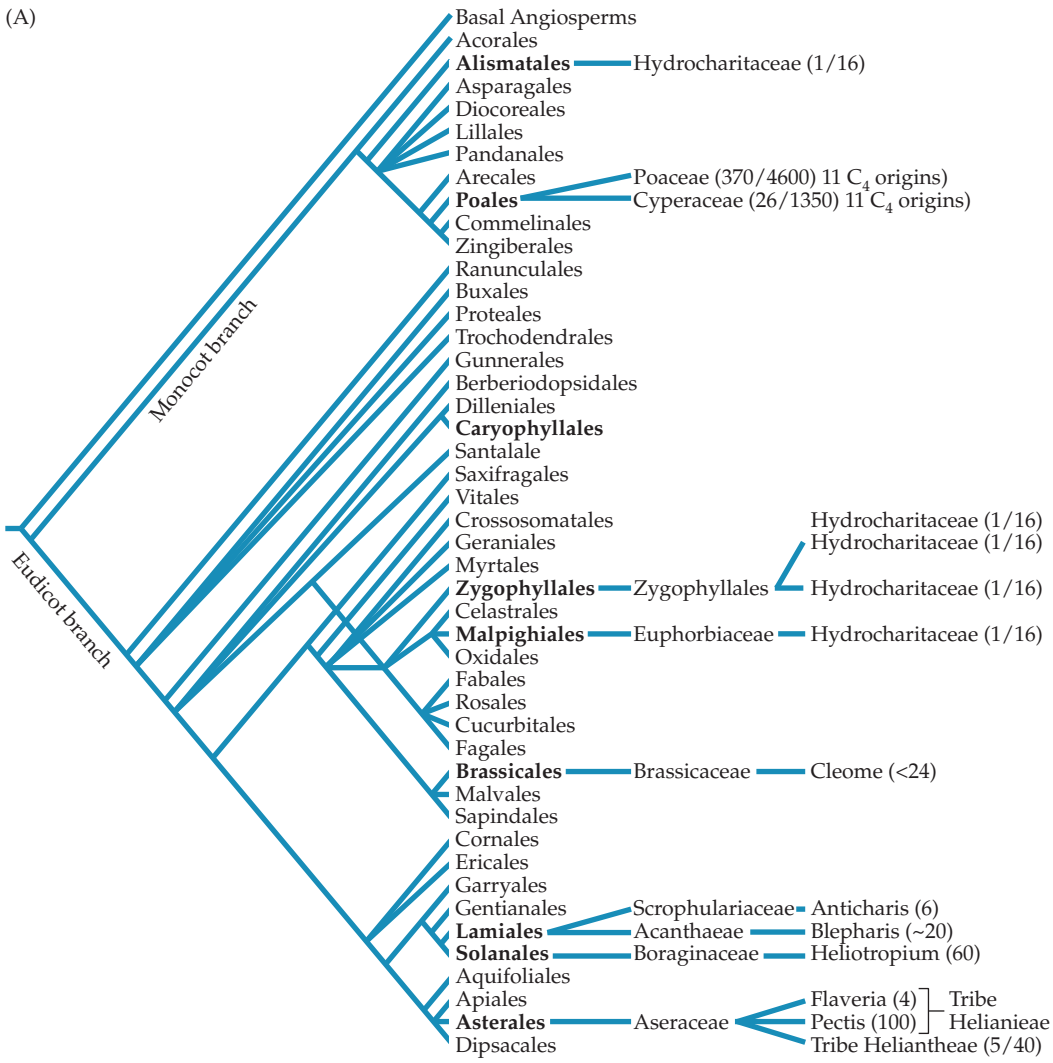
### $C_4$ and CAM evolved from $C_3$ photosynthesis many different times in many different plant families

The **phylogeny**, or evolutionary history of relationships of taxa by descent, reveals that $C_4$ photosynthesis and CAM have arisen independently from $C_3$ photosynthesis at least 65 times over the course of evolution (Figure 2.16, Sage 2016). There are apparent occasional reversals back to $C_3$ photosynthesis as well. Approximately 8145 angiosperm species (of approximately 300,000 angiosperm species total) in 418 genera belonging to 19 families use $C_4$ photosynthesis, including both monocot and eudicot species. No gymnosperms, ferns, mosses, or plants in other groups use $C_4$ photosynthesis. The monocots have the largest number of $C_4$ species, with over 1300 $C_4$ species in the Cyperaceae (the sedges and rushes) and over 5000 species in the Poaceae (the grasses). Approximately half the species in the Poaceae are $C_4$ plants. Other families with large numbers of $C_4$ species are the Amaranthaceae, Chenopodiaceae (now classified as a subfamily of the Amaranthaceae), and Euphorbiaceae. One family, the Gisekiaceae, has a single $C_4$ species. Typically, all of the species in a genus are either $C_3$ or $C_4$ plants, but there are several genera with both $C_3$ and $C_4$ members (e.g., Atriplex in the Amaranthaceae and *Panicum* in the Poaceae). A few species are intermediate between $C_3$ and $C_4$ in their structure and function.

### Photosynthesis first evolved about 2.5 billion years ago and has continued to evolve over Earth's history

Photosynthesis first evolved in bacteria in the Precambrian era, perhaps 2500 million years ago. Atmospheric $CO_2$ levels at that time, and for eons during the early evolution of land plants, were much greater than they are

(A)



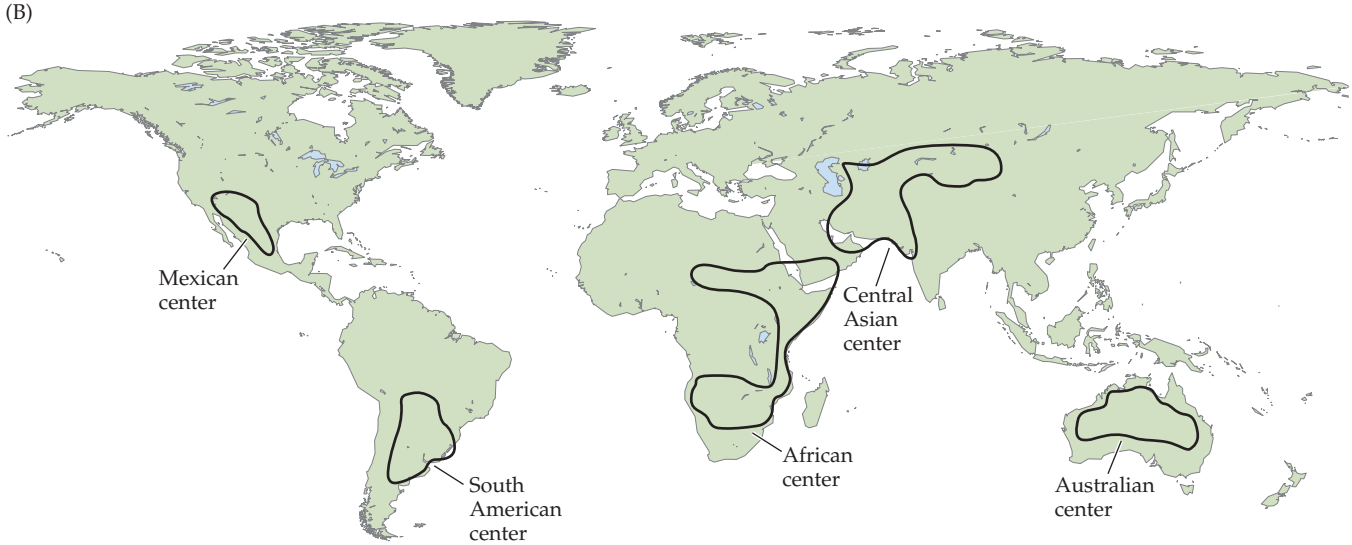
Courtesy of Carolyn Malmstrom

H. Zell/CC BY-SA 3.0

Stan Shebs/CC BY-SA 3.0

Courtesy of Laura Trejo

(B)



◀ **Figure 2.16** (A) The distribution of $C_4$ photosynthesis in the taxonomic orders of the angiosperms showing its multiple origins both within a family and among very distantly related families; those with $C_4$ photosynthesis are in bold. Lines to the right of these orders indicate families and principal $C_4$ genera within a lineage. Numbers in parentheses refer to estimates of genera/species numbers or, where relevant, just species numbers. Insets are photographs of plants from the families with the greatest numbers of $C_4$ species (from top): Poaceae (*Andropogon gerardii*, big bluestem), Amaranthaceae (*Amaranthus caudatus*, love lies bleeding), subfamily Chenopodiaceae (*Atriplex canescens*, fourwing saltbush), and Euphorbiaceae (*Euphorbia pulcherrima*, wild poinsettia). (B) Likely origins of different dicot lineages with $C_4$ photosynthesis. Some of the families with genera that have origins in these centers are: African (Scrophulariaceae, Aizoaceae), South American (Asteraceae, Chenopodiaceae/Amaranthaceae), Mexican (Asteraceae, Euphorbiaceae), Central Asian (Chenopodiaceae/Amaranthaceae). (A from R. F. Sage. 2004. *New Phytol* 161: 341–370; R. F. Sage. 2004. In: *A History of Atmospheric $CO_2$ and Its Effects on Plants, Animals and Ecosystems*, J. R. Ehleringer et al. [Eds.], pp. 185–213. Springer-Verlag: Berlin, Germany, based on P. Stevens. 2003. Angiosperm phylogeny website, version 4. http://www.mobot.org/MOBOT/Research/APweb/welcome.html; B from R. F. Sage. 2004. *New Phytol* 161: 341–370; R. F. Sage. 2004. In: *A History of Atmospheric $CO_2$ and Its Effects on Plants, Animals and Ecosystems*, J. R. Ehleringer et al. [Eds.], pp. 185–213. Springer-Verlag: Berlin, Germany, and references therein.)

now (see Figure 17.7). The evolution of stomata about 400 million years ago was a milestone in the early history of terrestrial plants because they permitted the regulation of water loss relative to carbon uptake. Stomata appear to have evolved from simple pores in the epidermis of plants that possessed a cuticle and intercellular gas spaces. Even before the evolution of functional stomata, these pores would have facilitated higher photosynthetic rates for the plants occupying a given area (Raven 2002).

During the Mesozoic era (248 to 65 million years ago), gymnosperms dominated; dinosaurs arose, became dominant, and went extinct; and flowering plants (the angiosperms) first appeared. At this time, atmospheric $CO_2$ levels were about 3.5 times greater than they are today (about 1260 parts per million partial pressure of $CO_2$). These levels would have resulted in limited photorespiration for $C_3$ plants, even at warm temperatures (Ehleringer and Monson 1993).

In the Eocene epoch (55 to 35 million years ago), when angiosperms began to dominate the landscape, $CO_2$ levels in the atmosphere became much lower (about 700 parts per million). The evolution of $C_4$ plants began about 30 million years ago in the Oligocene and has continued to recent times (Kellogg 1999; Christin et al. 2008; Sage 2016). As $CO_2$ levels dropped and photorespiration became an increasingly greater cost for plants with $C_3$ photosynthesis, $C_4$ photosynthesis would have been increasingly favored, particularly in warm climates and open habitats with high levels of sunlight. Climate change also played a significant role in the rise of $C_4$ grasses and grasslands, as many temperate and subtropical areas became more seasonal and drier. These changes (lower atmospheric $CO_2$, drier and more seasonal climates) accelerated during the Miocene, and by 10 million years ago many more $C_4$ lineages had evolved, and ecologically $C_4$ species became much more dominant (Strömberg 2011; Sage 2016). Today, as human activities rapidly push $CO_2$ levels up (see Chapter 16), plants using $C_4$ photosynthesis may lose some advantages over $C_3$ plants, although warmer temperatures and increasing drought in many parts of the globe may still favor $C_4$ species.

$C_4$ photosynthesis required the evolution of both plant biochemistry and anatomy and changes in a large number of genes. We are now beginning to understand the details of that transition from $C_3$ to $C_4$ photosynthesis (Christin and Osborne 2014; Bräutigam and Gowik 2016; Sage 2016). Anatomical changes occurred first, with greater vein density and changes in the physical relationship of mesophyll and bundle sheath cells bringing them into closer contact, followed by changes in the organelles in the bundle sheath cells. A critical next step was the development of the photorespiratory $CO_2$ pump (also known as $C_2$ photosynthesis), which brings the release of $CO_2$ from glycine in photorespiration into close physical proximity with the Calvin-Benson cycle (see Box 2B), at first in the same cell and eventually by separating the Calvin-Benson cycle into the bundle sheath cells. The photorespiratory $CO_2$ pump allows the recapture of some of the $CO_2$ that otherwise would be lost in photorespiration, improving the efficiency of photosynthesis and bringing carbon capture a step closer to $C_4$ photosynthesis. A number of plant species have intermediates between $C_3$ and $C_4$ photosynthesis, including *Moricandia arvensis* (colletxó or purple mistress, Brassicaceae, Figure 2.17), native to western North Africa and the Iberian peninsula, and species in the genera *Flaveria*, *Panicum*, *Mollugo*, and *Alternanthera*. All of them have a photorespiratory $CO_2$ pump and the greater vein density that is necessary for increased efficiency in recapturing photorespired $CO_2$. In the genus *Flaveria*, for example, $C_3$-$C_4$ intermediates have a photorespiratory pump that recycles $CO_2$ from photorespiration in the Calvin-Benson cycle into the bundle sheath cells. The next steps in the evolution of $C_4$ photosynthesis required ramping up of the activity of PEP carboxylase in the mesophyll cells, followed by increasing activity of the other enzymes of the $C_4$ pathway and declines and eventual disappearance in the enzymes of the Calvin-Benson cycle in those cells. The final step was the full anatomical separation of the $C_4$ and Calvin-Benson cycle pathways. Now,

© Photosteve/Alamy Stock Photo

**Figure 2.17** *Moricandia arvensis* (colletxó or purple mistress, Brassicaceae) uses carbon uptake that is intermediate between $C_3$ and $C_4$ photosynthesis.

plant scientists are trying to use the information learned about the evolution of $C_4$ photosynthesis to possibly genetically engineer $C_3$ food, forage, and biofuel crops to have a greater photorespiratory pump and eventually $C_4$ photosynthesis, with the goal of greatly increasing photosynthetic efficiency and ultimately crop yields. The greatest efforts are going into developing $C_4$ rice (rice, *Oryza sativa*, Poaceae, is a strictly $C_3$ species).

The evolutionary transition to CAM from $C_3$ ancestors was apparently much easier than the transition to $C_4$ photosynthesis, because fewer morphological and biochemical changes were necessary. More than 20,000 angiosperm species in approximately 25 families have evolved this pathway, including both monocots and dicots. There are also CAM species among the ferns, as well as in the aquatic, primitive vascular plant genus *Isoetes* (quillworts, Isoetaceae). It is not unusual for a plant family to have both $C_3$ and CAM members; the ecologically and taxonomically diverse genus *Euphorbia* (Euphorbiaceae) includes $C_3$, $C_4$, and CAM species. There are also species that appear to represent intermediates between fully expressed CAM and more primitive CAM cycling (that is, $C_3$ photosynthesis with the ability to recycle $CO_2$ from ordinary cellular respiration at night).

## 2.5 $C_3$, $C_4$, and CAM Plants Each Have Distinct Growth Forms, Phenology, and Distributions

### The three photosynthetic types dominate in different habitats and differ in growth form

The most striking evidence for the adaptive significance of the three photosynthetic pathways comes from patterns of species distribution and abundance. $C_3$ plants are by far the most abundant in terms of both number of species and total biomass. Only about 3% of the angiosperms are $C_4$ species; nevertheless almost 25% of terrestrial primary productivity is due to $C_4$ photosynthesis, because of the high productivity of $C_4$ photosynthesis and the ecological abundance of these species (Still et al. 2003). While there are more CAM species overall than $C_4$ species, CAM plants are much less abundant than $C_4$ plants in terms of both biomass and worldwide distribution. CAM plants are generally found in particular habitat types, and their importance, whether it is measured as the percentage of species in a regional flora or as the proportion of biomass in the vegetation, is small in most ecosystems. Each photosynthetic pathway, as we have just seen, offers advantages under particular conditions but has costs associated with it that put it at a disadvantage under other conditions.

$C_4$ grasses dominate, or are major components of, warm grassland ecosystems covering vast portions of the globe (Knapp and Medina 1999). Several important crop species are $C_4$ grasses (e.g., maize, millet, sorghum, and sugarcane, as well as other forage crops for domestic animals). It is no accident that $C_4$ photosynthesis is so important in the grasses. Grassland environments are often characterized by warm, bright conditions, and even outside of grasslands, grasses are commonly found in high-light environments. These are precisely the conditions under which $C_4$ photosynthesis is favored. While there are many $C_4$ annuals, perennials, and even shrubs, very few tree species use $C_4$ photosynthesis; likewise, while there are $C_4$ species that are native to shady, cool, and aquatic habitats, they are the exceptions. Some $C_4$ species are among the most important and globally abundant weeds.

CAM plants are found in two distinct habitat types with two distinctive growth forms: as succulent terrestrial plants growing in deserts or other arid habitats, and as epiphytes growing in the canopies of trees in tropical and subtropical habitats. CAM photosynthesis has evolved in many different families in arid environments (Figure 2.18), including cacti (Cactaceae), agaves (Agavaceae), species in the Euphorbiaceae from Africa, and species in the Crassulaceae. Succulents in this group often serve as examples of convergent evolution due to their similar external morphologies (see Figure 18.9).

**Epiphytes** are plants that rely for support on other plants, generally living on trees, but depend on minerals deposited from the atmosphere for nitrogen and other mineral nutrients (Figure 2.19A). Epiphytes have at least one thing in common with desert plants, being subjected frequently to severe shortages of water. Even in a tropical rainforest, life high in the canopy can be surprisingly dry (see Figure 18.11B). A number of families have epiphytic CAM species. Some familiar examples are Spanish moss

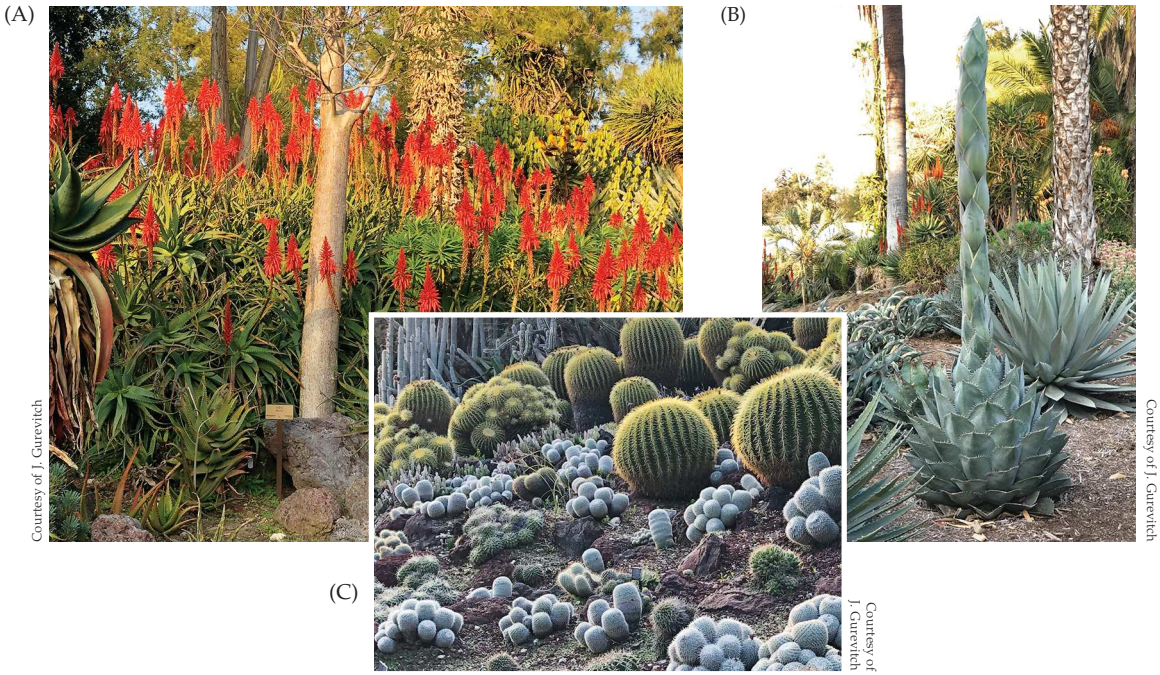



**Figure 2.18** Various CAM species. These succulents display different forms of fleshy leaves or other tissues. (A) *Aloe arborenscens* (candelabra aloe, Asphodelaceae, with red flowers). (B) *Agave palmeri* (blue or Palmer agave, Asparagaceae). (C) *Mammillaria ceminispina* (twin spined cactus, smaller and densely downy and spiny) and *Echinocactus grusonii* (golden barrel cactus, both Cactaceae).

(*Tillandsia usneoides*, Bromeliaceae; Figure 2.19B) and many orchids (the family Orchidaceae; Figure 2.19C) and ferns.

Perhaps surprisingly, CAM plants are rare in the very hottest, driest deserts, where rainfall not only is limited, but cannot be counted on in any given year. For example, there are few CAM plants in Death Valley, California, or in the Sahara of North Africa. In the Atacama Desert of Peru, where it may be years between rains in some places, cacti are almost entirely restricted to the zone that receives moisture as coastal fog. Why might CAM plants be kept out of the most arid deserts? Under these extreme conditions, rapid growth during the brief, unpredictable periods when water is available is strongly favored (see Figure 16.22), and CAM plants do not generally have the capacity to grow rapidly. Another factor may be the inability to survive a very prolonged drought. Both cellular respiration and water loss would be less in a completely dormant plant with no aboveground photosynthetic parts than in a succulent plant. Even with closed stomata, tiny amounts of water are lost from any green tissue, which may add up to a substantial amount over an extended time without any precipitation. Succulents, particularly those with CAM photosynthesis, are thus most abundant and diverse in deserts with both relatively higher levels and greater predictability of rainfall. In somewhat drier deserts with less predictable rainfall, perennials with either $C_3$ or $C_4$ photosynthesis are more likely to be dominant, depending on the season in which rain is most abundant and likely, as we will see shortly. In the driest deserts, the perennials decline in importance, and annual species tend to become more prevalent.

## $C_3$ and $C_4$ plants grow most actively in different seasons

The three photosynthetic types also tend to differ in their **phenology** (the seasonal timing of life history events, such as initiation of growth, flowering, and dormancy). $C_4$ photosynthesis is well represented among annual species that complete most of their growth in summer; such summer annuals are common in semiarid environments with seasonal rainfall, as well as among weedy species that colonize disturbed environments, such as roadsides and agricultural fields. In areas with mediterranean climates (such as California, southern France, and western Australia), where most of the rainfall occurs in winter, the native species are overwhelmingly $C_3$ plants, but the weeds that invade fields irrigated in summer may be $C_4$ species.

In the southwestern deserts of the United States, distinct patterns exist for $C_3$ and $C_4$ annual plants. In the deserts of eastern California, where the rainy season is in winter and early spring, the annuals are $C_3$ species and grow when temperatures are cool. In New Mexico and Texas, where there is a summer rainy season, the desert annuals are largely $C_4$ species. In the Sonoran Desert of Arizona and northern Mexico there are two separate rainy seasons, and annual $C_3$ species, including many

(A)

Courtesy of W. S. Judd

(B)

Matthew Townsend/CC BY-SA 3.0

(C)

JKguin/CC BY-SA 4.0

**Figure 2.19** CAM epiphytes. (A) An epiphytic bromeliad, *Guzmania monostachya* (Bromeliaceae). (B) *Tillandsia usneoides* (Spanish moss, Bromeliaceae) is an epiphyte that grows hanging from the branches of large trees in the southeastern United States. (C) *Prosthechea boothiana* (Florida dollar orchid, Orchidaceae), an endangered native of South Florida.

beautiful wildflowers, dominate in early spring, while $C_4$ annuals are found in summer. The periods of greatest growth for $C_3$ and $C_4$ perennials in these environments follow a similar pattern.

Similarly, the "cool season" perennial grasses and **forbs** (broad-leaved herbaceous plants) that predominate in spring and again in autumn in the mixed-grass prairie of South Dakota are largely $C_3$ species, while the "warm season" species that grow most actively in midsummer have a much greater representation of $C_4$ plants. However, in the shortgrass prairie farther south, while there are differences in phenology between the $C_3$ and $C_4$ species, the period of maximum growth for the two photosynthetic types does not differ (Hazlett 1992).

## $C_3$, $C_4$, and CAM species have different geographic distributions

Various studies have considered the broad geographic distributions of $C_3$, $C_4$, and CAM species and the climatic factors associated with their relative predominance. Not surprisingly, the physiological differences between $C_3$ and $C_4$ photosynthesis are associated with differences in distribution between plants of the two types. Across North America, the percentage of $C_4$ grass species is most strongly determined by summer temperatures: $C_4$ grasses are most common where summers are warm (Figure 2.20).

José Paruelo and William Lauenroth (1996) examined the relative abundances (in terms of biomass rather than number of species) of $C_3$ grasses, $C_4$ grasses, and shrubs in the western United States (Figure 2.21). $C_4$ grasses are increasingly dominant where temperatures, total rainfall, and the proportion of rain falling during summer are higher. $C_3$ grasses are favored by cool temperatures

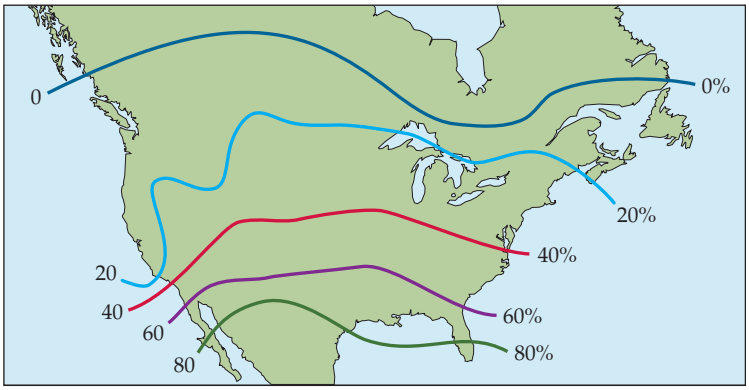



**Figure 2.20** Distribution of $C_4$ grass species as a percentage of the total grass flora across North America. (Data from J. A. Teeri and L. G. Stowe. 1976. *Oecologia* 23: 1–12.)

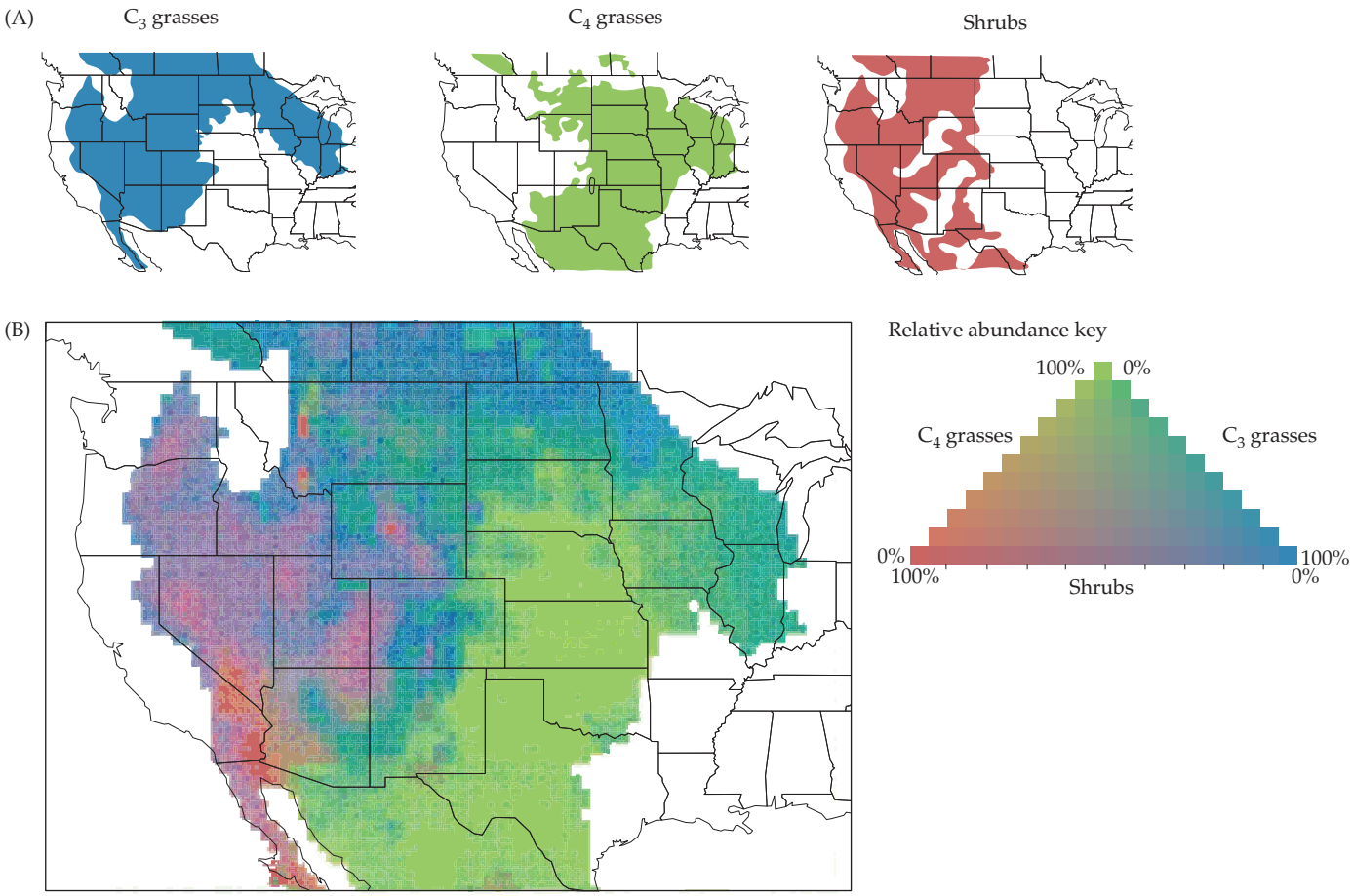



**Figure 2.21** $C_3$ grasses, $C_4$ grasses, and shrubs are the dominant functional groups in the central and western part of North America, which is dominated by grasslands and shrublands. The maps indicate the relative abundances of each as an approximate percentage of the total biomass. (A) $C_4$ grasses are the most abundant species in the southern plains, while $C_3$ grasses predominate farther to the west and north. Shrubs are most abundant in the far west. (B) The combined abundances of $C_4$ grasses, $C_3$ grasses, and shrubs, as indicated by the relative abundance key. (After J. M. Paruelo and W. K. Lauenroth. 1996. *Ecol Appl* 6: 1212–1224. Data from A. W. Kuchler. 1964. *The potential natural vegetation of the conterminous United States*. AGS: New York; J. L. Dodd. 1979. In *Perspectives in Grassland Ecology*, N. R. French [Ed.]. Springer Verlag: New York; R. Leemans and W. Cramer. 1991. *The IIASA database for mean monthly values of temperature, precipitation, and cloudiness on a global terrestrial grid*. IIASA: Luxembourg.)

and by a greater proportion of rainfall during the winter months. Shrubs, which have deep roots that increase their access to water stored deep in the soil, are most abundant as total rainfall decreases and as the proportion of winter rainfall increases. Consequently, in the great North American grasslands between the Mississippi River and the Rocky Mountains, the southeasternmost portions are dominated by $C_4$ grasses, while $C_3$ grasses increase to the north and to the west, and the westernmost, driest part of this region is dominated by (largely $C_3$) shrubs (Figure 2.22; see also Figure 18.7).

Other research has documented changes with elevation, from a vegetation dominated by $C_4$ species in the warm, often dry habitats at low altitudes to a $C_3$-dominated flora in the cooler, moister conditions at high altitudes, in Kenya, Costa Rica, Hawaii, Arizona, and Argentina. CAM species are, as one might expect from their superior water use efficiencies, most closely associated with aridity in temperate and semiarid regions, but not the most arid regions. In the Sonoran Desert of Arizona, the lowest elevation sites are dominated by $C_3$ shrubs and small trees, such as *Larrea tridentata* (creosote bush, Zygophyllaceae). CAM species such as the giant cactus *Carnegiea gigantea* (saguaro, see Figure 13.17) are found where rainfall is greater at somewhat higher elevations on rocky slopes.

The studies of correlations between climate and photosynthetic type all share a common limitation. It is possible that their results are due, to some unknown extent, to shared ancestry. Related species have shared history,

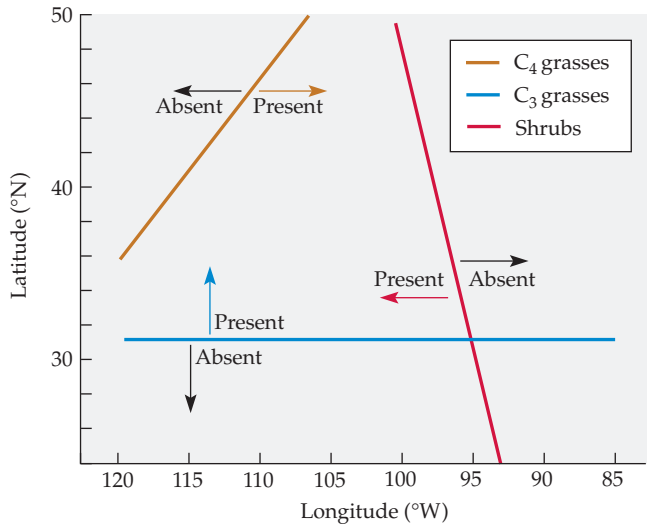



**Figure 2.22** Results of a model with lines showing the latitude and longitude where $C_3$ grasses, $C_4$ grasses, and shrubs are predicted to be absent or present within the central and western region of North America. The distributions of $C_4$ grasses and shrubs are largely influenced by longitude, while $C_3$ grasses are influenced by latitude. The predictions of this model were largely in agreement with the distribution patterns shown in Figure 2.21. (After J. M. Paruelo and W. K. Lauenroth. 1996. *Ecol Appl* 6: 1212–1224.)

with the same background climate and biogeography, regardless of photosynthetic type. $C_4$ photosynthesis evolved in families from warm environments. Therefore $C_4$ species in those families, like their $C_3$ relatives, should have numerous adaptations to living in those environments, regardless of the pathway they use to fix $CO_2$. Tropical $C_4$ grasses do not perform well in cold temperatures; neither do tropical $C_3$ grasses. Many $C_3$ species live in warm, dry places, and a few $C_4$ and CAM species live in colder climates. These factors make it somewhat difficult to disentangle the adaptive significance of photosynthetic type from all of the other adaptations of species of contrasting types. It is particularly difficult to disentangle the possible ecological significance of the different $C_4$ subtypes, which belong to different phylogenetic groups but have experienced convergent evolution in developing $C_4$ photosynthesis in somewhat different ways. We discuss this more general problem of convergent evolution in Chapter 18.

Nevertheless, there is abundant other evidence, derived from a variety of approaches, that $C_4$ and CAM photosynthesis are adaptations to environmental factors. We have a clear idea of what those factors are, and we understand the mechanisms behind the adaptations. While general patterns may be very revealing, they cannot explain everything, and photosynthetic type is only one of a host of factors that determine how a plant interacts with its environment.

## 2.6 Plants Possess Many Different Adaptations to Their Light Environments

### Many plants can detect the length of daylight and how it is changing seasonally

Photoperiodic responses control many different aspects of plant lives in temperate and arctic environments. **Photoperiod** is the number of hours of light and darkness in a 24-hour cycle. Photoperiod responses synchronize the internal circadian clock, which is genetically regulated. Many plants not only can detect, quite precisely, how long the day (or, actually, the night) is, but also can determine whether day length is increasing (as it does in spring) or decreasing (as it does in autumn). Even in tropical and subtropical environments some species can respond to very slight changes in day length and use these as cues (Borchert et al. 2005); of course, day length changes very little near the equator. Seed germination, the initiation of flowering, the formation of underground tubers to store food, the initiation of cold tolerance, and the color change and senescence of leaves in autumn are all often controlled by responses to day length. Day length is a more reliable indicator of season than is temperature or other more immediate cues, and plants can make use of this information to time important phenological events, including flowering, the onset of dormancy, or leaf fall, and to synchronize responses with other members of their populations. Day length is often used together with other environmental cues, including temperature and moisture. Climate change (see Chapter 16), however, is causing many species' day length cues to diverge from the seasonal weather that they are meant to signal. For example, in alpine systems snows are melting earlier, so less water may be available when plants start growing (Ernakovich et al. 2014).

Hal Mooney and colleagues found genetic variation in both the sensitivity and the nature of the response to day length in the widespread Arctic/alpine species *Oxyria digyna* (alpine sorrel, Polygonaceae) by growing plants together in controlled conditions. Populations from further south at high elevations in California, Wyoming, and Colorado flower when the day length is about 15 hours. In populations from farther north, near the Canada–United States border, some plants will flower at a 15-hour day length, but many will not. In populations from north of the Arctic Circle, flowering requires a day length of over 20 hours (Mooney and Billings 1961; Billings

Photo by Barbara Lily. Courtesy of Hal Mooney

Hal Mooney

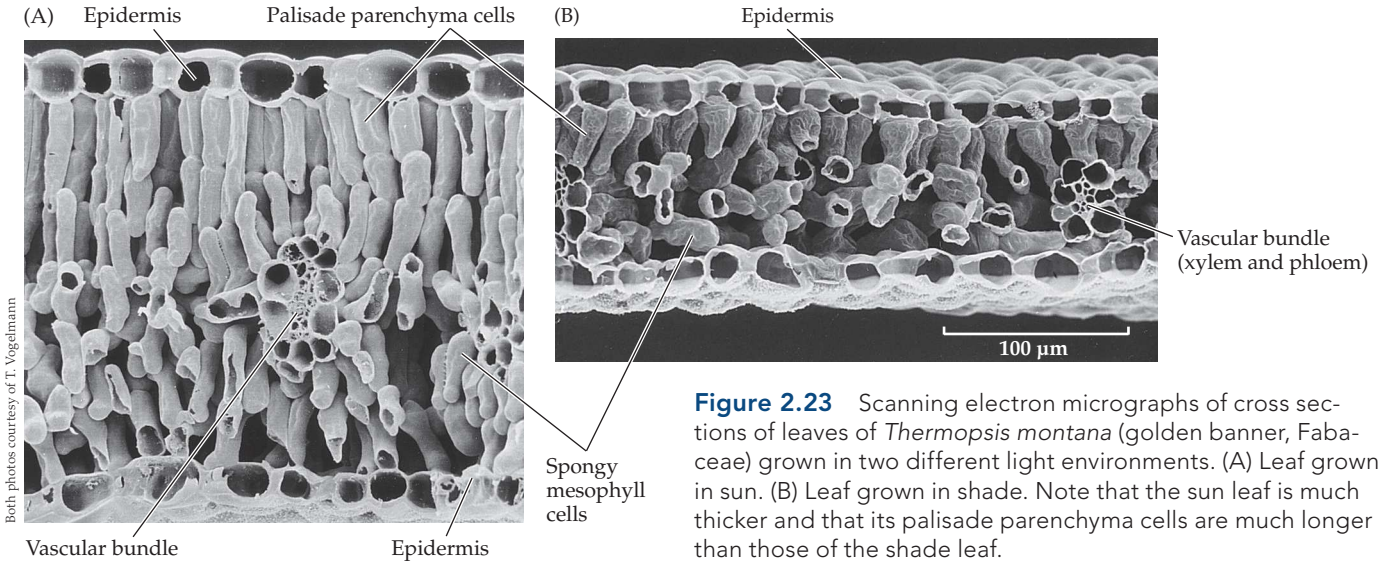



**Figure 2.23** Scanning electron micrographs of cross sections of leaves of *Thermopsis montana* (golden banner, Fabaceae) grown in two different light environments. (A) Leaf grown in sun. (B) Leaf grown in shade. Note that the sun leaf is much thicker and that its palisade parenchyma cells are much longer than those of the shade leaf.

and Mooney 1968). Thus, the onset of flowering occurs in mid-May both at the Canada–United States border and north of the Arctic Circle, although at the border day lengths are 15 hours at this time, while at the Arctic Circle they are 20 hours. The more southern, high-elevation populations do not experience 15-hour day lengths until mid-June or later. These alpine populations may still be under snow in mid-May, with a high probability of frost, and do not flower until about a month later than the more northern populations.

Plant sensitivity to day length was first discovered in the 1920s, but a great deal has been learned about photoperiod responses in recent years, particularly with respect to the intiation of flowering. Red, far-red, and blue light signals are detected by compounds including phytochromes, cryptochromes, and the blue-light receptor protein called ZTL, which interact with the plant's internal circadian clock. The *CO* gene and CO protein are major regulators of photoperiodic induction of flowering (a more detailed overview of these mechanisms can be found in Jackson 2009). As we shall see in Chapter 9, phytochromes also play other roles in allowing plants to detect and respond to their environments, including the detection of neighboring plants.

## Leaves grown in sunlit and shaded conditions can differ in structure and function

A leaf that functions well in bright sunlight usually will not perform well in deep shade, and the reverse is also true. Many species of plants produce different kinds of leaves in the sun and in the shade (Figure 2.23). Leaves at the top of a tree, for example, may differ in a variety of ways from those deep within the canopy. "Sun leaves" tend to have higher concentrations of rubisco, chlorophyll, and other key components of both the light reactions and carbon reduction per unit leaf area (Table 2.2). They are also typically smaller in leaf area and thicker than "shade leaves," with greater leaf mass per leaf area, or **LMA**. Leaf mass per leaf area is also called **specific leaf mass** or **specific leaf weight**, a measure of dry mass per unit leaf area. The inverse, **specific leaf area**, or leaf area per gram of dry mass, is also often reported (see Chapter 6).

In some plants, sun and shade leaves can also have different shapes, with deeper lobing, for example, in leaves produced in bright sunlight. As a consequence of the greater investments made in the machinery of photosynthesis, sun leaves usually have much higher light saturation levels and greater maximum photosynthetic

**TABLE 2.2** Contrasts between sun leaves and shade leaves of *Fagus sylvatica* (beech, Fagaceae)

| Character | Sun leaves | Shade leaves |
|---|---|---|
| Stomatal density (number/mm$^2$) | 214 ± 26 | 144 ± 11 |
| Leaf thickness (µm) | 185 ± 12 | 93 ± 5 |
| Leaf area (cm$^2$) | 29 ± 4 | 49 ± 7 |
| Fresh weight (g) | 0.5 ± 0.1 | 0.4 ± 0.1 |
| Dry weight (g) | 0.24 ± 0.03 | 0.12 ± 0.02 |
| Water content (% fresh weight) | 53 ± 4 | 70 ± 5 |
| Total chlorophyll (mg/g dry weight) | 6.6 ± 2 | 16.1 ± 2 |
| Total chlorophyll on a leaf area basis (mg/100 cm$^2$) | 5.5 ± 1.8 | 3.9 ± 0.4 |

*Source*: After H. K. Lichtenthaler et al. 1981. *Photosynth Res* 2: 115–141.
*Note*: Values are means of 9 leaves ± 1 standard deviation.

rates than shade leaves (Figure 2.24). Shade leaves often have lower rates of cellular respiration, perhaps because there is less "machinery" to maintain, and they may have higher net rates of photosynthesis at low light levels, with lower light compensation points.

The ability of individual plants to produce different kinds of leaves in different light environments is an example of **phenotypic plasticity**, the ability of an individual of a given genotype to produce different structures (such as leaf tissues) or to function differently under different environmental conditions (see Chapter 9). This flexibility in response to environmental conditions can also be called **acclimation** (a type of plasticity that is potentially reversible, also called acclimatization).

Not all acclimation to changes in light levels is through the production of different types of leaves. In the study of northeastern forest understory species discussed above (Rothstein and Zak 2001), the shifts in light compensation point in *Viola pubescens* and *Tiarella cordifolia* were due to physical and chemical re-adjustments within individual leaves, rather than to the production of different kinds of leaves in different seasons.

Species are adapted to their native light environments. Contrasts between the leaves of plants that are native to high-light environments and those of plants from shaded habitats are similar in many ways to contrasts between sun leaves and shade leaves produced by the same individual. Species adapted to bright habitats frequently incur costs or have specializations that would put them at a disadvantage in deep shade, and the reverse is also true. Many plants could not survive in environments with light regimes drastically different from the ones they live in. For example, the high investments in rubisco and chlorophyll and the higher respiration rates that generally characterize the leaves of plants from high-light environments would not provide sustainable returns in photosynthesis in deeply shaded habitats.

Unlike the more general differences between sun and shade leaves on the same plant, species native to high-light or low-light habitats often possess unique features that enhance their success in their own environment. Among high-light species, one such character is **solar tracking**. Irwin Forseth and James Ehleringer (1982) found that many desert annuals and some perennials from warm deserts—a high-light environment—have leaves that track the sun across the sky during the day. This strategy increases the amount of light available to the chloroplasts over the course of the day (Figure 2.25), thereby increasing photosynthetic rates.

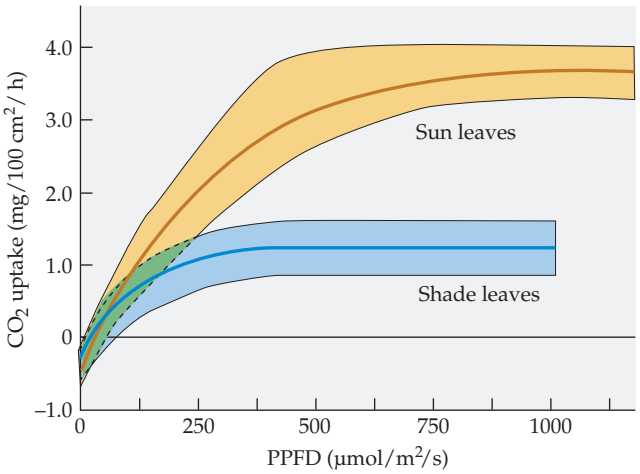



**Figure 2.24** Photosynthetic $CO_2$ uptake in response to light intensity for sun leaves and shade leaves of *Fagus sylvatica* (beech, Fagaceae) on a leaf area basis (mg $CO_2$/100 cm$^2$/h), showing the midpoint and range for nine different sun leaves and nine different shade leaves. The sun leaves become light saturated at much higher light intensities, and the maximum photosynthetic rates are much higher for sun leaves. (After H. K. Lichtenthaler et al. 1981. *Photosynth Res* 2: 115–141.)

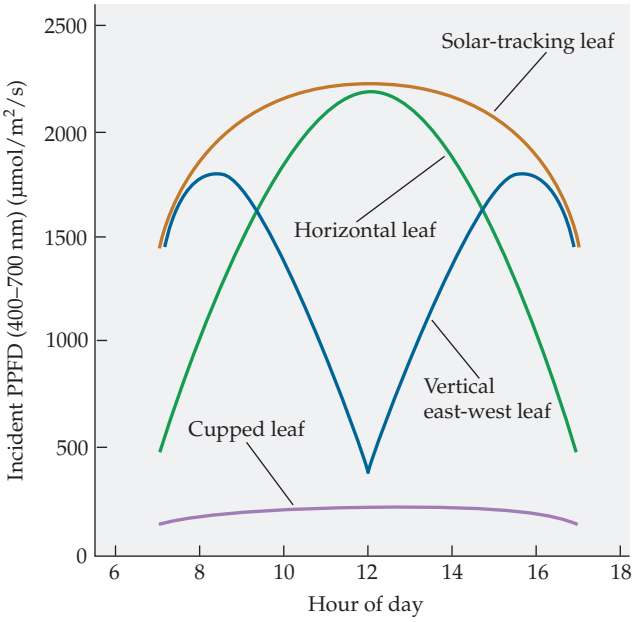



**Figure 2.25** The amount of photosynthetically useful light received over the course of a midsummer day in the desert at the leaf surface ("incident PPFD") by a solar-tracking leaf, a leaf held horizontally (in which case the light received in the early morning and late afternoon is sharply reduced), a leaf held vertically (in which case the light received during the middle of the day is sharply reduced), and a cupped leaf. The solar-tracking leaf receives at least 38% more light than the fixed horizontal or vertical leaves. (After J. Ehleringer. 1985. In *Physiological Ecology of North American Plant Communities*, B. F. Chabot and H. A. Mooney [Eds.], pp. 162–180. Chapman and Hall: New York; J. Ehleringer and I. Forseth. 1980. *Science* 210: 1094–1098.)

Why should solar tracking be advantageous in the bright, sunny environment of a desert? In desert environments, water is available for growth for only a small part of the year, and desert annuals, in particular, often germinate, grow, and reproduce in a very short time before water becomes unavailable and they die. Solar trackers are adapted to maximize their light capture in the short growth period available to them. In North American deserts, the proportion of species that are solar trackers becomes greater as the length of the growing season gets shorter. We explore adaptations to desert environments in more detail in Chapter 3.

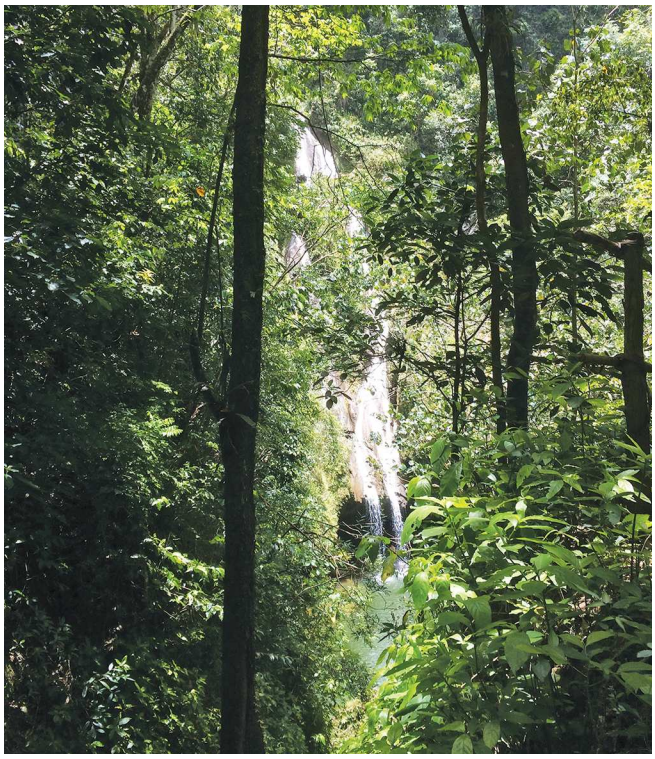
Some plants that live in deeply shaded habitats also have unique features. Many understory species that live in the deep shade of humid tropical forests possess a number of unique adaptations to the quality and quantity of light available to them. Light intensities at the rainforest floor may be less than 1% of the PPFD at the top of the canopy. Characters such as velvety or satiny leaf surfaces and blue iridescence are found in widely divergent species in understory habitats (Lee and Graham 1986; Lee 1997). Each of these characters appears to slightly enhance the light available for photosynthesis. That slight advantage in such an extreme environment may, however, make the difference between life and death, or between reproducing and failing to leave descendants.

In some species, a brilliant iridescence or velvety sheen is produced in epidermal cells, which may act in various ways to enhance both the quantity of light reaching the photosynthetic tissue and the quantum efficiency of photosynthesis. Matthew Jacobs and colleagues (2016) studied the intense, metallic blue iridescence found in the leaves of begonia species (Figure 2.26) native to the tropical rainforests of Malaysia. The blue iridescence in such leaves is not produced by blue pigments, but is a structural effect (Glover and Whitney 2010, Box 2D). The brilliant structural blue color of these leaves is due to unusual, highly modified chloroplasts called iridoplasts found in the epidermis (which in most plants has no chloroplasts and is clear). The iridoplasts increase light capture in the green wavelengths, normally of limited use in photosynthesis, due to highly structured crystalline properties of their membranes, which possess nanoscale structures that manipulate light interference. In this forest floor environment, light in the photosynthetically most usable range, 400–700 nm, is particularly depleted because it has been absorbed by the plants in the canopy above. The structural properties of the iridoplasts enable them to absorb more light at wavelengths that are less depleted at the forest floor but are still somewhat useful in photosynthesis. Other tropical forest floor species, from ferns to *Selaginella* species (spikemoss, Lycopodiopsida), also appear to have similar adaptations using various structural features and iridesence to enhance light capture.

Understory environments are not uniformly dark, however. The light environment is changing constantly, dim one moment and illuminated by the brilliant flash of a sunfleck the next (Figure 2.27). Shade-adapted species appear to be capable of using brief sunflecks with unusual efficiency. Studies by Robin Chazdon and Robert

With Permission by Harry Whitmore – Cloud Jungle Epiphytes

**Figure 2.26** *Begonia pavonina* (peacock begonia, Begoniaceae) has shimmering, iridescent blue leaves due to iridoplasts that can enhance light capture and increase photosynthesis in the deep understory of Malaysian rainforests.



Courtesy of J. Gurevitch

**Figure 2.27** Sunflecks and a light gap in a tropical forest in the mountains of central Cuba.

## BOX 2D Blue Color and Iridescence, Structural Coloration, and Anthocyanin Pigments

**Structural coloration** is caused by the way light bounces off physical surfaces with particular optical properties. It can be created by simple or very complex microscopic structures in both plants and animals. In the case of iridescent blue leaves, the optical effect is caused by thin-film interference within the leaf, which results from multiple layers (usually in the cell walls) of materials that transmit or reflect particular wavelengths of light. Structural colors are not due to pigments, and when the objects are seen from underneath or from other angles, they may appear to be black or another color. Plants utilize a variety of structural coloration; Glover and Whitney (2010) provide information and background about this amazing phenomenon.

Similar optical interference or, alternatively, scattering of blue light by small particles (responsible for the blue of the sky; see Box 3B) is responsible for the structural blues found in many animals as well. The blues of butterfly wings, the brilliant blue colors found in some beetles, the vivid blue feathers of peacocks and the more modest blues of blue jays, the blue rump of the baboon, and the blue eyes of some humans and cats are all due to these structural, optical phenomena rather than to blue pigment. Structural iridescence is very widespread in multicellular marine algae, serving an antioxidation function and preventing damage from UV radiation. No blue pigment has ever been found in a vertebrate.

Blue colors in plant flowers and fruit are often produced by pigments called anthocyanins. There are hundreds of different anthocyanins. These pigments are water-soluble, antioxidant flavenoids that range in color from scarlet to purples and blues, and they can also be colorless. Their colors depend on the specific form of the molecule and are altered by pH and association or complexing with other phenolic compounds (see Chapter 11) or metals such as aluminum and iron (which produce the most intense blues). Anthocyanins are what cause red roses to be red; blue violets to be, well, blue violet; red grapes to be red or purple; blue potatoes and blueberries to be blue; and black tomatoes (actually very deep red) to be almost black. They are also what cause autumn leaves in some trees and shrubs to turn red. Their antioxidant properties may protect plant tissues from UV damage, and they are important in the ability of plants to attract pollinators (by flower color patterns) and seed dispersers (via fruits). They appear to have important health benefits for humans and other vertebrates (anticancer and heart health), but this may be due to more complex metabolic effects rather than simple antioxidant properties.

Pearcy (1991) and their colleagues suggest that while understory plants in both temperate and tropical forests are capable of surviving well in very low light, they depend on the energy captured from occasional sunflecks to be able to grow and reproduce. A large proportion (up to half) of the carbon gained by understory plants may come from energy captured from sunflecks.

The ability of these forest floor species to capture the energy available in brief, unpredictable episodes of direct bright light depends on several adaptations that make them able to respond quickly and with great efficiency to that light. These adaptations include stomata that open even in dim light, and a high electron transport capacity (a function of the light reactions) relative to carboxylation capacity (the capacity to carry out the initial reduction of $CO_2$). These plants also exhibit a slow loss of **photosynthetic induction** (the necessary start-up time for plants to reach maximum photosynthesis after exposure to bright light), so once "primed" by a sunfleck, they remain capable of more fully responding to other sunflecks.

The ability of some shade-adapted species to capture and use sunflecks, as well as light coming into the canopy at steep angles in the early morning and late afternoon, may be enhanced by their having irregularly oriented grana stacks, in contrast to the neatly aligned grana of sun leaves, which are stacked perpendicular to the leaf's surface. That irregular orientation would mean that at least some grana would be directly aligned with incoming light, regardless of the direction from which the light came. These plants may also have various adaptations that prevent them from being damaged by the intense sunlight in patches and flecks that may briefly illuminate the deep shade.

## Summary

- The process of photosynthesis consists of two parts: the light reactions, in which light energy is captured, and the Calvin-Benson cycle, in which carbon is incorporated into organic molecules.
- The reduction of carbon is powered by the light reactions, and the light energy captured in photosynthesis is stored in the chemical bonds of carbohydrates and other organic molecules.
- The rate at which photosynthesis proceeds depends on the amount and quality of the light reaching the chloroplasts, the $CO_2$ concentration gradient between the air and the chloroplasts, and conductance to $CO_2$.
- There are three different photosynthetic pathways: $C_3$, $C_4$, and CAM photosynthesis. $C_3$ photosynthesis is ancestral to the other types and is the most common. Each offers advantages under particular environmental conditions and is at a disadvantage under other conditions.
- Plant species are often adapted to high-light or low-light environments. These adaptations confer advantages in the environment to which the species are adapted, but trade-offs and costs cause them to function more poorly in other environments.
- Leaves on the same plant in some species may develop to maximize photosynthetic carbon and energy capture in either high light ("sun leaves") or low light ("shade leaves").
- Plants may use light as environmental cues for a variety of reasons in addition to capturing light energy in photosynthesis.
- Many plants are capable of detecting day lengths and changes in day length, and these cues can be reliable indicators of changing seasons.

Courtesy of NPS Brad Sutton

3

# Water Relations and Thermal Energy Balance

Plants, like everything else alive, need water to live, grow, and reproduce. What makes different kinds of plants able or unable to obtain sufficient water to survive and grow in different environments? Like animals, plants use water as the medium in which all of their biochemical reactions occur; unlike terrestrial animals, most plants also rely on water to a greater or lesser extent for support of their nonwoody parts and for maintaining their physical structure. Much of the evolutionary history of plants on land is the story of increasing success in obtaining water, moving it to parts of the plant far from its source, and being able to reproduce in a very dry atmosphere. We begin this chapter by briefly reviewing this history, pointing out some of the essential structures that evolved during the course of the history of land plants.

Terrestrial plants not only must obtain water, but also are faced with the problem of restricting their water losses. What determines water loss, and how do plants control it? Is the loss of water always harmful? Intimately related to this problem is the challenge of maintaining leaf temperatures within an acceptable range. What factors are most important in determining leaf and plant temperatures, and what adaptations do plants possess for controlling these factors so that temperatures do not get too high or too low? In this chapter, we examine their evolutionary solutions to these interrelated problems and take a look at the physics of water and energy exchange between plants and their environments.

Adaptation to environmental conditions such as drought and extreme temperatures can occur through the evolution of fixed phenotypic characteristics in physiology, morphology, anatomy, life history, or other traits, or through the evolution of phenotypic plasticity (the ability of an individual to change its phenotype in response to changing

Above: *Yucca brevifolia* (Joshua trees, Asparagaceae) are found in the Mohave desert of eastern California, and are the largest members of their genus; they coexist with xerophytic shrubs, annuals, and perennial grasses and can live up to 500 years.

conditions; see Chapter 9). Phenotypic plasticity can occur at a variety of levels and time scales. Over very short time scales, plants might change the osmotic potential of their leaf cells. They may open and close their stomata. Other short-term changes include altering the form of an enzyme so that it functions optimally at a different temperature. Longer-term changes include the production of leaves with different structural characteristics, or the growth of roots into areas in the soil with more water or nutrients. The ability to **acclimate** (acclimatize) to changes in environmental conditions is important for plants because mature plants are not able to move when conditions change: they can neither step into the shade nor migrate to a cooler climate when it is too hot (although their seeds or spores may move).

### The ancestors of modern plants evolved to live in terrestrial environments

Photosynthetic organisms originated in marine environments, where they were continuously bathed by water in which mineral nutrients were dissolved. The ancestors of terrestrial plants were single-celled aquatic organisms that depended on having each cell in direct contact with water, obtaining nutrients from and being kept moist by the water in their immediate surroundings. Like modern photosynthetic aquatic organisms, these ancestors of terrestrial plants also depended on water for sexual reproduction, releasing their gametes into the water.

Early land plants did not differ dramatically from their aquatic ancestors in many of these characteristics. One of the earliest adaptations of land plants was the evolution of a **cuticle**, a waxy, nonliving covering over the exposed epidermal cells that prevents desiccation by the air. Land plants, like all multicellular organisms, need to move water and nutrients among their cells and from one part of their bodies to other parts. In bryophytes, such as mosses, water moves by diffusion from cell to cell. There are no specialized organs to take up water or to transport it to distant parts of the plant. Diffusion is slow and not very efficient; as a consequence, none of the cells in a moss plant can get very far away from the substrate that provides the water, so mosses are almost always very short.

A plant's life cycle alternates between a gametophyte (which is haploid and produces sexual gametes) and a sporophyte (which is diploid and reproduces by spores; see Chapter 16). The dominant generation—the plants that are longest lived, largest, and most obvious to us—in bryophytes such as mosses is the haploid gametophyte. Male gametophytes produce male gametes, which must swim to female gametes. Because these plants are terrestrial, this movement is possible only when the moss shoots are covered with a film of water from rainfall or dew. The sperm swim along this film of water from a male plant to a female plant, eventually making their way to a structure (an archegonium) that contains an egg. This works successfully only under certain restricted weather conditions, and males and females must live very close to each other.

Early vascular plants, which branched off from the ancestors of bryophytes in the Paleozoic era (see Chapter 17), had a number of enormous advantages over bryophytes in terrestrial environments. One major advantage was that their vascular systems allowed the transport of water throughout the plant much more rapidly and efficiently. Because they had evolved vascular tissues, they were able to develop specialized structures—roots—to take up water from an extensive volume of soil and move it efficiently to other tissues. This enabled the successful evolution of woody tissue capable of supporting a trunk and crown. Without vascular tissues to take up water and conduct it over long distances, trunks, branches, and crowns could not exist. Vascular plants also evolved a dominant sporophyte (diploid) generation. As the sporophyte stage evolved to become larger and longer-lived, the gametophyte stage became smaller and shorter-lived. However, the plants in these early vascular groups still depended on motile sperm to swim from the male to the female gametophyte and had not evolved seeds.

The seed plants—seed ferns (now extinct), gymnosperms, and angiosperms—represent the greatest advances in the adaptation of plant life to a dry environment. The tissue that conducts water in these groups, the xylem, is capable of moving large volumes of water efficiently over long distances. Particularly in the angiosperms—the flowering plants—there is a great diversity of structure and physiological function related to the ability to withstand dry conditions.

The evolution of pollination removed many restrictions on the ability to reproduce in the dry conditions that dominate terrestrial life, including the requirements that male and female gametophytes be located adjacent to each other and that reproduction occur only under a very limited set of weather conditions.

The evolution of seeds was another key innovation. Nonseed vascular plants, such as ferns, disperse as haploid spores, which are produced by the diploid sporophyte plant. Both seeds and spores can disperse from the parent plant and can survive various environmental hazards, including desiccation. However, seeds have numerous advantages over spores. Seeds include tissues (such as the seed coat) that allow for the detection of subtle variation in the environment (such as changes in day length and temperature) and the hormonal regulation of germination in response to those environmental conditions. Seeds allow the storage of food for the embryo. Seeds may also contain toxic compounds that deter seed predators (see Chapter 11). Seed plants, therefore, were able to evolve far greater variation in dormancy and in maternal provisioning of stored food than their

non-seed ancestors and have numerous other advantages besides those. You may be wondering, How can it be that in angiosperms, the diploid sporophyte parent produces diploid sporophyte seeds? What happened to the gametophyte? The embryo inside the seed is indeed a diploid sporophyte, with genes from both maternal and paternal parents. This genetic background has genetic, ecological, and evolutionary advantages. But the gametophyte hasn't disappeared. The gametophyte has been reduced in the evolution of angiosperms to just a few cells inside the parent sporophyte (see Figure 6.5). The gametophytes produce megaspores and microspores that unite to form the new sporophyte that becomes an embryo inside the seed (the seed also contains maternal parent tissues and triploid tissues; see Chapter 6).

## 3.1 Water Potential Provides a Framework for Understanding How Plants Interact with Water in Their Environment

The abilities of a plant to acquire water, to move that water to cells throughout the plant body, and to lose that water through the stomata in the leaves all depend on the **water potential** of the various parts of the plant and its immediate environment (Figure 3.1; Box 3A). Water potential is a concept taken from the physics of thermodynamics. One way of thinking about water potential is that it represents the difference in potential energy between pure water (which is defined as having a water potential of zero) and the water in some system, such as that in a plant cell or in the soil. When water is bound up in some way, or has materials dissolved in it, the potential energy it possesses is less than that of pure water, and the water potential of the system is negative. Analyzing the movement of water in plants in terms of water potential puts the study of plant water relations into the unifying context of thermodynamics. It makes it possible to study the movement of water in soils, in plants, and between plants and the atmosphere using the same concepts, terms, and units.

Water potential ($\Psi$) in plants and soils can be broken down into four major components,

$$\Psi = \Psi_{\pi} + \Psi_{p} + \Psi_{m} + \Psi_{g}$$

where $\Psi_{\pi}$ is the osmotic potential, $\Psi_{p}$ is the pressure potential, $\Psi_{m}$ is the matric potential, and $\Psi_{g}$ is the gravitational potential of the water in the system. Water potential is a measure of pressure and is expressed in the SI units megapascals (MPa).

The **osmotic potential** is the component of water potential that is due to solutes dissolved in the water. It results from the difference in potential energy between pure water and water containing dissolved substances. Osmotic potential is either zero (in pure water) or negative because the solutes reduce the capacity of the water to do work (it takes more heat to warm up salty water than pure water, for instance.) The osmotic potential is capable of acting as a driving force for water movement, as when the movement of the solutes is restricted by a semipermeable membrane that allows water, but not solute molecules, to pass through it. The osmotic potential is the major component of water potential in living cells (such as those in roots and leaves) and is thus the major driving force by which water moves into and out of those cells.

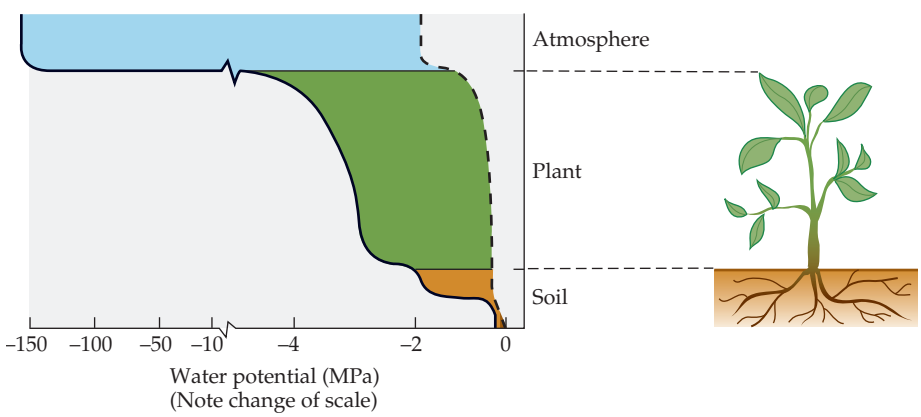



**Figure 3.1** Typical values (in megapascals, MPa) for water potentials in the soil, in a plant, and in the atmosphere in a mesic environment. The solid line at the left side of the colored region indicates the values for atmosphere, plant, and soil water potentials when the atmospheric relative humidity is low and the soil is dry. Under these conditions, soil water potentials range from close to 0 at deeper layers to about –1.8 MPa near the soil surface, plant water potentials range from about –2.0 to about –4.0 MPa, and atmospheric water potentials range from about –5.0 to –150.0 MPa. The dashed curve at the right side of the colored region indicates the values for moist soils and high atmospheric relative humidity; under these conditions, values for soil, plant, and atmosphere would be slightly less than 0 MPa, –0.01 to –0.75 MPa, and –1.0 to less than –2.0 MPa, respectively. Intermediate conditions would result in values falling within the colored regions; blue indicates the range of values for the atmosphere, green for the plant, and brown for the soil. (After J. R. Etherington. 1982. *Environment and Plant Ecology*. John Wiley and Sons, New York, NY.)

## BOX 3A Measuring Photosynthesis, Transpiration, and Water Potential

Chapter 2 and this chapter describe the complex systems plants use to extract $CO_2$ from the air and to limit the simultaneous loss of water. How do ecologists and plant physiologists measure these processes? Various types of technically sophisticated equipment have been developed over the past half century for these purposes. As you might imagine, there has been a revolution in making the devices smaller, more portable, and highly computerized over that time.

The most widely used methods for measuring photosynthesis involve direct measurement of the $CO_2$ concentration in a stream of air passing over a leaf or other photosynthetic tissue exposed to light (Field et al. 2000). These methods are based on the ability of $CO_2$ to absorb energy at wavelengths in the infrared part of the electromagnetic spectrum (this property is also what makes $CO_2$ a "greenhouse gas"; see Chapter 19). The gas sample is exposed to infrared light, and an electronic detector called an **infrared gas analyzer,** or **IRGA**, measures the amount of energy absorbed at each wavelength, which is proportional to the amount of $CO_2$ in the airstream. Measuring photosynthetic $CO_2$ uptake requires more than the IRGA itself. In the 1960s and 1970s, this equipment typically dominated an entire room, and "mobile" systems required a large, dedicated laboratory housed in a van. Modern systems are battery operated and designed for field-based measurements even in remote locations. Ecophysiologists can now take field measurements on entire populations, while two generations ago they could measure only a few individuals, usually in the laboratory.

Courtesy of Hans Verbeeck/PPSystems

Various commercial IRGAs and other instruments are small and lightweight, making it possible to conduct measurements on leaves in different field situations. This is a CIRAS-3 manufactured by PP Systems.

Besides measuring photosynthetic rates under a single set of conditions, scientists often want to estimate the carbon fixation rate, $A$, as a function of light intensity (resulting in **light response curves**; see Figure 2.3, Figure 2.13, and Figure 2.24), temperature, or intercellular $CO_2$ concentration, as shown in $\boldsymbol{A}$**-**$\boldsymbol{C_i}$ **curves** (see Figure 2.12). To facilitate these measurements, many systems come with built-in mechanisms to control light intensity, $CO_2$ concentration, and leaf temperature. Newer methods like flux towers are used to measure the $CO_2$ flux of large trees or even whole communities (see Chapter 5).

Transpiration rates, $E$, are often measured simultaneously with photosynthetic rates using an integrated IRGA system. The porometer is another tool that can be used to measure stomatal conductance and transpiration rates. The most common type of porometer is a diffusion porometer, which calculates leaf conductance to water vapor by the rate at which the air inside a small cuvette (chamber) becomes humidified by the diffusion of water from a leaf inserted into the cuvette. Stomatal diffusion is the major factor determining leaf conductance to water vapor, which controls leaf water loss.

An ecologist might also want to know the water potential ($\Psi$) of an entire plant, or of a part such as a leaf or branch. These measurements can be made with a relatively simple device called a Scholander pressure bomb (Scholander et al. 1965; Koide et al. 1989). The inventor, P. F. Scholander, is reported to have called his device a "bomb" because his early designs had a tendency to explode in his face; fortunately, designs have improved with time! When a leaf or stem is cut from a plant, water is pulled back into that structure by the negative water potential in the xylem. A pressure bomb works by measuring the pressure needed to push that water back out of the leaf or stem.

Water potential can also be measured using a psychrometer. This device has a small, thermally insulated chamber in which a leaf, or a disk cut from a leaf, is placed. Electricity is passed through a thermocouple to

Courtesy of J. Lee/USFS

Whole-canopy carbon flux can be measured using a variety of flux towers; these now exist in ecosystems all over the world. More information about these towers can be found at https://daac.ornl.gov/cgi-bin/dataset_lister.pl?p=9 and https://ameriflux.lbl.gov/about/about-ameriflux.

**BOX 3A** *(continued)*

lower the temperature until the dew point is reached. The rate of cooling of the thermocouple through evaporation of the resulting water droplet is used to calculate the internal water potential of the leaf. The psychrometer has the advantages of small size and that one can connect multiple psychrometers to a single data logger for measurement of populations in the field, but it has the disadvantage that larger plant parts or whole plants cannot be measured.

Water use efficiency (WUE), the ratio of carbon assimilation to water loss, can be measured over short or long times. Instantaneous WUE measurements are made with an IRGA. Long-term cumulative WUE can be determined by the ratio of $^{13}C$ and $^{12}C$ carbon isotopes ($\delta^{13}C_{leaf}$) in plant tissues; the more negative the $\delta^{13}C_{leaf}$ the lower the WUE, because the more efficient carbon capture is, the less carbon fractionation occurs and the more equivalent is the capture of $^{13}C$ and $^{12}C$.

The **pressure potential** is a second key component of plant water potential, and it is important both in living cells and in functioning xylem tissue (which is nonliving at maturity). It is a function of the hydrostatic or pneumatic pressure in the system and can be negative, zero, or positive in value. When water is enclosed by something that restricts its volume, such as a healthy plant cell enclosed by cell walls, the water presses on the walls, generating a positive pressure potential. This is the ordinary situation in a fully hydrated, or turgid, plant cell under normal conditions, in which the action of the water is similar to that of air inflating a balloon. The opposite situation occurs when water is pulled through an open system, such as a soda straw or a plant xylem vessel. In that case, the pressure potential is negative and actually pulls the walls of the vessel inward. Both positive and negative pressure potentials exist in plants, and both are important driving forces for the movement of water within the plant and between the soil, the plant, and the atmosphere.

**Matric potential** is a result of the cohesive force that binds water to physical objects, such as cell walls and soil particles; it is always negative because water bound to an object has less potential energy than free water. **Gravitational potential** results from the pull of gravity on water, and it is largely responsible for water draining from the largest pores of the soil in the first few days after a saturating rain (see Chapter 4). It is negative when the water is moving downward, since the water is losing potential energy. The matric and gravitational potentials are generally much more important in soils than in plants.

## 3.2 Water Moves through a Soil-Plant-Atmosphere Continuum

Water always moves from higher to lower values of Ψ (Keep in mind that a large negative number such as –100 has a lower value than a small negative number such as –5). In plants and in the soil, water is usually moving from a region of less negative water potential to one of more negative water potential. The gradient in water potential acts as the driving force for moving water from the soil, into and through the plant, and out to the atmosphere through transpiration.

The energy for moving water is provided without the expenditure of any energy by the plant itself, unlike in terrestrial multicellular animals, including humans, that need to use energy to pump blood throughout their bodies. Energy is still needed to lift water up from the soil to the top of the plant and to change the less energetic liquid water molecules inside the plant into the more energetic molecules of water vapor in the atmosphere. So where does this energy come from? The energy for moving the water comes directly from the sun, and the movement of water is driven by the gradient in water potential along this pathway. In this way, plants are able to move vast amounts of water, sometimes over tremendous distances (in large trees) at no direct energy cost to themselves, even though a huge amount of energy is required to do it.

In a plant growing in moist soil with leaves that have open stomata, the values for water potential and its components might look like those in Figure 3.2A. Water moves from the soil into the plant roots, which have a more negative Ψ, due primarily to the osmotic potential of the root cells. Then the water is drawn into the xylem vessels, which have even more negative Ψ values, due in this case to their negative pressure potential. Water moves into the leaves from larger to smaller veins, and then out into the substomatal cavity (Figure 3.3). From there, when the stomata are open, water moves out into the atmosphere. Even in humid air (Figure 3.2B), the atmosphere surrounding the leaf has a highly negative water potential. The liquid water molecules ultimately gain the energy needed to move into a vapor state from the energy in sunlight, and the water throughout the rest of the pathway is pulled upward by the transpiration stream escaping from the leaves of the plant to the atmosphere through the stomata.

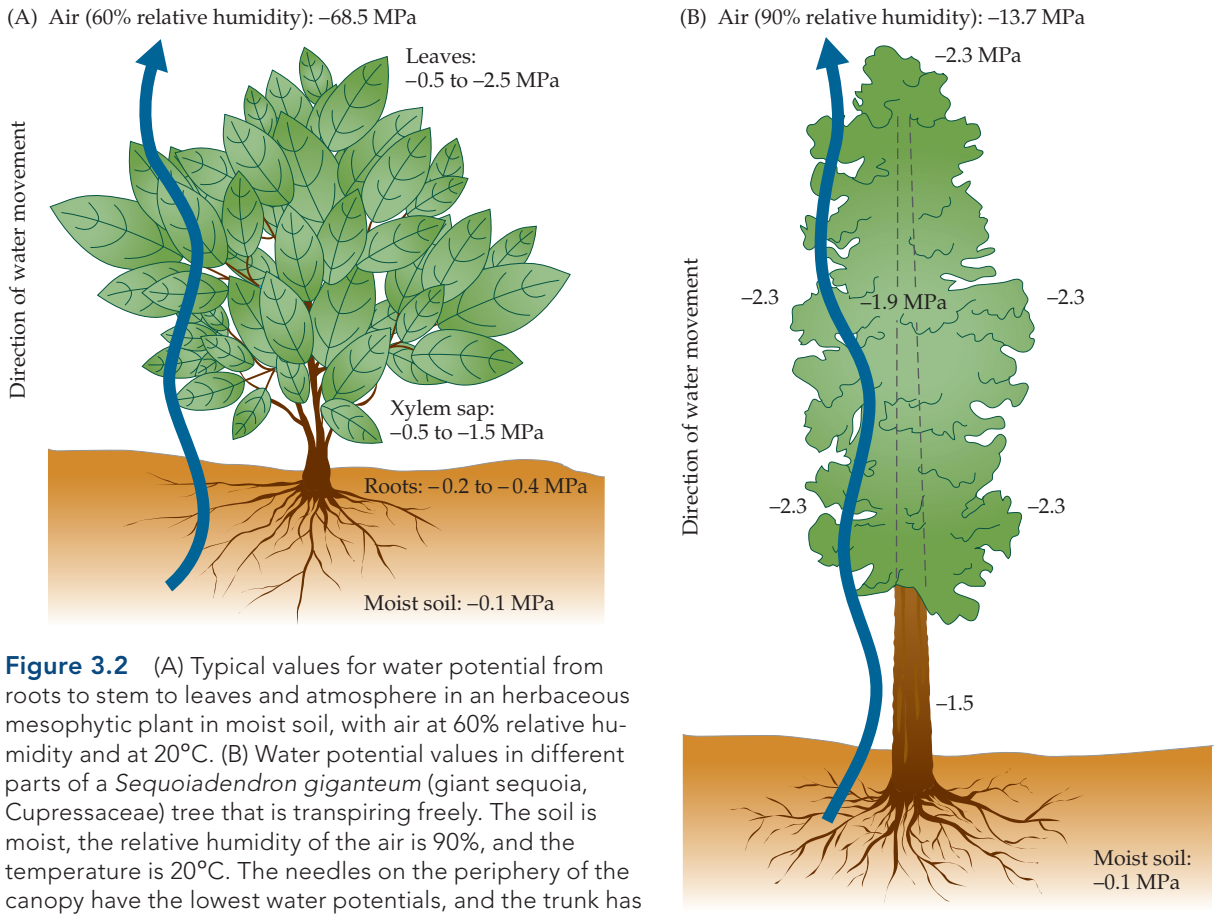



**Figure 3.2** (A) Typical values for water potential from roots to stem to leaves and atmosphere in an herbaceous mesophytic plant in moist soil, with air at 60% relative humidity and at 20°C. (B) Water potential values in different parts of a *Sequoiadendron giganteum* (giant sequoia, Cupressaceae) tree that is transpiring freely. The soil is moist, the relative humidity of the air is 90%, and the temperature is 20°C. The needles on the periphery of the canopy have the lowest water potentials, and the trunk has a higher water potential at the base than high in the canopy. (After H. Mohr and P. Schopfer. 1995. *Plant Physiology*. Springer: Berlin; C. A. Price. 1970. *Molecular Approaches to Plant Physiology*. McGraw-Hill: New York; H. Richter. 1972. *Ber Deutsch Bot Ges* 85: 341–351.)

Water moves upward in the soil by a combination of being pulled by the transpiration of plants, water cohesion, and capillary action, and it moves downward primarily by gravity. The cohesion is a consequence of the hydrogen bonds in water molecules, whereby the negatively charged $OH^-$ end of one water molecule is attracted to the positive $H^+$ end of another water molecule, holding them together. The capillary action is the result of the adhesion between the water molecules and the walls of the narrow spaces through which they are moving in the soil and subsequently within the xylem tissue of the plant. The movement upward occurs mostly via the small pores and micropores in the soil, and in the film of water surrounding soil particles, while the downward movement occurs largely via the largest pores between particles (see Chapter 4).

As water evaporates from the soil surface, capillary action and water cohesion pull water from deeper in the soil. The roots of transpiring plants also contribute to pulling water upward. Sometimes large amounts of water can be transferred from deep horizons to dry surface layers by deeply rooted plants. During the day, roots generally have more negative water potential than soils. At night, the dry upper layers of the soil may have more negative water potential than roots and may be capable of removing water from roots. In this way, deeply rooted plants may take up water from moist deep horizons, from which it travels up the root system and moves out of the upper part of the root system into the dry soil in the upper horizons at night. There, the water can become available to other, more shallowly rooted plants. This phenomenon, **hydraulic lift**, may in some cases allow the survival of these other plants during periods of drought or in arid environments. But water may be moved downward by roots as well.

**Hydraulic redistribution** refers to rapid water redistribution in the soil in any direction by roots, occurring mainly at night when transpiration is halted (Ryel et al. 2002). Hydraulic redistribution may alter the transpiration of the entire plant canopy substantially, particularly in arid environments, by making water available where it was not previously accessible (Ryel et al. 2003). In a study on the effects of hydraulic redistribution during normal dry seasons and extreme drought events in the Amazon rainforest, Guiling Wang and colleagues (2011) found community- and ecosystem-wide implications of

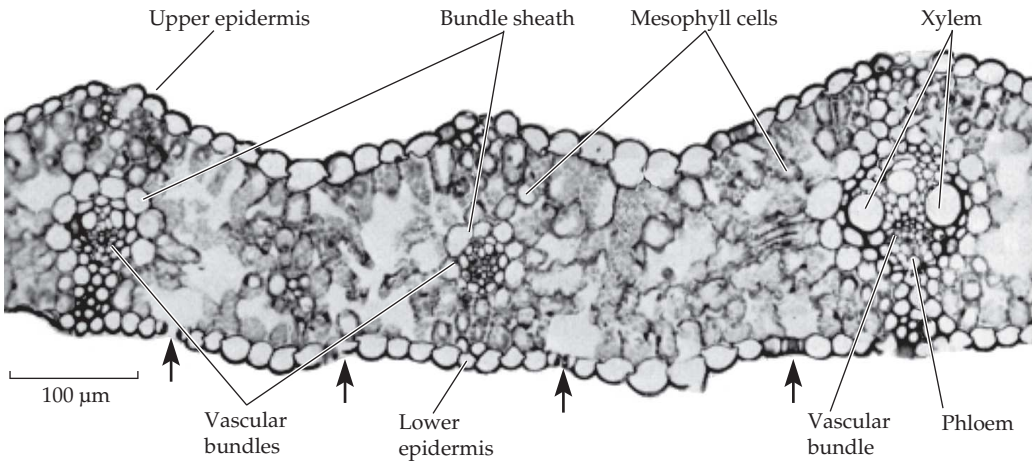



**Figure 3.3** Cross section of a leaf of $C_3$ grass *Hordeum vulgare* (barley, Poaceae), showing the upper and lower epidermis, which are each one cell thick; photosynthetic mesophyll cells; and vascular bundles each with xylem and phloem surrounded by a single-layered, nonphotosynthetic bundle sheath. Stomata are shown by arrows, and the substomatal cavities are visible inside the leaf, above the guard cells of the stomata. (From K. Esau. 1977. *Anatomy of Seed Plants*. John Wiley: New York; K. Esau. 1957. *Hilgardia* 27: 15–69.)

hydraulic redistribution. The Amazon forest is dominated by evergreen angiosperm trees that maintain green leaves throughout the year (see Chapter 18). Drought-deciduous trees, which lose their leaves in the dry season, are much less common in this region. During normal dry seasons, hydraulic redistribution increases water availability across much of the Amazon rainforest, resulting in increased transpiration and water loss from the soil. During extreme droughts, in contrast, soil moisture becomes severely depleted, hydraulic redistribution stops, and the evergreen trees lose their leaves. Because the drought-deciduous trees have already lost their leaves as a normal part of their response to the dry season, they are not as seriously affected, but the evergreen trees face slow recovery and sometimes death. There may be sudden losses of a substantial part of the tree canopy. Ultimately, the forest community may change to one where drought-deciduous trees become dominant, with major consequences for both the plant and animal communities.

## 3.3 Plants Manage Transpiration and Water Loss

Plants transpire vastly more water than they actually use metabolically. A large rainforest tree with a plentiful water supply might transpire as much as 1000 L of water in a day (weighing 1000 kg—a metric ton of water lifted from the soil, transported up the trunk, and through the leaves into the air). Temperate broad-leaved trees may lose up to 140 L of water per day, but a typical needle-leaved tree loses only 30 L in a day. While most plants transpire fairly modest amounts of water, that is not always the case: Jan Cermák and colleagues (1984) measured the loss of a remarkable 463 L per day from individual *Salix fragilis* (brittle willow, Salicaceae) plants growing in moist soil.

One reason that plants are capable of moving so much water from soil to air is that the water potential of air is so much lower than that of moist soil. Another reason is that large plants take up water from a large volume of soil and release it from a large surface area of leaves. It may at first be surprising that the volume of water released through transpiration can potentially far exceed the amount that would evaporate from an open water surface of comparable area. Consider, though, that the water is taken up and transferred ultimately to the air across some surface area. The surface area of the roots of all the trees taking up the water and the surface area of all the leaves losing the water in a 1 ha forest is far greater than the flat surface of a 1 ha lake.

The loss of water through transpiration can be described by a flux equation,

$$E = g_{\text{leaf(wv)}} \times [C_{i(\text{wv})} - C_{a(\text{wv})}]$$

where $E$ is the transpiration rate (in mmol $H_2O/m^2/s$). This equation is similar to the one that describes the diffusion of $CO_2$ into a leaf during the process of photosynthesis (see Chapter 2). The meters-squared measure refers to the surface area of the leaf, usually just the area of the underside of the blade but for leaves with stomata on both surfaces it includes both upper and lower surfaces of the leaf. The transpirational flux is driven by the difference between the concentration of water vapor inside the leaf, $C_{i(\text{wv})}$, and that in the external atmosphere surrounding the leaf, $C_{a(\text{wv})}$ (expressed in $mol/m^3$). Sometimes the vapor pressure of water is used instead of vapor concentration in this equation, and the difference between water vapor in the leaf and

the air is called the **vapor pressure deficit** (measured in kilopascals, KPa). The vapor pressure deficit of water is related directly to the water vapor concentration difference. The leaf conductance to water vapor, $g_{leaf(wv)}$, is a measure of how readily water vapor moves into or out of the stomata and is defined as the rate at which water vapor flows into the leaf for a given difference between ambient and intercellular concentration; it is expressed in mmol $H_2O/m^2/s$ and is the inverse of leaf resistance. The leaf conductance for water vapor flux consists of the boundary layer conductance ($g_a$), cuticular conductance (due to the waxy leaf cuticle; $g_c$), and stomatal conductance ($g_s$) to water vapor. Boundary layer conductance to water vapor is generally high. The thickness of the boundary layer of air surrounding a leaf depends on leaf size and shape as well as on ambient conditions, so it is not under the immediate control of the leaf and has little effect on impeding water vapor flux. Cuticular conductance is small; only limited water vapor leaves the leaf through the cuticle. In contrast, stomatal conductance is highly variable and is under the control of the plant; it is high when the stomata are open, and very low when the stomata close. Stomata are therefore very important in determining the amount of water lost by the plant through transpiration (Figure 3.4). Water exits the leaf through the same stomatal pores through which $CO_2$ enters in photosynthesis, so the movement of water vapor out of the leaf and $CO_2$ into the leaf are closely linked.

The loss of large volumes of water through transpiration does not have negative effects on the plant when soil water is freely available. Transpiration is what moves water, along with dissolved mineral nutrients, from the roots to the leaves without the need for the plant to expend energy. As the soil becomes drier, however, the plant encounters a dilemma. As $CO_2$ moves into the leaves through the stomata, water moves out through transpiration. A plant that keeps its stomata open faces desiccation, but a plant that keeps its stomata closed to restrict water loss eventually faces what is essentially starvation, as it can no longer take in $CO_2$. Angiosperms have evolved a greater diversity of mechanisms for coping with this dilemma, and for living with drought, than any other group of plants.

## Plants have different strategies for adapting to water availability

Plants experience drought in a variety of ways, depending on both the environmental conditions where they live and their adaptations to those conditions. Drought may be brief or prolonged, mild or extreme, occasional and unpredictable or regular in its onset and duration. It may occur when conditions are otherwise favorable for growth, or it may occur when it is too cold for plants to grow.

Plants have evolved various means of coping with limited water availability. These evolutionary solutions, or sets of coordinated adaptive traits, are known as **strategies** (discussed further in Chapter 10). Ecologists use different terms to categorize some of these sets of traits, but it is worth remembering that these categories are designated for convenience rather than being fixed, rigidly defined entities. In reality, each species has a unique combination of characters for coping with the conditions in its environment. One contribution to the global ecological and evolutionary dominance of the angiosperms is their unique leaf adaptations, with high densities of vascular tissue, high stomatal densities, and high rates of photosynthesis and transpiration (Baresch et al. 2019).

Photos © Walter Dioni, with permission from Microscape Magazine www.microscopy-uk.org.uk

**Figure 3.4** Stomata in the lower epidermis of (A) *Aptenia cordifolia* leaf (baby sun rose, Aizoaceae) and (B) *Rhoeo spatifolia* (Commelinaceae). Notice that the guard cells have chloroplasts, but the epidermal cells are clear with no chloroplasts.

**Mesophytes** (Figure 3.5) are plants that live in moderately moist soils and generally experience only an occasional, mild shortage of available water. Mesophytes typically transpire as long as soil Ψ is greater than about –1.5 MPa, although some need soils wetter than –1 MPa, while others may be able to continue withdrawing water from the soil at water potentials as low as –4 MPa. When the soil becomes drier, mesophytic plants close their stomata and wait until conditions improve. In a

(A)

(B)

Photos courtesy of J. Gurevitch

**Figure 3.5** Mesophytes grow where soil water is usually available and droughts are mild and short-lived. (A) *Solidago canadensis* (Canada goldenrod, Asteraceae) in a meadow on Long Island, New York, U.S.A. (B) *Rosa acicularis* (Alberta wild rose, Rosaceae) in a montane woodland in Banff National Park, Alberta, Canada.

typical **mesic** (moist) habitat, that may be a few hours or a few days. If the drought extends for many days (a few weeks is all that most of them can tolerate), many mesophytes begin to die. Most agricultural crops, ornamental plants, forest trees, common meadow and forest wildflowers, and other plants common in temperate parts of the world are mesophytes. Agricultural crops are generally the least drought tolerant and need soils that do not dry out too much in order to grow well.

**Aquatic plants** grow in standing water, and **hygrophytes** are found in permanently moist soils. Plant ecologists and amateur naturalists alike have always been fascinated with the adaptations of **xerophytes** (Figure 3.6), plants that live in regions with frequent or extended drought. Some xerophytes continue to transpire in soils that are phenomenally dry, as low as –6 MPa. One of the earliest experimental studies of the seasonal water relations of desert plants was conducted by Edith Shreve (1923), who contrasted the strategies and stresses experienced by annual, perennial, and cultivated species.

Courtesy of Desert Laboratory on Tumamoc Hill

Edith Shreve

(A)

Courtesy of Sonie Sampson

(B)

Courtesy of J. Gurevitch

**Figure 3.6** (A) Two xerophytes growing in the Mojave National Preserve in eastern California, U.S.A.: *Larrea tridentata* (creosote bush, Zygophyllaceae), the larger green shrub, and *Ambrosia dumosa* (burro bush, Asteraceae) the small grayish shrub. *Larrea* is able to tolerate extreme drought as a mature plant. It has small, sclerophyllous, resinous leaves that give the desert flats and alluvial fans where it lives a distinctive and intense scent, particularly when it rains. (B) *Artemesia tridentata vaseyana* (mountain big sagebrush, Asteraceae) is a xerophyte common in the cold desert, steppe, and mountain habitats in the Great Basin of North America.

Marco Vinicio Olla/CC BY-SA 4.0

**Figure 3.7** *Salicornia europea* (glasswort, Amaranthaceae) is a succulent halophyte. *Salicornia* species are found all over the world in saline coastal soils in salt marshes, under coastal mangroves, and on beaches in North America, Europe, South Africa, and South Asia.

**Halophytes** (Figure 3.7) are plants that live in saline soils and possess unique adaptations for tolerating salt, as well as sharing some characteristics with xerophytes. Halophytes experience a dilemma similar to that of xerophytes in that plants in both dry and saline soils must be able to extract water from soils with very negative water potentials. Mohamed Sankary and Michael Barbour (1972) studied the water relations of the halophytic shrub *Atriplex polycarpa* (desert saltbush, Amaranthaceae). They found that it could survive and recover from remarkably low leaf water potentials—lower than –5 MPa. Even under well-watered conditions, its leaf water potentials were as low as –2 MPa—a condition that would be fatal to most mesophytes.

## Water use efficiency is a measure of carbon gain versus water loss

As plants take up $CO_2$ through their stomata, water is lost. The ratio of the gain of carbon in photosynthesis to the loss of water in photosynthesis is the instantaneous **water use efficiency** (WUE) of carbon assimilation. The rate of evaporation of water through the stomata is greater than that of $CO_2$ uptake for several reasons. The gradient in water vapor density from within the leaf to the air is about 20 times steeper than that for $CO_2$ entering the leaf, and water molecules are smaller and diffuse more quickly than $CO_2$ molecules. In addition, $CO_2$ must dissolve upon entering the leaf and diffuse all the way into the chloroplast, whereas water vapor evaporates out directly from the moist substomatal cavity.

Water use efficiency is highest for $C_3$ plants when their stomata are partially open. $C_3$ plants can also increase their water use efficiencies by increasing the concentration of rubisco and other photosynthetic enzymes (see Chapter 2), as more rapid fixation of carbon leads to a steeper diffusion gradient for $CO_2$ entering the leaf. The water use efficiencies of $C_4$ plants are much higher than those of $C_3$ plants because $C_4$ plants are more efficient at capturing carbon. CAM plants have the highest water use efficiencies, due to the combination of the efficiency of the $C_4$ step in capturing carbon and their greatly reduced water losses. Water loss in CAM photosynthesis results from decoupling carbon uptake and fixation, which allows CAM plants to open their stomata only at night when the air is cooler and more humid.

On a whole-plant or canopy basis, the WUE of productivity is defined as the total dry matter production over the course of a growing season for a plant or plant stand, divided by the total water loss. There are substantial differences among different crops, and among different natural communities, in this measure of WUE. Many factors influence WUE at this larger scale, including stand density, environmental conditions, the functional group of the plants, and the physiognomy of the vegetation (see Chapter 18).

## Plants have different adaptations for coping with reduced water availability

A plant experiences its environment very differently than one might expect from a cursory look at the average conditions in that environment. Annuals in arid environments, particularly short-lived ephemerals, complete their entire aboveground lives during the brief rainy season characteristic of these habitats. These plants spend most of their lives as seeds, which are able to survive extended dry periods. The seeds germinate only after prolonged rains, and the plants grow and photosynthesize rapidly, have high rates of transpirational water loss, and set seed and die when the rainy season ends and the soil dries out (Figure 3.8). By "living fast and dying young," they experience the desert as an environment with plentiful

Courtesy of Gordon A. Fox

**Figure 3.8** Annual plants flower in very large numbers when sufficient moisture is available in springtime, in Namaqua National Park, Northern Cape, South Africa. These are mainly *Osteospermum pinnatum* (African daisy, Asteraceae).

water for most of their lives as green plants. This strategy has been called **drought avoidance**.

Many desert shrubs and trees lose all or some of their leaves during the long dry season and grow new leaves after the rains begin again. These **drought-deciduous** plants (Figure 3.9A) are also practicing a form of drought avoidance, minimizing their water loss through transpiration by reducing their leaf area when water is not available. Seasonal tropical forests (see Figure 18.13), which also experience predictable dry periods alternating with rainy seasons, have many species of drought-deciduous trees (Figure 3.9B). Some evergreen species have a similar strategy: They lose only some of their leaves during dry periods, retaining some photosynthetic capacity but reducing overall water loss. Sandra Bucci and colleagues (2005) investigated the water status of six evergreen or largely evergreen woody savanna species in Brazil over the course of the 5-month dry season. They found that the plants all maintained almost unchanging minimum leaf water potentials despite the large seasonal change from high rainfall to almost no rainfall and the large changes in atmospheric evaporative demand on the plants. The plants kept a consistent water status from the wet to the dry season by shedding a substantial proportion of their leaves, reducing both their total leaf area and the stomatal conductance to water vapor in their remaining green leaves.

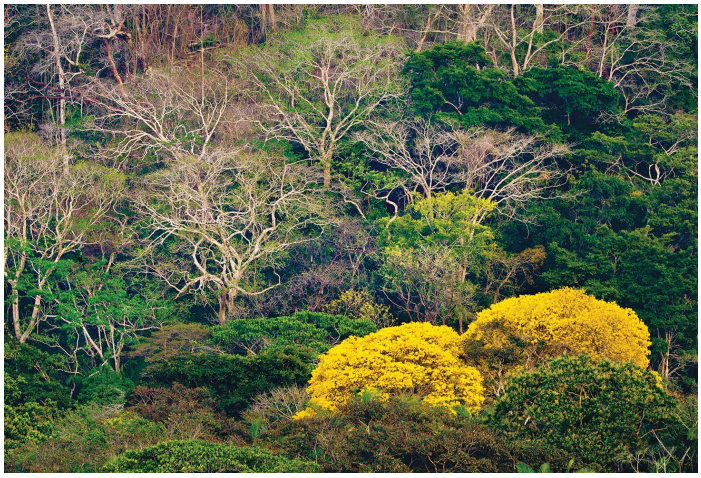
Many herbaceous perennials in deserts, arid grasslands, and other **xeric** (dry) environments are largely dormant during the dry season, when most of their living

(A)

Both photos https://tedRmyers.com

Both photos https://tedRmyers.com

(B)



©Panama Landscapes by Oyvind Martinsen/Alamy Stock Photo

**Figure 3.9** (A) *Fouquieria splendens* (ocotillo, Fouquieriaceae), a drought-deciduous woody plant of the Sonoran Desert found in southern Arizona and northern Mexico. Left, without leaves in dry conditions; right, with leaves and flowers in rainy season. Insets show the leafless stems and stems with new leaves and a red inflorescence. (B) *Tabebuia guayacan* (tabebuia, Bignoniaceae). Trees flowering in the dry season in Soberania National Park, Republic of Panama. The yellow flowering trees are *Tabebuia guayacan* (Guayacan trumpet tree or tabebuia, Bignoniaceae). Many tropical trees flower during the dry season; water loss from the flowers is limited because they are short-lived.

material is found below ground or at the ground surface. The aboveground parts of many perennial grasses, for example, die back to the ground during the dry season. The advantage of dormancy during the dry season is, of course, reduction of water loss, enhancing survival during the unfavorable dry season. Its major disadvantage is that there is a long time during which the plant cannot photosynthesize, grow, or reproduce. Plants that use this drought avoidance strategy generally have higher maximum photosynthetic and growth rates during the time when they are active than drought-tolerant plants, which help to compensate for the time when they are dormant. Many of these plants also have underground rhizomes (underground stems), tubers, bulbs, or thick roots that are capable of storing food (primarily carbohydrates) and some water for long periods of time. These structures enable them to survive extended drought periods and allow rapid production of leaves and stems when conditions become favorable again. Another way in which some species avoid drought is to grow only in the wettest places within a dry environment, such as in low spots where water temporarily collects after rains or along temporary streams. Such plants typically lack adaptations for extreme aridity, and if their seeds germinate in drier microhabitats, the seedlings will die.

There are, however, plants that are able to live and grow under remarkably dry conditions. These **drought-tolerant** species, or true xerophytes, use a variety of means to "tough it out" in arid environments. From the whole-plant to the cellular and molecular levels, the morphology, physiology, and anatomy of these plants are adapted for life in xeric habitats. Even these species, however, depend on the brief time of the year when water is available for most of their growth and for reproduction. Forrest Shreve (1911, 1917) was one of the first ecologists to carefully document the establishment of seedlings of desert plants. He found that for a wide range of perennial species living in the North American Sonoran Desert, germination occurs only during the rainy season, but most seedlings are not able to survive the following dry periods. As Shreve and many subsequent researchers have found, the successful establishment of most desert perennials is a relatively rare event.

Courtesy of Desert Laboratory on Tumamoc Hill

Forrest Shreve

Plant form can also be an adaptation to limited water availability. **Succulents**, such as cacti, agaves, and yuccas (Figure 3.10), have the capacity to store large amounts of water in their tissues. The succulent tissues may be located in leaves, stems, or other plant parts. The stems of cacti are their photosynthetic organs (the spines are actually nonphotosynthetic, highly modified leaves). As a result, cacti have reduced surface-to-volume ratios, decreasing the total surface area capable of water loss and increasing the volume that retains water. Many succulent species (but certainly not all) use CAM photosynthesis and have low

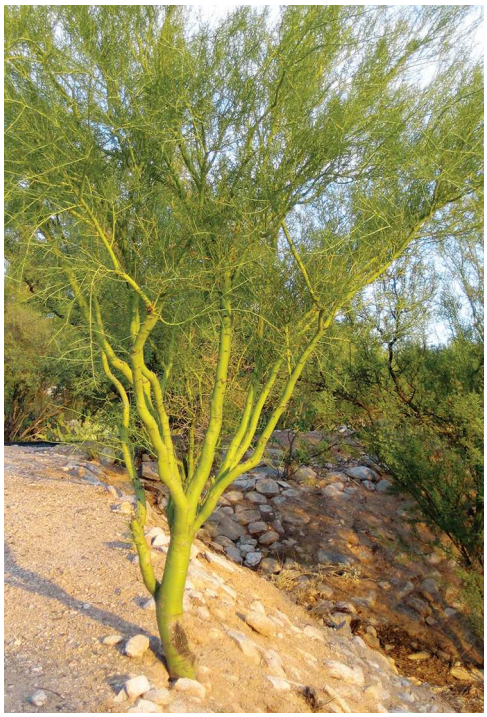
(A)

© Pete Gregoire

(B)

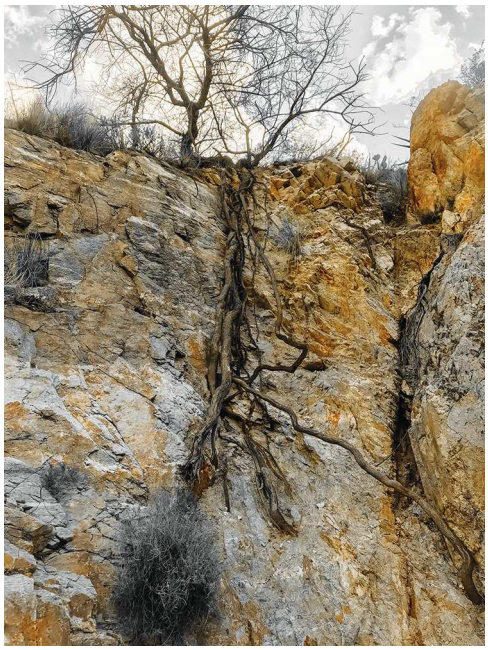
Courtesy of Charlie McDonald, USDA

**Figure 3.10** (A) *Carnegiea gigantea* (saguaro, Cactaceae) is a long-lived succulent that characterizes the upper Sonoran Desert in Arizona, U.S.A., and Sonora, Mexico. It generally is found on rocky slopes, and occasionally on sandy flats. (B) *Agave parryi* (Parry's agave, Asparagaceae) is a monocarpic succulent with CAM photosynthesis. These agaves are in an oak-juniper woodland in the Gila National Forest in southwestern New Mexico, U.S.A.



Courtesy of Jaime McGuigan

**Figure 3.11** *Parkinsonia microphylla* (foothill palo verde, Fabaceae) has extremely small leaflets, which are lost during the dry season, but photosynthesis continues at low rates in the green bark.



©Johnny Coate/Shutterstock.com

**Figure 3.12** *Prosopis velutina* (mesquite, Fabaceae) in Sabino Canyon, Arizona, U.S.A. is a phreatophyte with roots that may go down 30–50 m to reach water.

maximum photosynthetic rates. Other desert plants also may use organs other than leaves to carry out photosynthesis. For example, trees in the genus *Parkinsonia* (palo verde, Fabaceae) (Figure 3.11; and see Figure 2.8) in the Sonoran Desert retain their leaves only during the rainy season but remain capable of photosynthesis in their green bark at low rates during dry periods of the year.

Root morphology can be adapted to dry environments in a variety of ways. Roots can be extensive and shallow, as they are in many cactus species, enabling the plant to take up water after a brief rain that saturates only the top layers of soil. Other species have roots that extend deep into the soil, where it is slow to dry out. **Phreatophytes** (Figure 3.12) have roots that extend so deeply into the soil that they may even reach the water table, gaining relatively permanent access to water—the ultimate drought avoidance mechanism.

The ratio of dry mass invested in roots to the mass of aboveground tissue (primarily leaves and stems) is called the **root-to-shoot ratio**. It varies widely among plant species and may even vary for the same species in different environments. Xerophytes typically have much higher root-to-shoot ratios than do mesophytes—most of the plant may be below ground. Consequently, they are able to take up more water, and they lose less water through transpiration. Annual plants usually have much lower root-to-shoot ratios than perennial plants. Annuals invest heavily in growth and reproduction, while perennials store material below ground to regrow in future years. The cost of a high root-to-shoot ratio may be a reduction in maximum growth rates, both because of the reduction in total leaf area, and therefore in the amount of photosynthetic tissue, and because of the metabolic cost of maintaining a large root mass.

In parts of the world where the soil freezes in winter, it can be difficult for plants to obtain liquid water from the soil for many winter months. Plants in these environments can experience severe drought stress in winter. This is one reason why deciduous trees and shrubs in the northern temperate zone and farther north lose their leaves in winter, and it is one of the factors that sets the northern boundary of the distribution of broad-leaved evergreen species.

## Plants have complex physiological adaptations to drought

When a plant begins to experience a water deficit, a series of physiological events is triggered. Hormones are produced and travel throughout the plant, signaling the onset of various changes in plant function. Cell growth and most protein synthesis slow, then cease. As the plant experiences more prolonged water deficits, the allocation of materials to roots and shoots is adjusted, the stomata begin to close, and photosynthesis is inhibited. Leaves may begin to wilt. In some species, older leaves are allowed to dry out and die, freeing up

available water for younger leaves that have a greater photosynthetic capacity.

In some species, specific measures for **osmoregulation**, the regulation of osmotic potential, are initiated under dry conditions. During drought these plants synthesize certain soluble compounds such as proline, other low-molecular-weight nitrogen compounds, and soluble carbohydrates. The resulting increase in solutes lowers the osmotic potential of the cells, leading to an influx of water by osmosis, which prevents turgor loss and wilting. In saline soils, soil water has a negative water potential largely due to a negative osmotic potential caused by dissolved salts. For many halophytes growing in these soils, osmoregulation is particularly important in maintaining a favorable water potential gradient. Other halophytes have the ability to excrete salt.

Some plants, known as "resurrection plants," have highly unusual adaptations that allow them to survive complete, extended desiccation (Figure 3.13). These plants include many lichens and some mosses, ferns, and other non-seed plants, as well as a number of angiosperms in the Scrophulariaceae, Lamiaceae, Poaceae, and Boryaceae (monocots endemic to Australia) and lycopods in the Selaginellaceae. They are found in various parts of the world but are most diverse and abundant in southern Africa. All of these different plants survive cellular dehydration by a highly coordinated set of processes. Drought-stable proteins are synthesized, phospholipid-stabilizing carbohydrates are incorporated into the cell membranes, and the cytoplasm may become gelated (turned into a gel state). With the plant in the dehydrated state, metabolism is brought almost to a halt. Rehydration when water is again available also occurs as a highly coordinated process in which the cell components are reconstructed step by step. The mechanisms by which these processes are carried out have been the subject of research among plant molecular physiologists (e.g., Petersen et al. 2012; Griffiths et al. 2014).

At the other end of the water availability spectrum, adaptation to flooding is critical for survival in some habitats (Figure 3.14). Plants may be subjected to variation in the depth to which they are submerged (from waterlogged soil around the roots to total submergence of the shoot) as well as in the frequency, season, and duration of exposure to flood conditions. Coastal species in habitats such as salt marshes and species found in floodplains and at river and stream margins are predictably exposed to flooding and adapted to tolerate it. If, however, the frequency and severity of flooding become greater (see Chapter 16), these new conditions may overwhelm the ability of the plants to tolerate them. Other habitats, including low-lying areas in forests, meadows, and fields, also expose plants to occasional submersion, particularly in sites with clay soils that do not drain readily after heavy rains.

When the soil is flooded, plants are faced with a number of conditions that make it difficult for them to function normally, or even to survive. The greatest problem is a lack of soil oxygen. In ordinary, well-drained soils, oxygen diffuses through the soil to the roots, which use it for respiration. In waterlogged soils, the diffusion of oxygen is severely diminished because the soil pores are filled with water rather than air. Longer-term exposure to flooding results in compaction of soils. When this happens, larger aggregates of soil particles collapse,

© Wayne P. Armstrong

**Figure 3.13** *Selaginella lepidophylla* (spike moss, Selaginellaceae) is a "resurrection plant." The plant has biochemical mechanisms that allow it to slow its metabolism to a virtual halt during an extended drought. Although the brown and desiccated plant (left) appears to be dead, the advent of rainfall rehydrates the tissues, which quickly become green and photosynthetic (right).

Courtesy of S. Scheiner

**Figure 3.14** An intermittently flooded swamp in southern Illinois is dominated by *Taxodium distichum* (bald cypress, Taxodiaceae). This species produces "knees" (visible in the foreground) that project from the root system upward above the water to facilitate root gas exchange.

leading to a reduction in soil pore space and soils that are harder for growing roots to penetrate (see Chapter 4). Any oxygen that remains in the soil may then be quickly consumed by plant roots and soil organisms. Oxygen-dependent microbial processes are severely inhibited, and toxic substances from bacterial anaerobic metabolism begin to accumulate.

Plants that have evolved to tolerate flooding exhibit a variety of physiological, morphological, anatomical, and life history characteristics that allow them to function in flooded environments. One physiological mechanism for surviving for an extended time with little or no oxygen available to the roots is an increased reliance on glycolysis. The NADH that is generated by glycolysis can be oxidized to NAD by ethanolic fermentation, allowing some ATP production to proceed. The oxygen deficiency caused by flooding also stimulates a series of metabolic responses, including the production of polypeptides called anaerobic stress proteins. The molecular genetics and physiology of these adaptations are now beginning to be understood (e.g., Bailey-Serres and Voesenek 2008; Liang et al. 2019).

## The anatomy and physiology of stomata shape plant responses to water loss

A wide variety of anatomical and morphological adaptations enable plants to survive and grow under very dry or very wet conditions. Among the most important are variations in stomatal number, arrangement, size, and behavior, which vary widely among plant species. Stomatal number can even vary for the same plant under different growing conditions. Most of the water that a plant loses goes out of the stomata as water vapor.

Small amounts of water can escape through the waxy cuticle on the leaf epidermis. Structures such as flowers may sometimes lose considerable amounts of water. Other green plant parts with stomata (see Figure 2.8), like green bark, sepals, and the green skin of fruit, will typically lose much less water than the stomata in leaves. Stomatal conductance to water vapor (as well as to $CO_2$) varies directly with pore width (the size of the stomatal opening). Pore width is controlled by the **guard cells** on either side of the opening, which continually change their shape, widening or narrowing the pore opening (see Figure 3.4).

The guard cells open and close the stomata primarily in response to three stimuli: light, $CO_2$ concentration, and water availability. (Because leaves are vulnerable to pathogens entering through the stomata, the guard cells may also close in response to an attempted entry by a pathogen.) The response to light is an active process that occurs by proton pumping from ATPase, and this occurs throughout land plants and is likely to have evolved very early. Stomatal opening and closing in response to $CO_2$ is an inherent characteristic of the guard cells. Those cells respond to their own internal concentration of $CO_2$ by altering membrane potassium channels and internal ion concentrations and by polarizing the guard cells' membrane potentials (Ainsworth and Rogers 2007). These changes result in reduced turgor pressure that causes stomata to close. The response to water deficits, in contrast, occurs by a variety of mechanisms in different species and is controlled by both active and passive mechanisms. Passive closure of stomata in response to water deficit occurs as part of the loss of turgor in the epidermis and guard cells due to water deficits. Active closure in response to drying is a result of a series of reactions mediated by the plant hormone abscisic acid (ABA) that cause ions to move out of the guard cells, causing a loss of turgor.

While the evolution of the stomatal response to water deficits is controversial, evidence suggests that it evolved after seed plants evolved (Sussmilch et al. 2017). The evolution of better regulation of water loss in the earliest seed plants may have been a major factor contributing to their increasing dominance and the decline of ferns and lycophytes toward the end of the Paleozoic (McAdam and Brodribb 2012) (see Chapter 17).

The action of the guard cells differs among plants with $C_3$, $C_4$, and CAM photosynthetic pathways (see Chapter 2). Light stimulates the stomata of $C_3$ and $C_4$ plants to open. The partial pressure of $CO_2$ in the intercellular leaf space has a major influence on the guard cells, signaling stomata to close as the partial pressure goes up and to open when it goes down. At night, when photosynthesis ceases in $C_3$ and $C_4$ plants, the $CO_2$ concentration in the intercellular space increases, and the stomata close. During the day, photosynthetic capture of $CO_2$ reduces the partial pressure of $CO_2$ in the intercellular space, and the stomata open. In contrast, for CAM plants, the partial pressure of $CO_2$ drops at night as $CO_2$ is captured by PEP carboxylase, and the stomata open. During the day, organic acids are decarboxylated ($CO_2$ is removed) and accumulate in the leaf, signaling the stomata to close (see Figure 2.14).

Plant water relations are critical in determining the behavior of the stomata. Stomata respond to both ambient humidity and leaf water potential (Figure 3.15). A declining water potential in the leaf will override other factors, such as $CO_2$ partial pressure, to close the stomata. Desiccation can cause immediate death, so preventing it is more pressing than maintaining photosynthetic rate.

Typical leaf anatomy of mesophytes include elongated palisade cells, responsible for much of the photosynthesis; photosynthetic spongy mesophyll cells adjacent to large substomatal cavities where air enters from the stomata; and a single layer of non-photosynthetic epidermis and veins consisting of xylem and phloem (Figure 3.16). The stomata of different plant species vary in their sensitivity to the three factors that control stomatal opening and closing—light, $CO_2$ concentration, and water availability. Mesophytes, particularly agricultural crop species, often

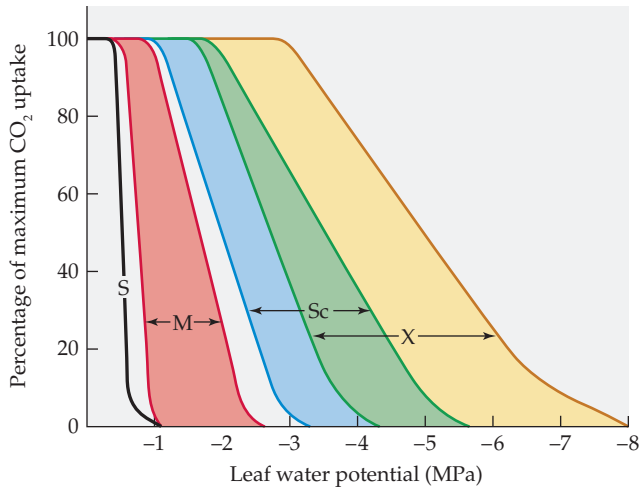



**Figure 3.15** Stomata respond to a decline in leaf water potential by closing either quickly as the soil dries, or more slowly, depending on species and habitat; this results in a reduction in photosynthesis. The graph shows the decline in photosynthetic rate, expressed as the percentage of maximum $CO_2$ uptake, with decreasing leaf water potential in plants with different strategies. Succulents, S, close their stomata most readily with the onset of drought; mesophytes, M, follow (range of water potential values shown in red shading); sclerophyllous plants, Sc (trees and shrubs of mediterranean and other semiarid climates, blue and green shading) and xerophytes, X (herbaceous dicots, grasses, and shrubs of deserts and other very arid environments, green and yellow shading), close their stomata at the most negative values. Sclerophylls and xerophytes have partially overlapping values. (After W. Larcher. 1995. *Physiological Plant Ecology*, 3rd ed. Springer: Berlin. Data from numerous authors.)

close their stomata during the middle of the day when water losses would be greatest, even in moist soil. The stomata of xerophytes behave differently than those of mesophytes. Most mesophytes quickly close their stomata as soon as the soil dries and leaf water potential drops. In contrast, many xerophytes are able to keep taking up water from the soil, transpiring and photosynthesizing well into a period of drought (see Figure 3.15). The stomata of xerophytes typically remain open (although at low conductance values) at much lower water potentials and for a longer period of time in drying soil in comparison to mesophytes. Mesophytes' stomata usually have much greater maximum pore widths, higher maximum stomatal conductances, and higher maximum rates of both transpiration and photosynthesis than those of xerophytes (Table 3.1).

In environments where dry conditions alternate with a predictable wet season, some plants are able to photosynthesize and grow rapidly during periods when water is available, and shut down when it is not. By maximizing carbon gain when water is most available and stomata can open without costly water loss, the plant gains the most carbon and energy over the course of the year without adverse consequences.

## Leaf anatomy can be adaptive for survival and growth in arid environments

Specializations in leaf anatomy can help some plant species achieve high rates of photosynthetic carbon uptake during favorable periods. Many plants native

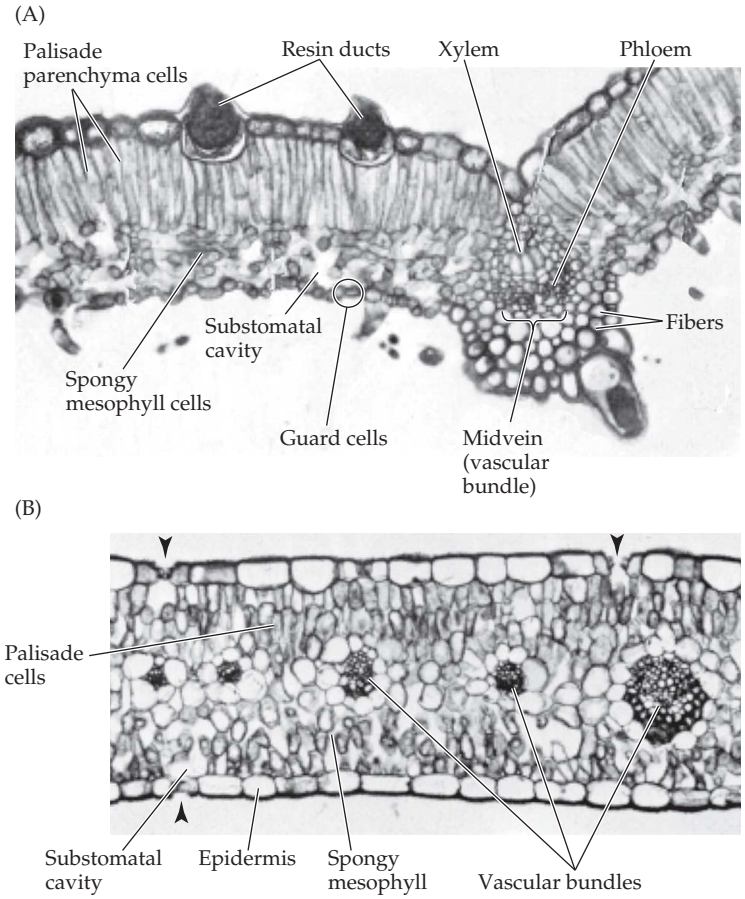



**Figure 3.16** Cross sections showing leaf structure in two mesophytes. (A) *Cannabis sativa* (marijuana, Moraceae). The photosynthetic cells are the elongated palisade parenchyma cells and the smaller, irregularly spaced spongy mesophyll cells. The midvein consists of xylem and phloem cells and has a protective cap of fiber cells. (B) A leaf of *Dianthus* sp. (pink, Caryophyllaceae). The micrograph shows a large vein (at right), several smaller veins, and palisade and spongy mesophyll cells. There are stomata on both the upper and lower surfaces of the leaf (arrowheads)—a condition more typical of xerophytes than of mesophytes such as this species. (From K. Esau. 1960. *Anatomy of Seed Plants*. John Wiley: New York.)

**TABLE 3.1** Properties of stomata on the lower leaf surface in plants from mesic and xeric environments

| Type of plant | Stomatal density (number/mm$^2$ leaf area) | Pore length (μm) | Maximum (μm) pore width | Pore area (% leaf area) | Maximum conductance for $CO_2$ (mmol/m$^2$/s) |
|---|---|---|---|---|---|
| Mesic environments | | | | | |
| Herbaceous plants (sunny habitats) | 150 (200) | 15 (20) | 5 | 0.9 (1.0) | 425 (700) |
| Herbaceous plants (shaded habitats) | 75 (100) | 23 (30) | 6 | 1.0 (1.2) | 130 (200) |
| Grasses | 75 (100) | 25 (30) | 3 | 0.6 (0.7) | 270 (460) |
| Tropical forest trees | 400 (600) | 18 (25) | 8 | 2.3 (3.0) | 100 (300) |
| Deciduous trees | 200 (300) | 10 (15) | 6 | 0.9 (1.2) | 150 (250) |
| Xeric environments | | | | | |
| Xerophytic desert shrubs | 225 (300) | 13 (15) | — | 0.4 (0.5) | 155 (260) |
| Sclerophylls | 300 (500) | 13 (15) | 2 | 0.4 (0.5) | 175 (250) |
| Succulents | 33 (50) | ca. 10 | 10 | 0.3 (0.4) | 95 (130) |
| Evergreen conifers | 80 (120) | 18 (20) | — | 0.7 (1.0) | 200 (—) |

*Source*: After W. Larcher. 1995. *Physiological Plant Ecology*, 3rd ed. Springer: Berlin.

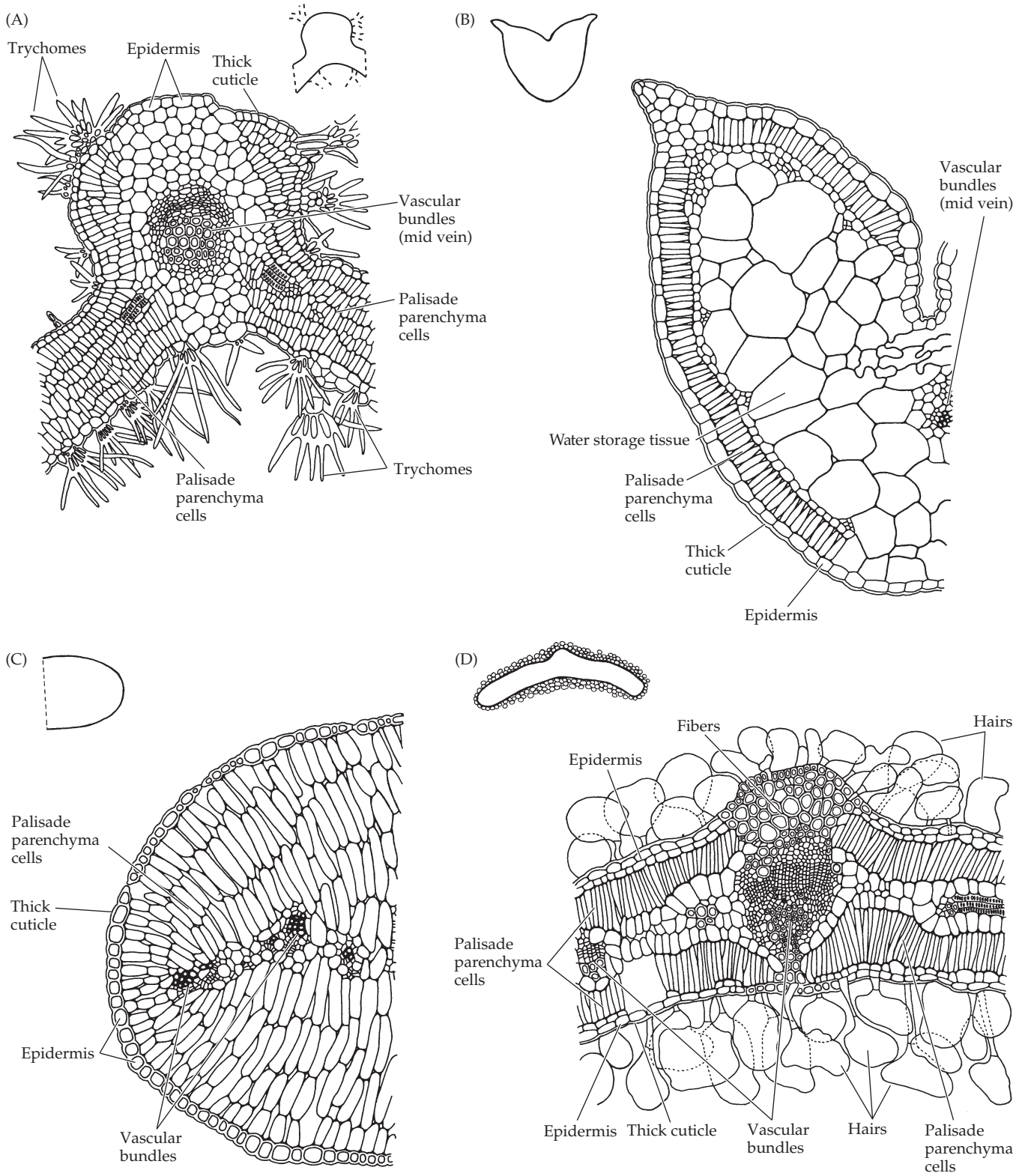
*Note*: Medians and average maxima (in parentheses) are shown for each stomatal characteristic. While there are broad overlaps between the values for plants from different environments, stomatal characteristics reflect environmental adaptations in a number of ways. Grasses are found in a wide range of habitats, from mesic to xeric. Succulents, many of which use CAM photosynthesis, have very different stomatal properties from other plants of xeric environments. Evergreen conifers tend to live in habitats that are dry due to reduced rainfall or low soil water holding capacities but have unique leaves (needles) with unusual characteristics.

to arid habitats are **amphistomatous** (have stomata on both sides of the leaf; mesophytes typically have stomata only on their undersides). Over 90% of species in North American deserts are amphistomatous. Many desert plants are also **isobilateral** (have a distinctive, symmetrical internal leaf architecture with palisade mesophyll tissue on both upper and lower sides of the leaf) (Figure 3.17). These two features together are likely to be an adaptation to making the most of the high light levels available to plants growing in arid habitats. Because intense sunlight can penetrate more deeply into the interior of a leaf than can dimmer light, maintaining these packed layers of photosynthetic cells deep inside the leaf allows the plant to achieve high maximum photosynthetic rates when opening the stomata is possible. Having stomata on both sides of the leaf reduces the average distance that $CO_2$ must diffuse before reaching chloroplasts, which also enhances photosynthesis when conditions are favorable.

Although xerophytes usually have a greater number of stomata per square millimeter of leaf surface area than do mesophytes, the average total pore area as a percentage of the total leaf area is about half as great for xerophytes as for mesophytes (see Table 3.1). This leaf structure may be an adaptation for maintaining tight control of water loss when the soil begins to dry, while still minimizing the distance $CO_2$ must travel to reach any photosynthetic cell.

Another feature of some plants from arid environments is stomata that are sunken below the leaf surface, or even located within deep crypts (Figure 3.18). This feature lengthens the distance that water must travel to diffuse away from the leaf surface, increasing the leaf's resistance to water loss. Transpiration is also directly reduced because the water vapor concentration difference between the leaf and the air immediately outside the stoma is greatly reduced by water vapor retained in the crypt.

The waxy cuticle covering the epidermis of the leaves may be especially thick in xerophytes, increasing cuticular resistance. Cuticular resistance becomes important when the stomata are closed. A plant with low cuticular resistance could leak considerable amounts of water over a prolonged dry period, even with closed stomata; preventing




such loss is critical for survival in an arid habitat. The cuticle may also have a specialized construction that reflects much of the light falling on the leaf, reducing heat from solar radiant heat input (discussed in more detail below).

The leaves of many species adapted to dry conditions are **sclerophyllous**, or tough and leathery. Various structures make the leaves tough and hard, including thick cell walls (especially in the epidermis) and increased

◀ **Figure 3.17** Cross sections of leaves with various xeromorphic features. Outlines of leaf cross sections are shown at top of each figure. (A) *Sphaeralcea incana* (desert mallow, Malvaceae) with mesophyll consisting only of elongated palisade cells. A large midvein is visible in the center of the illustration. Also note the numerous trichomes (spelled "trychomes" here, needlelike hairs; see Figure 11.12) on the surface of the leaf. (B) *Salsola kali* (tumbleweed or Russian thistle, Amaranthaceae). Anatomy of the succulent leaf, showing large-celled water storage tissue surrounded by a single layer of palisade parenchyma. (C) *Nerisyrenia camporum* (fanmustard, Brassicaceae). Leaf anatomy showing low surface-to-volume ratio and mesophyll consisting only of elongated palisade parenchyma cells. Several small vascular bundles are visible in the middle of the leaf. (D) *Atriplex canescens* (four-wing saltbush, Amaranthaceae). Leaf anatomy showing isobilateral mesophyll and large, abundant hairs on both the lower and upper epidermis. The large midvein is in the center of the leaf, with a cap of thick-walled fibers between it and the upper epidermis. Note that the leaves of all these plants have thickened cuticles. (From K. Esau. 1960. *Anatomy of Seed Plants*. John Wiley: New York; based on L. M. Shields. 1951. *Am J Bot* 38: 175–190; L. M. Shields. 1950. *Bot Rev* 16: 399–447.)

amounts of fibers and other structural tissues. Stronger, tougher leaves serve several functions, from providing mechanical strength that reduces wilting, to protection against herbivory, pathogen attack, and even damage from wind (which is likely to be stronger in open, arid habitats). Mesophytes may have plastic responses to sunny, dry conditions that are similar to some of the features of xerophytes, but not to the same degree.

## Roots, stems, and their tissues have adaptations for controlling plant water relations

There is a vast array of morphological differences among the root systems of different plants. Roots, of course, are what most vascular plants use to obtain water. Root systems differ according to the environment to which each species is adapted and the evolutionary history and relationships of that species as well. Plant roots also vary plastically in response to the environment in which the individual plant is growing.

There are several different types of root systems. **Fibrous root systems** form a dense network, capable of exploring a large volume of soil for water and nutrients. Most grasses, and many other monocots, have fibrous root systems (Figure 3.19). Other species, generally dicots, have a single **taproot** that extends deep into the soil and may also be thickened and capable of storing food.

Roots also possess various anatomical adaptations to the environments in which they grow. Roots have been the focus of far less research than leaves, perhaps

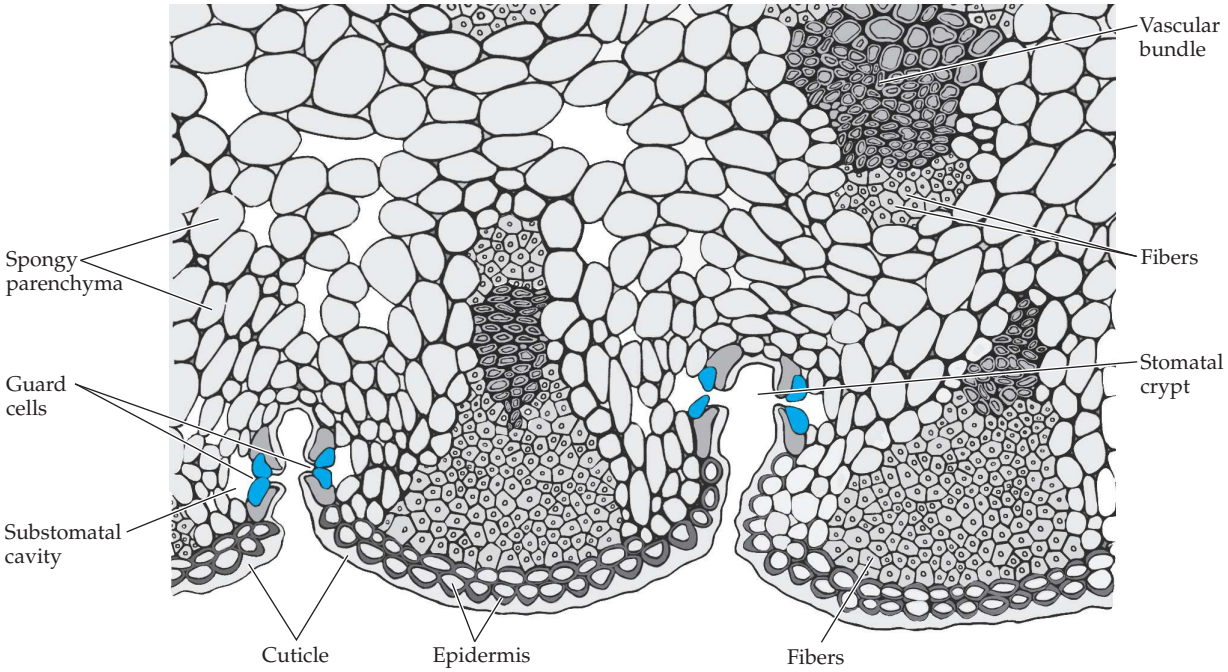



**Figure 3.18** Stomatal crypts in a leaf of *Yucca* sp. (yucca, Agavaceae). There are two stomata within each crypt; the ones in the crypt at the left are closed, and the ones in the crypt at the right are open. Other xerophytic features of this plant include thick leaves, large bundles of fibers, and a thickened, waxy cuticle. (After T. E. Weier et al. 1974. *Botany, An Introduction to Plant Biology*, 5th ed. Wiley: New York.)

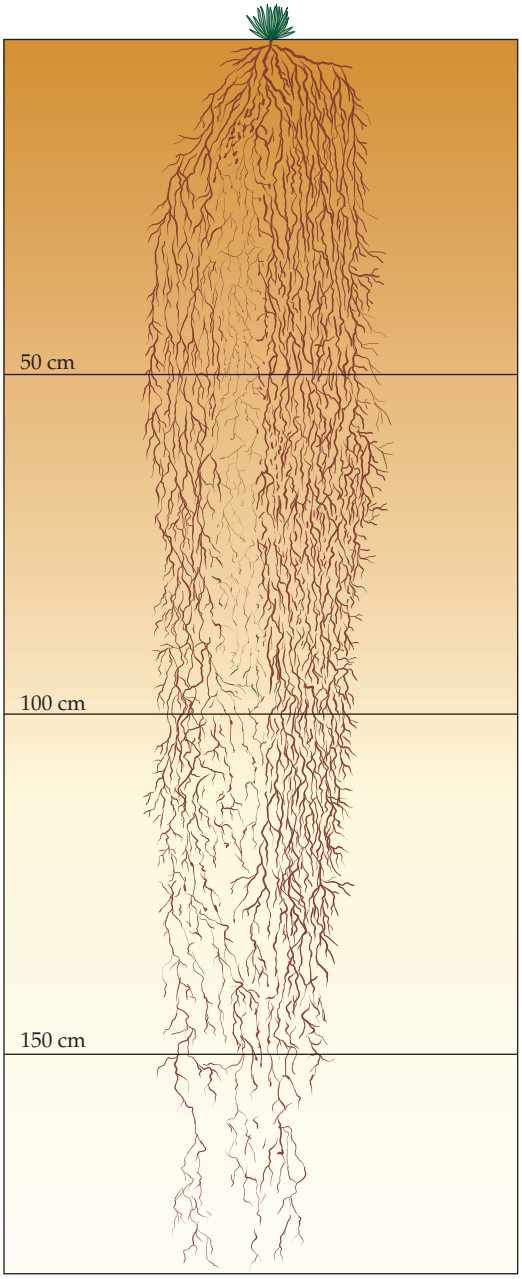
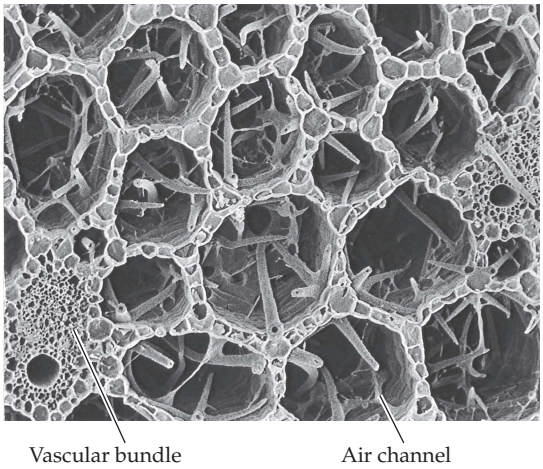



**Figure 3.19** The fibrous root system of *Triticum aestivum* (wheat, Poaceae), a cultivated monocot species. The shoot of this plant has been cut off just above the soil surface. (After H. Mohr and P. Schopfer. 1995. *Plant Physiology*. Springer: Berlin; J. E. Weaver. 1926. *Root Development of Field Crops*. McGraw-Hill: New York.)

because they are simply harder to observe in situ, although the development of rhizotrons—instrumented soil tubes or other setups to allow visualization and study of roots and their environment—has accelerated their study. One of the most dramatic adaptations of root anatomy to habitat conditions is the **aerenchyma** (aerated tissues) of aquatic plants and other plants adapted to flooding (Figure 3.20). The problem facing plants




© J. N. A. Lott/Biological Photo Service

**Figure 3.20** Scanning electron micrograph of aerenchyma tissue in *Nuphar* sp. (yellow water lily, Nymphaeaceae). The air channels allow gases, including essential oxygen, to move through plants growing in inundated conditions.

rooted in flooded soils is not so much that there is too much water, but that oxygen is very low when the soil is inundated. The formation of aerenchyma is a way of avoiding the low-oxygen conditions of flooded soil by providing an interconnected system of air channels for gas movement between the aboveground parts and the roots. Oxygen from the external air, as well as oxygen released by photosynthesis, diffuses to the roots through the aerenchyma, while other gases (such as ethylene and methane) escape from the roots and out through the aboveground parts of the plant. Other plants growing in inundated soils produce specialized aboveground roots, or roots that bend upward and emerge from the soil into the air, presumably to obtain access to oxygen (Figure 3.21).

Courtesy of J. Gurevitch

**Figure 3.21** Mangrove stand on the Gulf of Mexico in Florida. *Rhizophora mangle* (red mangrove, Rhizophoraceae) is in the foreground. Note the aerial roots, which act as props and as a means of channeling air to the submerged portions of the roots.

**Xylem**, the tissue responsible for carrying water and minerals throughout a vascular plant, is found in roots, stems, and leaves. Xylem tissue can differ greatly among species, depending on their evolutionary relationships and their ecological adaptations. In all plants, mature xylem cells are dead, and the cell contents have been removed to produce hollow vessels through which water is transported. The size and shape of xylem vessels and the connections between them differ among taxa, and these differences strongly affect their function. Xylem evolution illustrates the adaptation of plants to moisture availability and loss, to the predictability of water stress, and to the need for mechanical strength.

The water-conducting tissue in earlier-evolved vascular plants (such as ferns, horsetails, and lycopods) consists of elongated cells called **tracheids** (Figure 3.22A). These cells are connected to one another by the thinner parts of their end walls, through which water diffuses. Water transport in such a system is limited; much less water is transported, and it moves more slowly, than in more recently evolved taxa.

Gymnosperms also depend on tracheids, although their tracheids display various advances over those of the earlier vascular plants. Angiosperms have tracheids as well, but they have also evolved another type of cell, called a **xylem vessel element** (Figure 3.22B, C; Figure 3.23). These cells make up long water-conducting tubes called **xylem vessels**. In comparison with tracheids, vessel elements are shorter, and they may be fully open at the ends, allowing more water to move rapidly from root to shoot. The xylem vessel elements of eudicots from dry environments tend to be particularly short and narrow, with exceptionally thick walls. This structure offers an adaptive advantage for conducting water under the very negative pressure potentials necessary for maintaining a favorable water potential gradient when water is being extracted from very dry soils. The extreme negative pressure that the xylem is exposed to in these xerophytic plants would be enough to collapse the vessels of species not possessing this specialized kind of xylem. In some deciduous temperate trees, xylem vessels produced in the spring are wider than those produced in the summer, and the alternate pattern of wider and narrower vessel elements produces characteristic tree rings in these "ring porous" trees (although not all temperate trees show this pattern).

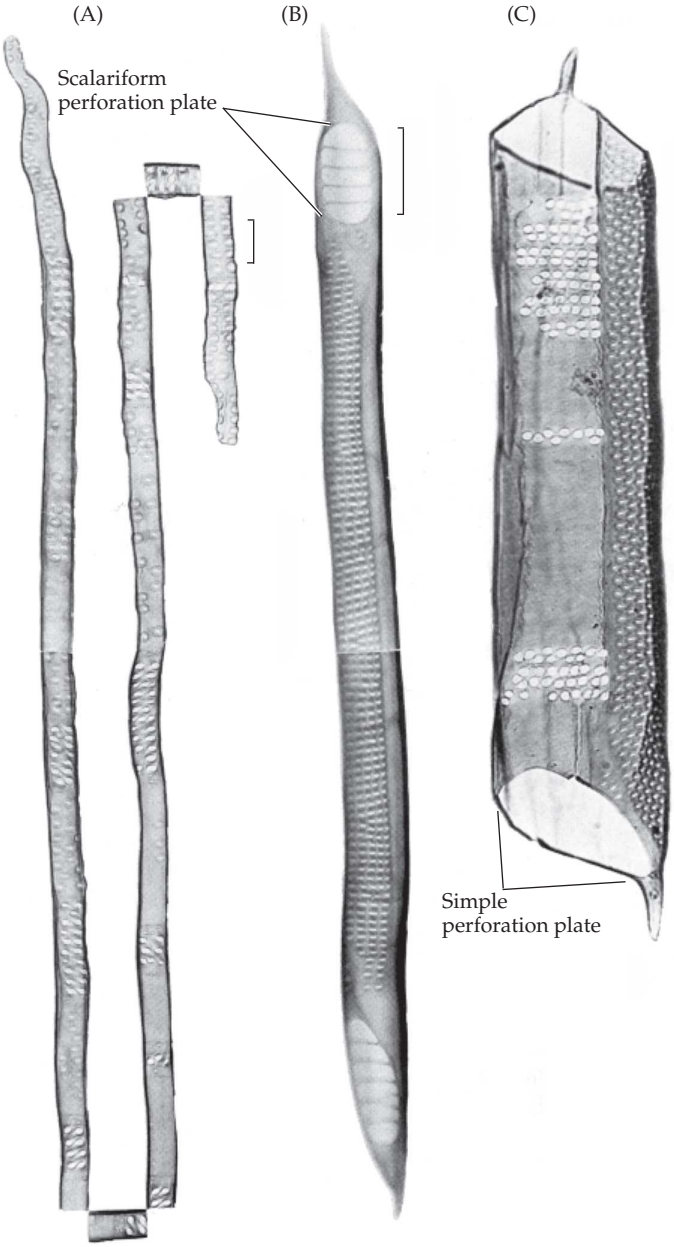



**Figure 3.22** Water-conducting xylem cells from the wood of a gymnosperm and several angiosperm trees. (A) A tracheid of *Pinus lambertiana* (sugar pine, Pinaceae), bent in several places to show its full length. (Notice that the ends of the tracheid are not open; water diffuses through these end walls as it moves up to the tracheid above it.) (B) Xylem vessel element of *Liriodendron tulipifera* (tulip tree, Magnoliaceae), showing openings in the end wall (the scalariform perforation plate; water can move directly through these openings as it travels up the xylem to the next vessel element). (C) Xylem vessel element of *Populus trichocarpa* (black cottonwood, Salicaceae), with a completely open end wall (simple perforation plate). From left to right in the figure, the xylem cells become progressively shorter and wider and have larger openings in the end walls, all contributing to facilitating the passage of greater volumes of water more rapidly. Scale bars are both 100 μm. (From K. Esau. 1960. *Anatomy of Seed Plants*. John Wiley: New York; C. H. Carpenter and L. Leney. 1952. *382 Photomicrographs of 91 Papermaking Fibers*. College of Forestry at Syracuse: Syracuse, NY.)

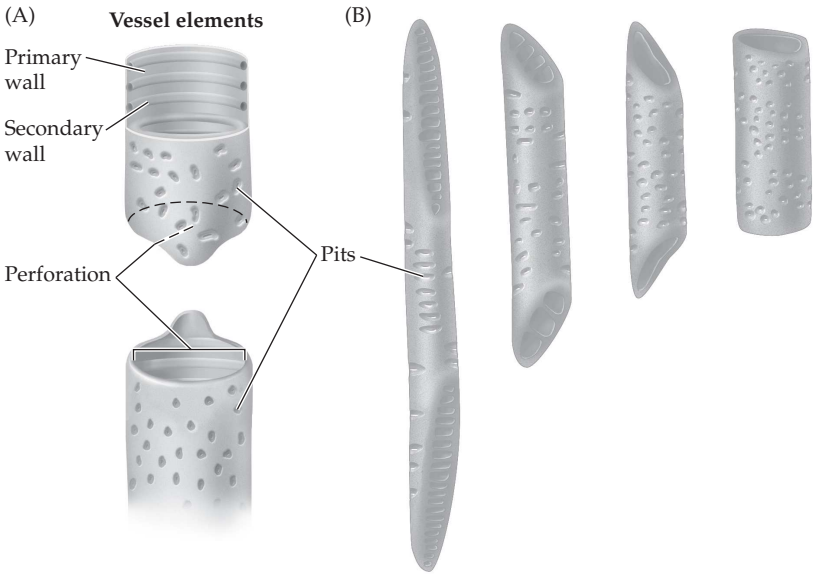
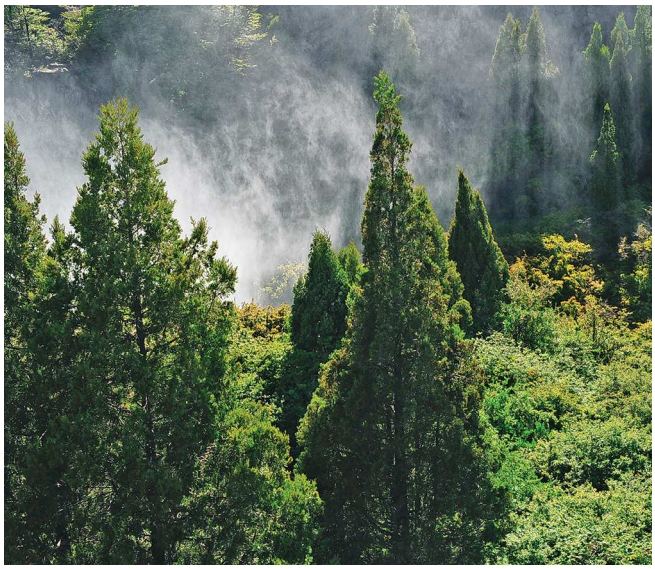



**Figure 3.23** (A) Structure of a xylem vessel element. Note that the end wall of the vessel element is perforated and open to the passage of water. The secondary wall is laid down inside the primary wall and strengthens it. The wall of the vessel element has pits—thinner parts of the cell wall through which some diffusion occurs—on its outer surfaces. (B) Xylem vessel elements from four different angiosperms arranged, left to right, from the earliest to evolve (similar to magnolia) to the most recently evolved (similar to oak). The earliest vessel elements are long and thin and closely resemble tracheids (see Figure 3.22). The evolutionary trend has been toward shorter, wider elements with thicker cell walls, which can conduct much larger volumes of water. Because of the greater volumes of water, these wider vessel elements are under greater stress; additional structural support is provided by surrounding fibers. (After J. D. Mauseth. 1991. *Botany: An Introduction to Plant Biology*. Saunders College: Philadelphia, PA.)

A severe restriction on the ability of xylem vessels to conduct water is the potential for **cavitation**—breakage of the water column in xylem vessels or tracheids. Severe water stress and freezing are the most likely conditions to cause cavitation. This can happen under extreme water stress, when bubbles of air enter the xylem through the pits in the cell walls of the vessels (see Figure 3.23) due to the very low water potential inside the xylem capillaries (the smallest vessels). The water adjacent to the bubble then rapidly evaporates into the bubble, expanding it until the column of water is disrupted, breaking the continuum of water from the soil through the plant to the atmosphere and greatly reducing the ability of the vessel to conduct water. Recordings of the popping sounds of water columns snapping when xylem cavitation occurs can be used to document the rate of cavitation. The high vulnerability to cavitation of the conifer *Austrocedrus chilensis* (Chilean cedar, Cupressaceae) is consistent, for example, with its drought avoidance strategy. This species has various morphological adaptations that generally maintain a consistently high (favorable) water potential over an exceptionally broad soil moisture gradient throughout its range in southern Chile and Argentina; cavitation can occur in its tracheids more readily than in many other plants if water potentials fall below those high values (Gyenge et al. 2005; Figure 3.24).

When the liquid sap in a xylem vessel or tracheid freezes, the gases dissolved in the sap are forced out as bubbles and trapped in the ice. During subsequent thawing, the bubbles may join together and become large enough to cause cavitation. Large, wide vessels are more likely to experience cavitation due to water stress or freezing than are smaller vessels or tracheids, but the ability of a vessel to conduct water rapidly during normal transpiration (that is, its hydraulic conductance) is also much greater in wider vessels (hydraulic conductance is proportional to the fourth power of vessel diameter). Thus, there is a trade-off between the capacity to transport large amounts of water rapidly and susceptibility to cavitation (due to either water stress or freezing) in plants with large vessels.



LBM1948/CC BY-SA 4.0

**Figure 3.24** *Austrocedrus chilensis* in Laguna del Laja National Park in the Chilean Andes.

## 3.4 Everything in the Universe Has a Thermal Energy Balance

The survival of any living thing depends on maintaining a temperature that is neither too cold nor too warm for its tissues and organs to function. Clearly, plants must avoid freezing or cooking. They must also maintain their leaves at temperatures at which they can carry out metabolic processes optimally, or at least adequately, most of the time. How can a plant manage to do this? One does not ordinarily think of plants as being capable of regulating their own temperature. Animals can shiver or bask in the sun; they can retreat to the shade or to an underground burrow when it is too hot, too sunny, or too windy—but none of these options is available to plants. What is possible for a plant? To understand this better, we need to look at the way plants, and particularly their leaves, interact with their energy environment.

All physical objects in the universe are continually exchanging energy with their surroundings, and plants are no exception. The factors that determine the temperature of any object, such as a leaf, are best understood in terms of the **energy balance** (also called the energy budget) of the object. **Temperature** is a measure of the average random kinetic energy of the molecules in a substance. Calculating the energy balance of an object means accounting for the amount of energy absorbed by the object, the amount of energy leaving it, and the amount of energy stored within it. To do this, we must first take a more specific look at the components of the energy balance of an object such as an organism. We take an intuitive rather than a mathematical approach to this topic. Lambers et al. (2008) and Nobel (2009) provide a detailed look at the biophysics of the energy balances of organisms, as well as plant water relations and photosynthesis.

In order to get an intuitive feel for the components of the energy balance of an organism—a groundhog, a leaf, or yourself—we begin by describing an ordinary experience; precise definitions will follow. Imagine that you are sitting on a rock at the beach on a warm summer day, soaking up the rays. Your skin is warmed by the sun's **radiant energy**. You can also feel heat reaching your skin by **conduction**—direct transfer—from the rock you are sitting on, which has also been warmed by the radiant energy of the sun. A breeze cools you by **convection**—heat transfer by a fluid (air or water), in this case to the air.

Now suppose you jump into the water. The water is cooler than you are, and this time the water moving against your skin carries your body heat away from you by convection to the liquid water, so you feel chilly. You do not get cold as quickly as a smaller animal would, however, because your greater mass stores more heat energy. Your **boundary layer** reduces heat transfer to the surrounding fluid, and it is much larger than that of a small animal. You start to swim vigorously, and the **metabolic energy** produced by your muscles burning calories stored in your body (originally obtained from food) warms you up. When you get out of the water, you feel chilled again: the **latent heat loss** caused by the transfer of heat energy from your skin to the evaporating water molecules is cooling your body, and convection from the breeze cools you further.

This scenario summarizes the major terms of the energy balance equation and its regulation. An understanding of the values for these terms can explain many aspects of the way organisms exchange energy with their environments and their adaptations for maintaining appropriate temperatures.

Most leaves are too small to store appreciable amounts of energy, and unlike animals, most plants do not generate much metabolic heat. Because it is usually unnecessary to account for metabolic energy and energy storage, it is relatively simple to calculate the energy budget of a typical leaf. The energy balance equation for a leaf at a constant temperature can be written as

$$R_n - H - \lambda E = 0$$

In other words, Rn, the net amount of radiant heat energy coming into the leaf, minus H, the **sensible heat** (heat that can be sensed—that is, conduction + convection) being lost, minus λE, the latent heat being lost, is equal to zero for a leaf at energy balance: The energy gained by the leaf is exactly balanced by the energy lost. If there is more energy coming in than going out, the temperature of the leaf will rise; if more energy is leaving the leaf than coming in, its temperature will decline.

While these terms are generally adequate for leaves, they are not sufficient for determining the energy balance of certain massive plant organs, such as the thick stems of some cacti, which can store considerable heat energy. There are also a few plant organs and unusual plants, such as *Amorphophallus titanum* (voodoo lily, Araceae; see Box 6A) and *Symplocarpus foetidus* (skunk cabbage, Araceae), that are capable of generating substantial metabolic energy. In such cases, additional terms for heat storage and metabolic heat must be included in the energy balance equation.

We now explore what these terms mean and what determines how large or how small their values are under different conditions. We also discuss what a plant can do to modify these values, on an immediate basis or over evolutionary time, to control its exchange of energy with the surrounding environment.

## Radiant energy is always being exchanged between plants and their surroundings

Radiant energy is that energy transferred from one object to another by **photons**, discrete packets of energy that travel at the speed of light. Physical objects are continually absorbing and emitting photons. The emission of a photon (by an electron undergoing a transition in energy state) produces electromagnetic radiation at a single wavelength. The photon flux from a source—the amount of energy emitted—depends on the wavelengths of the photons emitted; shorter wavelengths have higher energy levels. Visible light, with shorter wavelengths, has more energy than infrared radiant energy, with longer wavelengths, for instance.

The radiant energy emitted by the surface of an object, per unit surface area, is given by the **Stefan-Boltzmann equation**,

$$\Phi = \varepsilon\sigma T^4$$

where $\Phi$ is the energy emitted (the emitted flux density, in $W/m^2$), $\varepsilon$ is the emissivity of the surface, $\sigma$ is the Stefan-Boltzmann constant ($5.67 \times 10^{-8}$ $W/m^2/K^4$), and $T$ is the absolute temperature in kelvins (K; equal to the temperature in °C + 273). **Emissivity** is a measure of how efficient a body is at emitting energy. If an object absorbs all the radiation that falls on it and, in turn, emits, or radiates, all of the energy possible for an object at its temperature, it is called a **perfect blackbody radiator**. The emissivity of a perfect blackbody radiator is therefore 1.0. No real objects are perfect blackbody radiators over the entire electromagnetic spectrum, but most natural objects have emissivities of between 0.90 and 0.98 (although this can vary quite a bit according to the wavelength of the emissions).

Notice that one of the most important quantities in the Stefan-Boltzmann equation is temperature, because it is raised to the fourth power. The surface of the sun, with a temperature of approximately 6000 K, emits 73,000 $W/m^2$; the sun is close to being a perfect blackbody radiator. The energy reaching the surface of the Earth from the sun is reduced by, among other things, the scattering of the photons by particles in the atmosphere (Box 3B; see also Figure 16.1). The sun emits the greatest number of its photons in the visible range of the spectrum (see Figure 2.3B), with the peak at yellow light (that is why the sun looks yellow), although more total energy comes from its emission of higher-energy photons (shorter wavelengths) in the ultraviolet (UV) range. The Earth is much, much cooler, with an average temperature of about 290 K. It emits much less energy, radiating about 400 $W/m^2$ in the less energetic infrared part of the spectrum.

You may be surprised to learn that the Earth, like the sun, radiates energy. It does, as do all objects in the universe, including you and the leaves of plants. The wavelengths of these emissions are in the infrared part of the electromagnetic spectrum and not in the visible wavelengths, because the Earth's temperature (as well as your temperature, and that of a leaf) is too cool for it to emit visible light.

The net radiation of a leaf ($R_n$) accounts for all of its incoming and outgoing radiant energy (Figure 3.25). In sunlight, $R_n$ is positive, but at night it

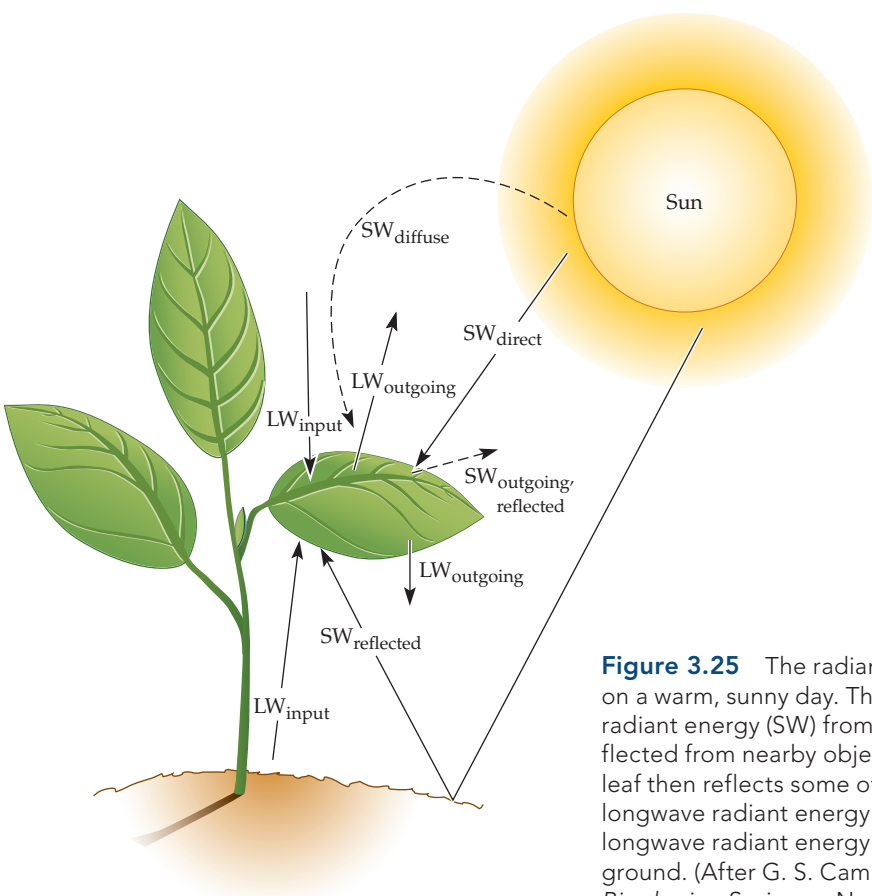



**Figure 3.25** The radiant energy flux ($R_n$) of a leaf held horizontally on a warm, sunny day. The leaf absorbs direct and diffuse shortwave radiant energy (SW) from the sun, as well as shortwave energy reflected from nearby objects, such as the ground (or other leaves). The leaf then reflects some of that shortwave radiant energy. It absorbs longwave radiant energy (LW) from the ground and the sky and emits longwave radiant energy upward to the sky and downward to the ground. (After G. S. Campbell. 1977. *An Introduction to Environmental Biophysics*. Springer: New York.)

## BOX 3B Why the Sky Is Blue and the Setting Sun Is Red

Some of the visible radiant energy from the sun is scattered before it reaches the Earth's surface. The gas molecules and small particles of the atmosphere scatter light primarily at short wavelengths (called Rayleigh scattering). This explains why the sky is blue: Blue wavelengths are the shortest visible ones common in the solar spectrum. When we see this scattered blue light from the Earth's surface, the sky appears blue to us.

At day's end, the angle of incoming sunlight is much steeper relative to the Earth's surface, and blue light is refracted (bent) high above the surface by the atmosphere. With the blue wavelengths removed, what remains are the longer wavelengths (reddish colors), which produce the reds and oranges of a sunset. In addition, large-particle scattering of light is most effective at longer wavelengths. That is why volcanic dust, particulate pollution, and other sources of suspended material in the atmosphere can result in particularly intense red sunsets.

Courtesy of S. Scheiner

will often be negative as the leaf radiates more energy than it absorbs from its surroundings. By convention, we divide incoming radiation into **shortwave radiation** (wavelengths <700 nm, including visible and ultraviolet wavelengths) and **longwave radiation** (wavelengths >700 nm, including infrared and longer wavelengths). Shortwave radiation includes direct sunlight shining on the leaf, light reflected from the surroundings, and diffuse shortwave light from various sources, including the background light from the sky. A leaf held directly perpendicular to the sunlight will receive more shortwave energy than one held at a steep angle. Longwave inputs include the radiant energy emitted by the ground, the sky, and other objects surrounding the leaf. The leaf absorbs only some of the radiant energy impinging on it; how much depends especially on its own surface characteristics. The remainder is reflected back to the sky or toward other objects in the environment.

### Energy flows between plants and air, water and soil via conduction and convection

**Sensible heat exchange**—energy exchange that results in a change in temperature—includes two major components, conductive and convective heat transfer (Figure 3.26). The transfer of sensible heat between an organism and its surroundings is directly proportional to the temperature difference between them, and inversely proportional to the organism's resistance to heat transfer. This relationship can be modeled using flux equations. Heat is transferred from warmer objects to cooler ones. The molecules in a warmer solid object vibrate faster than those in a cooler one (molecules or atoms in warmer liquids slide around over one another more rapidly, and those in gases bounce around more rapidly in three-dimensional space). When the molecules of two objects are in contact, the molecules with more kinetic energy speed up those with less. In the process, the faster molecules slow down a bit, until eventually both objects have the same temperature.

Conduction is the direct transfer of heat energy from the molecules of a warmer object to the molecules of a cooler one. It occurs when objects of different temperatures are in direct contact with each other. Conduction is a slow and fairly ineffective way of transferring energy, and it is not usually very important in either water or the atmosphere. It may be important in the soil, however, as well as inside leaves, in the boundary layer of air surrounding a leaf, and in other parts of the plant under certain conditions.

Convection is heat transport by a volume of fluid (generally air or water in the case of plants and other living organisms) moving as a unit. It is much more rapid and effective than conduction in transferring heat energy between organisms and their environment. If your body is warmer than the surrounding air, you experience convective heat loss as a cooling breeze. That is why wind chill is important in how you experience the cold: the greater the wind speed, the more rapidly

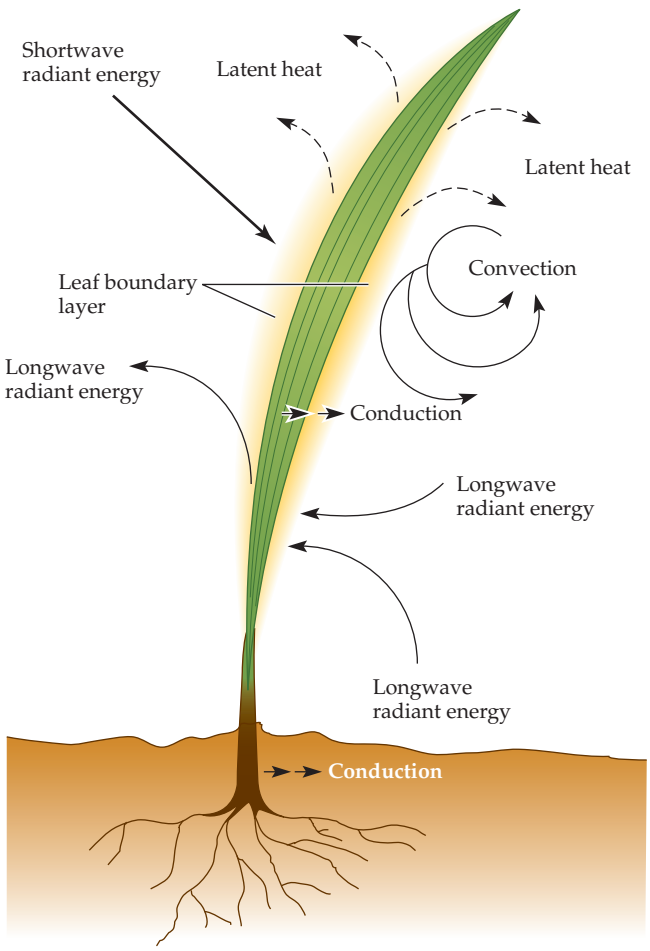



**Figure 3.26** Major components of the energy balance of a grass leaf, showing the major sources of energy entering and leaving the leaf. The leaf has a net gain of shortwave radiant energy from the sun, and it also gains longwave radiant energy from the atmosphere and its surroundings (especially the ground). It loses energy by convection, by conduction (primarily through the boundary layer and under ground to the soil), and by latent heat loss from transpired water, and it radiates energy out to its surroundings in the longwave part of the spectrum. The boundary layer of the leaf—a blanket of still air that surrounds it and reduces convective heat loss—is thickest at the central part of the leaf.

your body is cooled by convective heat loss. In unusual environments in which the surrounding air is warmer than your body, the breeze will feel like a blast of heat from a furnace. In this case, you are experiencing convective heat gain.

The fur of mammals and the feathers of birds can offer considerable resistance to convective heat exchange, but leaf **pubescence** (leaf hairs) does not generally have a big effect on reducing convection. However, leaves have another option. Like all objects in air or water, leaves are surrounded by a blanket of relatively still fluid, called the **boundary layer**. Heat (and mass) transfer across a boundary layer is much less efficient than heat transfer elsewhere in the fluid. The thicker the boundary layer, the greater the resistance to heat transfer posed by a leaf.

The boundary layer of a leaf is determined mainly by wind speed and by the size and shape of the leaf. The greater the wind speed, the more eroded the boundary layer becomes. Larger, rounder, more entire leaves carry a thicker boundary layer than leaves that are smaller or narrower or have more dissected (broken-up) shapes. The **characteristic dimension** of a leaf is a measure that takes into account both its size and shape and represents the effective width of the leaf with respect to fluxes of energy and mass entering and leaving the leaf. It is directly proportional to the boundary layer resistance of the leaf.

## Water loss is accompanied by latent heat loss

**Latent heat exchange** is the transfer of energy that occurs during evaporation—the process of converting water from a liquid to a gaseous state. (It is called latent—concealed—because there is no change in temperature during this energy transfer process, in contrast with sensible heat exchange, which can be "sensed" due to a change in temperature.) It is equal to the product of $E$, the evaporation rate, and $\lambda$, the latent heat of vaporization of water (2.45 MJ/kg at 20°C). Like convection, latent heat exchange can be a highly efficient cooling mechanism for organisms. The molecules in a substance in the gaseous state have much more energy than those in the liquid state at the same temperature. At 20°C, water vapor has about 2450 J/g more energy than liquid water. In contrast, warming liquid water by 1°C takes only about 4 J/g. Where does all of the extra energy come from as water is transformed from liquid to gas in the process of evaporating? It is taken from—and thus cools—the surface that is losing the water: the tongue of a panting animal, the skin of a sweating athlete, or the tissue of a transpiring leaf.

## Putting it all together: what determines leaf and whole-plant temperature?

The ability of a leaf to survive and function depends on its ability to maintain its temperature within an acceptable range. The leaf temperature, in turn, depends on the energy balance of the leaf. The large amount of radiant energy input to a leaf in direct sunlight, for example, must be balanced by adequate convective and/or latent heat loss. Leaves that have a large characteristic dimension and a high boundary layer resistance can potentially become decoupled from the air temperature around them, and they can be many degrees warmer or cooler than leaves with a small characteristic dimension (Figure 3.27).

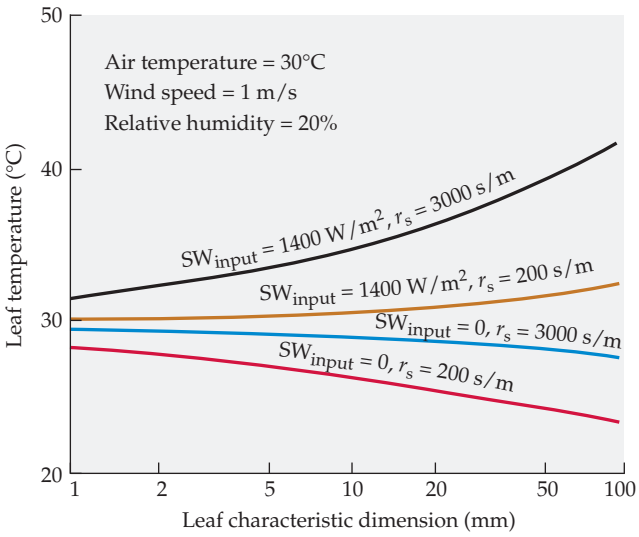



**Figure 3.27** Leaf temperatures at a constant air temperature of 30°C, with a constant wind speed of 1 m/s (a fairly stiff breeze), and relative humidity of 20%. Leaf temperatures were calculated for leaves of increasing characteristic dimension (approximately equal to leaf width) at different shortwave radiant energy input levels ($SW_{input}$, expressed in W/m$^2$ of leaf surface) and stomatal diffusive resistances ($r_s$, expressed in s/m). Small leaves (those with low values for leaf characteristic dimension) remain close to air temperature, regardless of sunlight intensity and stomatal resistance (which affects water loss, and thus latent heat loss), while large leaves can become much hotter or much cooler than the air, depending on shortwave input and stomatal resistance. (After G. S. Campbell. 1977. *An Introduction to Environmental Biophysics*. Springer: New York.)

Courtesy of S. Scheiner

**Figure 3.28** *Gunnera insignis* (poor man's umbrella, Gunneraceae), an herbaceous species native to Central America with impressively large leaves. This photo was taken in Costa Rica, in a cloud forest.

A plant with large leaves in bright sunlight (large radiant energy input), still air (minimal convective heat loss), and dry soil is faced with a critical dilemma. Opening the stomata to allow cooling through latent heat exchange risks severe wilting and even death. But stomatal closure to restrict water loss will also reduce latent heat loss, potentially causing the leaf to heat to lethal temperatures. With adequate water available and even a light breeze, however, latent and convective heat losses can cool a large leaf to well below air temperature (Figure 3.28). Small leaves are likely to remain close to air temperature, even with closed stomata, as the slightest breeze will minimize their boundary layer resistance. Under high winds, leaves of any size will have substantial rates of convective heat exchange. Conversely, at equal stomatal conductances, a warm leaf will lose a lot more water than a cool leaf because water evaporates more quickly from a warm surface (the warm surface supplies more energy to the evaporating water).

Thus, on a short time scale, plants (at least those with large leaves) can directly control latent heat loss from their leaves, and thus leaf temperature, by controlling their transpiration rate. Over many generations, adaptation to warm, dry environments can involve the evolution of leaves of different sizes and shapes, changing their boundary layer properties and altering convective heat exchange. Many desert plants, for example, have small or narrow leaves that have a low boundary layer resistance to heat transfer and remain close to air temperature even when the stomata are closed. Large leaves of some species may be broken up into small segments with small boundary layers, keeping their temperatures always close to air temperature.

Radiant heat input is also, to some extent, under the control of the plant. The leaves of many desert shrubs are held at a steep angle (Figure 3.29), which reduces their radiant heat gain. There are also plants that change leaf angle in response to leaf temperature. Leaf coverings such as reflective white pubescence and shiny, waxy coatings serve a similar function, reducing radiant heat input to the interior of the leaf (at the cost of reducing the light available for photosynthesis).

Some cacti have dense coverings of highly reflective spines; while these structures may also act to deter herbivores, a main benefit is reducing radiant heat input. *Encelia farinosa* (brittlebush, Asteraceae) is a desert shrub that produces two different kinds of leaves, depending on the season. In summer, the plant puts out densely pubescent, highly reflective leaves covered with a mat of white hairs. In spring, when temperatures are cooler and water is more available, it produces green, nonpubescent leaves. James Ehleringer and colleagues (1976) determined that the spring leaves of *E. farinosa* absorb much more shortwave radiation, and consequently have

brewbooks/CC BY SA 2.0

**Figure 3.29** Leaves of *Simmondsia chinensis* (jojoba, Simmondsiaceae), a desert shrub, The thick, leathery leaves are held at an upright angle to reduce radiant heat gain.

Courtesy of S. Schwinning

**Figure 3.30** Variation in the appearance of leaves of *Encelia farinosa* (brittlebush, Asteraceae). The leaves in the center of the plant are being produced as the environment is becoming warmer and drier; those surrounding it were produced earlier in the season, when conditions were cooler and moister. The central leaves are thicker, smaller, and covered with a dense white pubescence (hairs) that reflects sunlight; those toward the outside are greener because they lack the thick pubescence.

higher photosynthetic rates, than the reflective summer leaves (Figure 3.30). In contrast, the highly pubescent summer leaves reflect much of the incoming solar radiation; although this reduces their photosynthetic rates, it also reduces leaf temperatures. Consequently, these leaves are able to remain alive and photosynthetic even in very hot, sunny conditions, enhancing the total carbon gain of the plant and indirectly reducing water loss (Ehleringer and Mooney 1978). A number of other plant species also produce different types of leaves in different seasons, each with morphological or physiological adaptations for functioning optimally in its own season.

A common adaptation of many grasses (and some other plants) is to roll their leaves into narrow cylinders when facing water stress. This change serves to modify the leaf energy balance in several ways, minimizing both boundary layer resistance and radiant heat input. In addition, rolling up the leaf directly reduces water loss by reducing the amount of leaf surface exposed to the dry surrounding air. There are also many whole-plant adaptations for managing water and energy losses, including partial or total shedding of leaves in dry or cold seasons, as we saw earlier.

## Plants may have adaptations to extreme temperature regimes

Extremes of heat and cold can damage plants in a number of ways. At cold temperatures, membranes lose their elasticity, biochemical processes slow and the energy required to activate them becomes greater, and the formation of ice crystals can destroy cells. At high temperatures, many biochemical processes are also reduced, membranes can become damaged, and proteins are denatured. The range of temperatures at which a plant can function effectively, its **activity limit**, is narrower than the range of temperatures at which it can remain alive, the **lethal limit**. Plants are capable, to some extent, of adjusting their biochemistry, metabolism, and morphology to increasing or decreasing temperatures. Plant species may evolve greater tolerance of higher or lower temperatures over the course of many generations, so species characteristic of different parts of the world have very different tolerances for extremes of heat and cold (Table 3.2); different tolerances also exist among populations of some widespread species. These adaptations can be morphological, biochemical, phenological, and physiological (for examples see Larcher 1995).

Contrasts in the environment can affect plant distributions globally, regionally, and locally. Maximum and minimum temperatures vary greatly over the course of a day and a year in different places across the globe (see Chapter 16). The length of time that plants are exposed to extreme temperatures also varies with the nature of the surrounding plant community (e.g., forest vs. grassland) and the microhabitat. For example, local plant distributions in temperate regions can be greatly affected by **cold air drainage**—sinking of cold air masses to low-lying spots due to the greater density of colder air. Freezing can thus occur late in the spring in local "frost pockets"

**TABLE 3.2** Temperature extremes (°C) for leaf damage for plants from different environments

| Environment | Threshold temperature for cold damage | Threshold temperature for heat injury |
|---|---|---|
| Tropical habitats | | |
| Trees | +5 to –2 | 45–55 |
| Alpine plants | –5 to –15 | about 45 |
| Temperate habitats | | |
| Deciduous trees and shrubs | –25 to –35 | about 50 |
| Herbaceous species | –10 to –20 | 40–50 |
| Grassland species (grasses and sedges) | less than –30 | 60–65 |
| Habitats with cold winters | | |
| Evergreen conifers | –40 to –90 | 44–50 |
| Boreal deciduous trees | less than –30 | 42–45 |
| Herbaceous alpine and Arctic species | less than –30 | 18 (20) |

*Source*: After W. Larcher. 1995. *Physiological Plant Ecology*, 3rd ed. Springer: Berlin.
*Note*: For determination of cold tolerance, plants were exposed to cold and allowed to acclimate ("harden") before measurement; heat tolerance was determined during periods of active growth.

long after the surrounding areas are frost-free, causing sometimes dramatic differences in the vegetation.

The question of the role of extreme temperatures in determining plant species distributions has attracted the attention of plant ecologists for a century or more. Forrest Shreve (1911) hypothesized that the northern distribution and upper elevation of *Carnegia gigantea* (saguaro, Cactaceae; see Figure 3.10A) was limited to places where freezing temperatures did not persist for longer than 24 h.

Park Nobel and his associates have conducted many studies on the effects of extreme temperatures on the distributions of desert species in North America. In one study, Nobel (1980) found that the apical meristems of six species of giant columnar cacti were the most vulnerable parts of those plants. Those parts experienced the lowest temperatures because their position at top of the plant, exposed to the night sky, made them the most likely to freeze. Because these parts are the growing points of the main stems, freezing damage would prevent further growth of the cactus and limit the distribution of the species. The diameter of these columnar cactus species increases northward from Mexico into the United States, and Nobel showed that the greater the diameter, the higher the minimum temperature of the vulnerable stem top. Death resulting from a single freezing event can take from 3 to 9 years to occur in these massive plants (Steenbergh and Lowe 1976a, b). Guillermo Goldstein and Park Nobel (1994) showed that *Opuntia humifusa* (devil's tongue, Cactaceae), a cactus species native to more northern environments, exhibits adaptations at the cellular level that increase its tolerance of freezing. For example, it increases its osmotic potential by synthesizing simple sugars and mannitol and thus prevents intracellular ice formation and tissue dehydration due to freezing. Cactus species native to warmer environments do not show such adaptations.

Courtesy of Park S. Nobel

Park Nobel

At the largest temporal and spatial scales, the sizes of populations and the distributions of species change over time, in part as a consequence of the interaction between the organisms and their responses to factors in their abiotic environment, such as temperature extremes and changing temperature regimes (see Chapter 16). Populations increase, decrease, or go extinct, and plant distributions change as the climate changes. Following glaciations, for instance, plant species gradually recolonized formerly glaciated lands (see Chapter 17). At the longest temporal scales, evolutionary changes may alter both the activity limit and the lethal limit for heat and cold, sometimes within species, sometimes by the evolution of new species with different tolerances.

## Summary

- Terrestrial plants are descended from organisms that lived in the oceans, and their ability to live on land has required the evolution of anatomical, physiological, and other adaptations to life in dry terrestrial environments.
- Plants' abilities to obtain water, and the movement of water throughout the plant, depend on the water potential of plant tissues and of the surrounding soil and atmosphere.
- Water moves from less negative to more negative water potentials, so we can predict the direction of water movement if the water potential of various components is known.
- Plants lose much more water in transpiration than they use metabolically.
- Plant strategies for obtaining and retaining water and tolerating its loss differ greatly. Mesophytes, aquatic plants, hygrophytes, xerophytes, halophytes, phreatophytes, and desert ephemerals differ in how they obtain and retain water, and survive water loss.
- Plants are continually exchanging energy with their surroundings. The temperature of leaves and other plant organs depends on their energy balance. Radiant energy input and loss, sensible heat exchange (conduction and convection), and latent heat loss are the major components of leaf energy balance and together determine leaf temperature.
- The long-term adaptations of plant species and the shorter-term physiological processes of individuals modify many of the critical environmental factors that plants experience, including temperature, light, and moisture.

© Sonerbakir/Dreamstime.com

4

# Soil and Terrestrial Plant Life

Most terrestrial plants are rooted in the soil and depend on it for support, water, and mineral nutrients. What is soil made of, and how is it created? What role do living organisms play in making soils what they are, and what role do soils have in determining the nature of plants and plant communities? What do plants get from the soil, and how does soil structure and composition affect plants? How do soils in different parts of the world differ? In this chapter, we look down and examine some of the surprisingly complex things going on right under our feet. We look at the characteristics and properties of soils and see how those properties affect the availability of water to plants. We will also consider some of the interactions between plant roots and the organisms in soils.

Above: Soil erosion and loss is a major environmental threat in forests and other terrestrial systems in many parts of the world.

## 4.1 Soils Have Distinct and Varied Composition, Characteristics, and Structure

It is easy to overlook soils. If you have not had the opportunity to examine soils in detail, you might think of soil as undifferentiated "dirt." But soils are acually complex, often highly structured systems that are the unique product of the interaction between living organisms and a physical matrix. Because soil is formed by interactions among living organisms and their physical substrate, it is found only on the Earth and nowhere else in our solar system (at least as far as we know). In fact, another English word for soil is *earth*. (Some people call the loose substrate that has been collected from the moon and Mars "soil," but this material does not have either the development or properties of soils on Earth.) The plants growing in soil affect its structure and composition. At the same time, soil structure and composition are important factors determining plant growth and the plant species that occur in a given location. We begin to explore this reciprocal interaction in this chapter, and we hope that this will give you more of an awareness and appreciation for the diversity, complexity, and importance of soils.

In the 1870s the Russian scientist Vasily Dokuchaev first formulated the modern concept of soil as a product of climate, living organisms, and the physical substrate, rather than just being disintegrated rock. In the early and mid-twentieth century, Hans Jenny greatly advanced Dokuchaev's ideas, unifying concepts of physical chemistry with those of the microscopic and macroscopic physical structure of soils, conceiving of soils as a complex, dynamic open system. Jenny, born in Switzerland, moved to the U.S. as a young man and collaborated early in his career with the Swiss ecologist Braun-Blanquet (see Chapter 12). He formalized the understanding of soil formation, and he integrated soil science with what we now call ecosystem ecology, particularly advancing the science of the role of soils in the carbon and nitrogen cycles (see Chapter 5) and developing our knowledge and understanding of global patterns of soil geography. Jenny once said, "Many ecologists glibly designate soil as the abiotic environment of plants, a phrase that gives me the creeps" (Jenny as quoted by Stuart [1984]). He was instrumental in soil and plant community conservation and fostered an aesthetic appreciation of soils. He said:

> *"Soil appeals to my senses…. I like to sit on bare, sun-drenched ground and take in the fragrance of soil…. Warm brownish colors characterize fields and roofs in Cezanne's landscape paintings of southern France, and radiant red soils of the tropics dominate canvasses of Gauguin and Portinari [the great Brazilian painter]…. I have seen so many delicate shapes, forms, and colors in soil profiles that, to me, soils are beautiful."*
>
> —Jenny as quoted by Amundson (2005)

What materials are found in soil, and what is it made of? To begin with, there are mineral particles, derived from rock. The largest components—stones and boulders—are not really part of the soil but may be part of the parent material from which it forms. In progressively smaller sizes, we can see pebbles and gravel. The mineral soil itself consists of sand particles (>2.0–0.02 mm in diameter), silt particles (0.02–0.002 mm), and clay particles (<0.002 mm). Besides these mineral particles, there is organic material in varying states of decomposition. Air and water containing dissolved minerals are found in the pores between the mineral particles and around the organic matter.

Soil structure is a property of soils that has a major influence on plants (as well as on the other organisms living in soils). **Soil structure** describes the physical arrangement of soil particles into larger clusters (called aggregates or peds). One of the most important aspects of soil structure is the **porosity** of the soil, which measures the total volume of pores and their size, shape, and arrangement between and within the aggregates (the aggregates are not solid and may be crumbly, "puffy," or packed to varying degrees). Soil porosity determines how much, and how easily, water and air can be held by and move around in the soil and thus is an important factor in the water-holding capacity of the soil.

Finally, soil is, to one extent or another, the environment of various kinds of living things: fungi, bacteria, photosynthetic prokaryotes such as cyanobacteria, single-celled eukaryotes such as diatoms and protists, visible and microscopic insects and other arthropods (see Chapter 5), and many other kinds of animals small (such as slugs, earthworms, and nematodes) and large (such as groundhogs in the eastern U.S. and prairie dogs in the western U.S.). And, of course, soil is filled with the roots, rhizomes, and other underground parts of plants. The living things in the soil interact with one another and with the soil to continually alter its physical and chemical properties, including its pH and nutrient composition. Plants take up large volumes of water from the soil, changing the amount and distribution of air and water within it (see Chapter 3), and release and remove materials to and from the soil surrounding their roots. The various soil organisms act and interact in a wide variety of ways. Some benefit one another; others consume one another or defend themselves to avoid being consumed, or decompose others' remains. In the process of metabolism, soil organisms respire, altering the chemistry of their environment, and produce waste products and

other substances, which become part of the soil. Their activities affect plants in many ways, from making nutrients available to causing plant diseases.

Soils are important determinants of the kind of plant community found, and the plants in a community influence the soil they are growing on. Sometimes the soil is so distinctive that it has a dominant effect on the vegetation growing in an area. For example, distinctive and unusual plant communities may be found on an unusual kind of soil called serpentine soil, derived from serpentine rock (Box 4A). Serpentine soils have characteristics which may not be tolerated by the surrounding plant communities, including very low or very high pH values and high concentrations of toxic heavy metals.

### It takes many thousands of years to create soil

Soils have different origins, and this is a major cause of the great differences among soils. The origins and formation of soils are important factors contributing to differences in vegetation among regions. What processes create soil? Soils are formed from **parent material**, the upper layers of the heterogeneous mass that is left over after the action of weathering and other forces on rocks. The rocks from which the parent material is originally derived are classified as **igneous** (of volcanic origin), **sedimentary** (from the deposition and recementation of material derived from other rocks), or **metamorphic** (changed by the action of great pressures and temperatures on igneous or sedimentary rocks deep underground).

Large and small fragments of rock and rock-derived debris break up in place or are moved by ice (particularly by glaciers), wind, and water. They may then be redeposited in the form of glacial till and outwash, **loess** (pronounced "luss") and other eolian (wind) deposits, alluvial (river and stream) deposits, and lake and marine sediments (which may develop into soil long after the area becomes dry land). Chemical reactions occurring in the disintegrating rock material accelerate the physical breakdown of the fragments and begin the process of decomposition and chemical alteration (Figure 4.1). The parent material can be further broken up by the mechanical action of temperature changes, including repeated freezing and thawing; by the direct action of water, wind, and ice in abrading and eroding rock fragments; and by the actions of plants, fungi, and animals. These factors grind up, move, and leach (dissolve minerals from) the parent material until it begins to develop into true soil.

## BOX 4A Serpentine Soils

Serpentine soils occur all over the world, often in patches of limited extent, or "islands," surrounded by more ordinary nonserpentine soils. They have a number of very unusual characteristics. These "serpentine barrens" often support distinctive plant communities that contrast strongly with the vegetation of the surrounding region. For instance, serpentine vegetation in California typically occurs as species-rich patches with many showy dicot species, surrounded by lower-diversity, nonserpentine grasslands. Serpentine vegetation often includes many endemic species. Serpentine soils are derived from serpentine rock, which is a magnesium silicate rock of metamorphic origin. These soils are poor in many nutrients—particularly calcium, nitrogen, and phosphorus—and they may have either very low or very high pH. Toxic elements, particularly heavy metals, are also characteristic of serpentine soils.

The species from surrounding areas are unable to tolerate the toxic, low-nutrient soils, and this excludes them from serpentine patches. In addition, Kruckeberg (1954) showed that the serpentine species appear to be excluded from the surrounding regions by competition (echoing Bradshaw's findings for plants growing on mine tailings; see Chapter 6). Important ecological research has been conducted over the years on various aspects of the ecology of vegetation on serpentine soils (Kruckeberg 1954; Whittaker 1954; Hooper and Vitousek 1998).

Courtesy of David Hooper

Serpentine grassland in the state of California.

The process of soil formation from parent material takes thousands of years. A young soil may be only 10,000 years old; an old soil may be 100,000 years old or even older. As a soil ages, the primary minerals that came from the parent material undergo chemical changes, and secondary minerals are formed. The soil develops structures and properties that differ physically and chemically with depth. Materials are leached, redeposited, and lost from the soil as well as being physically and chemically altered. The mineral constituents of the soil shift as minerals that are most stable at the Earth's surface gradually come to dominate. Young soils are found where the parent material is still present, as in areas that were glaciated relatively recently (such as much of the northern part of North America) or in areas of recent alluvial deposits (such as the bottomlands of the world's great river systems). The oldest soils are most commonly found in the tropics and subtropics. Africa and Australia have some of the world's oldest soils. Because soils take so long to create, they are not easily restored if they are damaged or destroyed.

Five major factors are responsible for determining the kinds of soils that develop in an area: climate, the nature of the parent material, the age of the soil, the topography, and living organisms. These factors do not operate in isolation, but rather interact in complex ways.

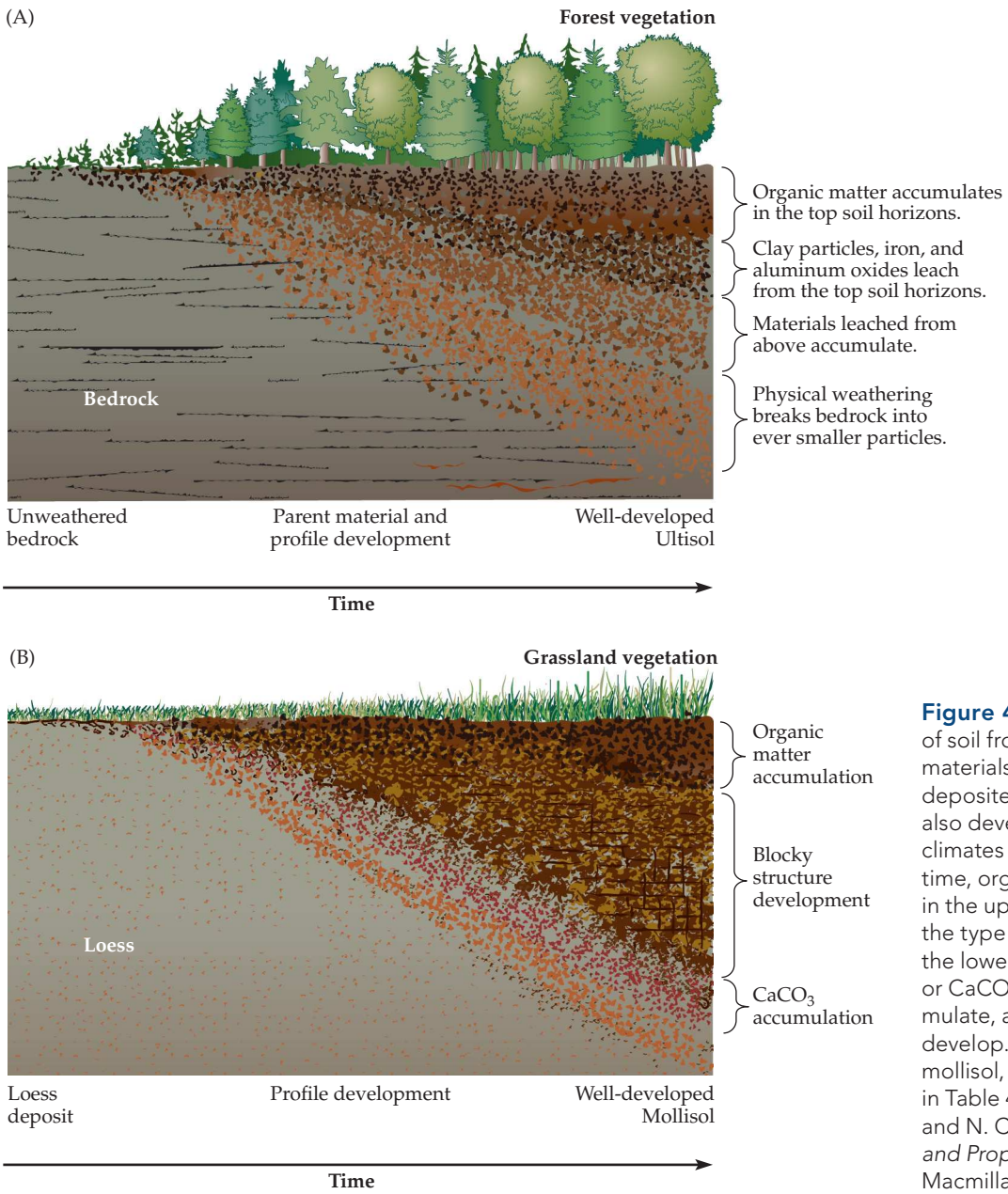



**Figure 4.1** The development of soil from two different parent materials: (A) bedrock and (B) loess deposited by wind. These soils are also developing under two different climates and vegetation types. Over time, organic matter accumulates in the upper horizon, depending on the type of vegetation present. In the lower horizons, clay, iron oxides, or $CaCO_3$ are deposited and accumulate, and characteristic structures develop. The soil orders (ultisol and mollisol, respectively) are described in Table 4.1. (After H. O. Buckman and N. C. Brady. 1969. *The Nature and Properties of Soils*, 7th ed. Macmillan: New York, NY.)

Vegetation, with its associated soil organisms, has an especially strong influence on soil development, but the nature of the vegetation that is present is also dependent on both soil and climate. Land use by humans also plays an important role in modifying the physical and chemical properties of soils.

Soils are classified according to a comprehensive taxonomic system developed in the United States, although some soil scientists and ecologists still rely on earlier classification systems. The broadest category in the modern system is the **soil order**. The major soil orders worldwide (Table 4.1) are characterized by rainfall,

**TABLE 4.1** Soil orders and some of their characteristics

| Soil order | Development and characteristics | Found in |
|---|---|---|
| Mollisols | Develop under prairie vegetation; many under cultivation and constitute some of the most productive agricultural soils | Great Plains of the United States and Canada; large areas in the heartland of Russia, Mongolia, and northern China; northern Argentina, Paraguay, and Uruguay |
| Spodosols | Generally develop in cold-temperate humid regions under northern forests; usually strongly acidic and of low fertility | Northeastern and northern midwestern United States and adjacent areas in Canada; northern Europe, Siberia; some areas in southern South America |
| Alfisols | Develop in humid regions under deciduous forests; weathered somewhat less than spodosols but more than inceptisols; contain silicate clay-dominated horizon; often cultivated and highly productive agriculturally | Large area from Baltic states to western Russia; southern half of Africa; eastern Brazil; England, France, central and parts of southern Europe; Michigan and Wisconsin to Pennsylvania and New York in the United States |
| Oxisols | Usually tropical soils, often supporting tropical rainforest vegetation; oldest and most highly weathered soil order, with a deep subsurface horizon high in clays and other colloidal particles; low in silica, removed by intense leaching, leaving a high proportion of oxides of iron and aluminum, often with a red color; one of the most widespread soil orders globally and among the most important in terms of size of human population supported; also supports the greatest biodiversity of many groups of plants and animals | Much of Africa and South America |
| Entisols | Recent soils with lack of significant profile development; highly variable in fertility | Sand hills of Nebraska; areas of northern and southern Africa; northern Quebec; parts of Siberia and Tibet; many areas that were covered by the most recent glaciation |
| Inceptisols | Young soils exhibiting limited weathering but with greater profile development than entisols; many under agricultural production | All continents—northern Africa, eastern China, western Siberia, Spain, central France, central Germany, northern South America, northwestern United States |
| Aridisols | Develop in arid regions where soils are dry through most of the year and there is little leaching; high pH and often with calcic horizons | Southwestern United States and northern Mexico; southern and central Australia; Gobi and Taklamakan deserts in China; the Sahara and also southwestern Africa; Pakistan; other arid regions |
| Vertisols | High content of certain clays; sticky and plastic when wet and hard when dry, with extensive cracking; unstable for building and difficult to farm | Large regions of India, Sudan, and eastern Australia; small areas in southeastern Texas and eastern Mississippi in the United States |
| Ultisols | Develop under warm to tropical climates under forest or savanna vegetation; generally highly weathered but still retaining some minerals; somewhat acidic, with clay horizons; fairly old soils | Southeastern United States; northeastern Australia; Hawaii; southeastern Asia; southern Brazil |
| Histisols | Wetland and bog soils; develop in a water-saturated environment; very high in organic matter content; sometimes also high in clay; important not only in present-day wetlands, but in areas of ancient swamps and bogs that are currently forested, cultivated, or used for mining peat | Not widespread but may be locally important, as in the Everglades in southern Florida and in many areas of northern Europe; largest areas in Canada just north of Hudson and James Bay and in northwestern Canada into eastern Alaska |

temperatures, soil age, and to some extent underlying geology (Figure 4.2). For example, Gelisols occur in very cold climates and possess permafrost, permanently frozen lower soil layers. Almost 25% of the Earth's terrestrial area is characterized by soils with permafrost. As a result of climate change (see Chapter 16), permafrost is melting rapidly in many northern areas. The effects on human and natural systems of melting permafrost are dramatic:

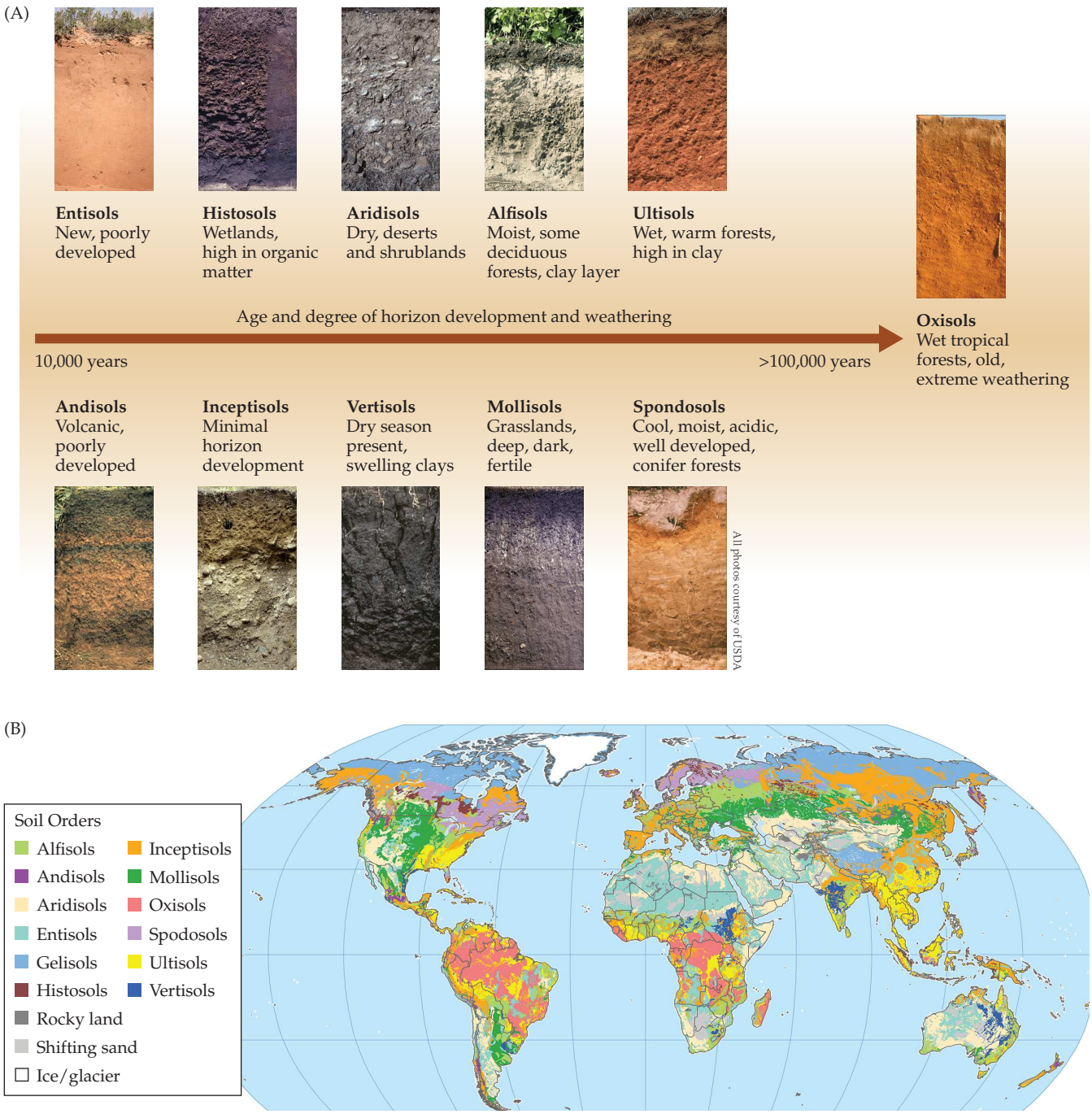



**Figure 4.2** The global soil orders. (A) Weathering and development of the global soil orders. (See https://passel.unl.edu/pages/informationmodule.php?idinformationmodule=1130447032&topicorder=3&maxto=16&minto=1 for typical soil profiles for each soil order.) (B) Geographic distribution of the global soil orders. (A after N. C. Brady and R. R. Weil. 2008. *The Nature and Properties of Soils*, 14th ed. Pearson: New York, NY; B from USDA-NRCS Soil Survey Division, World Soil Resources: Washington D. C.; https://www.nrcs.usda.gov/wps/portal/nrcs/detail/soils/use/?cid=nrcs142p2_054013.)

buildings, roads, and trees can become unstable and collapse and fall as the frozen layers supporting them melt. Soils can turn into waterlogged mud when upper layers of permafrost melt but the water has nowhere to go because deeper layers are still frozen. This can result in erosion and mudslides on steeper terrain. Decomposition of the organic material that was previously trapped in the permafrost releases large amounts of methane and $CO_2$, exacerbating climate change.

The soil classification categories most commonly used by plant ecologists are those at the regional and local levels: **soil series** (Figure 4.3) and **soil type**. A soil series is usually named after a local geographic area, feature, or person and generally consists of a dominant type with several other associated types. Soil types are based on topography, parent material, and the vegetation under which the soils were formed. Soil series occur at regional scales, and soil types are landscape-scale features. In the United States, soil series are mapped for almost every county (websoilsurvey.sc.egov.usda.gov/); detailed maps are available for almost all locations, showing the soil series and type and often providing extensive information on the characteristics of local soils. Such data can be highly valuable for many kinds of ecological studies. The availability of information on local soils in other countries varies widely.

## Soil texture determines many of the properties of soils that affect plants

The properties of soils, and the way soils affect plants, depend, first, on **soil texture**: the proportions of the different-sized particles making up the soil. Soil texture varies greatly among ecosystems, regions, and biomes. The texture of the soil is a major factor in determining plant growth and survival, especially because it has large consequences for plant water and nutrient relations.

Depending on which particle size—sand, silt, or clay—dominates the character of a soil, the soil texture is categorized as sandy, silty, clayey, or loamy (Figure 4.4). **Loamy soils** have a balance between sand, silt, and clay particles and are generally considered the most desirable soils for agriculture. **Sandy soils**, with more than about 50% sand particles, have a coarse, gritty-feeling texture. These soils drain rapidly after a rain, and they hold water and cations poorly. Water and air penetrate sandy soils easily. Because of these characteristics, in temperate regions they warm readily in spring and cool quickly in autumn. Mineral nutrients in cation form are lost quickly, as are some pollutants.

Even a soil with as little as 35%–40% clay will exhibit the distinctive properties of clay (a little clay goes a long way). **Clayey soils** can be muddy or sticky in texture, can hold a large volume of water, and retain water and minerals exceptionally well. Similarly, clay-dominated soils retain pesticides, pollutants, and other substances. They are much less permeable to air and water than sandy soils, which can result in puddling, greater runoff, poor drainage, and poor aeration (due to water filling up the pores and excluding air). As a result of the water retained in the soil, in temperate regions clayey soils are slow to warm in spring and cool more slowly in autumn; this slows plant growth early in the growing season. **Silty soils** have a powdery (when dry), silky texture and tend to be intermediate in their characteristics and properties between sandy and clayey soils. Soils often possess characteristics of more than one of these texture classes.

To understand why different soil textures have the properties they do, we need to dig a little deeper. Sand and silt particles are generally irregular in shape, ranging from somewhat rectangular or blocky

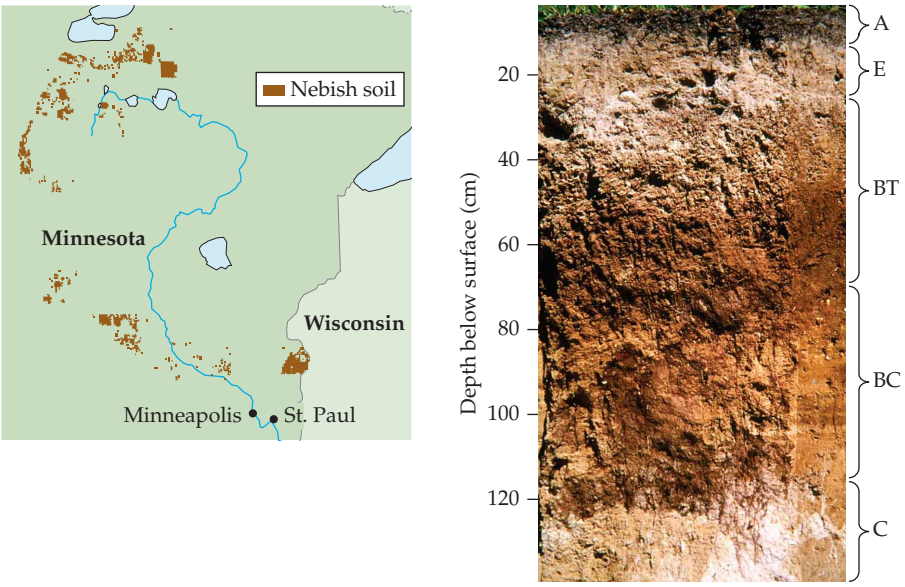



**Figure 4.3** An example of a soil profile of the Nebish soil series. This very deep, well-drained soil was formed on uplands in late Wisconsin calcareous glacial till and is in the order Alfisols. Although Alfisols are usually fertile and productive, the Nebish series is somewhat pathetically ineffectual in this regard, most likely because it is in a region with frigid temperatures that limit crop growth. It is mostly found in mixed deciduous-coniferous forests. Letters at right indicate soil horizons (see Figure 4.9). (Map from Soil Series Extent Explorer https://casoilresource.lawr.ucdavis.edu/see/#nebish California Soil Resource Lab University of CA, Davis. Soil profile from A. R. Aandahl. 1982. *Soils of the Great Plains: Land Use, Crops, and Grasses* by permission of the University of Nebraska Press: Lincoln, NE.)

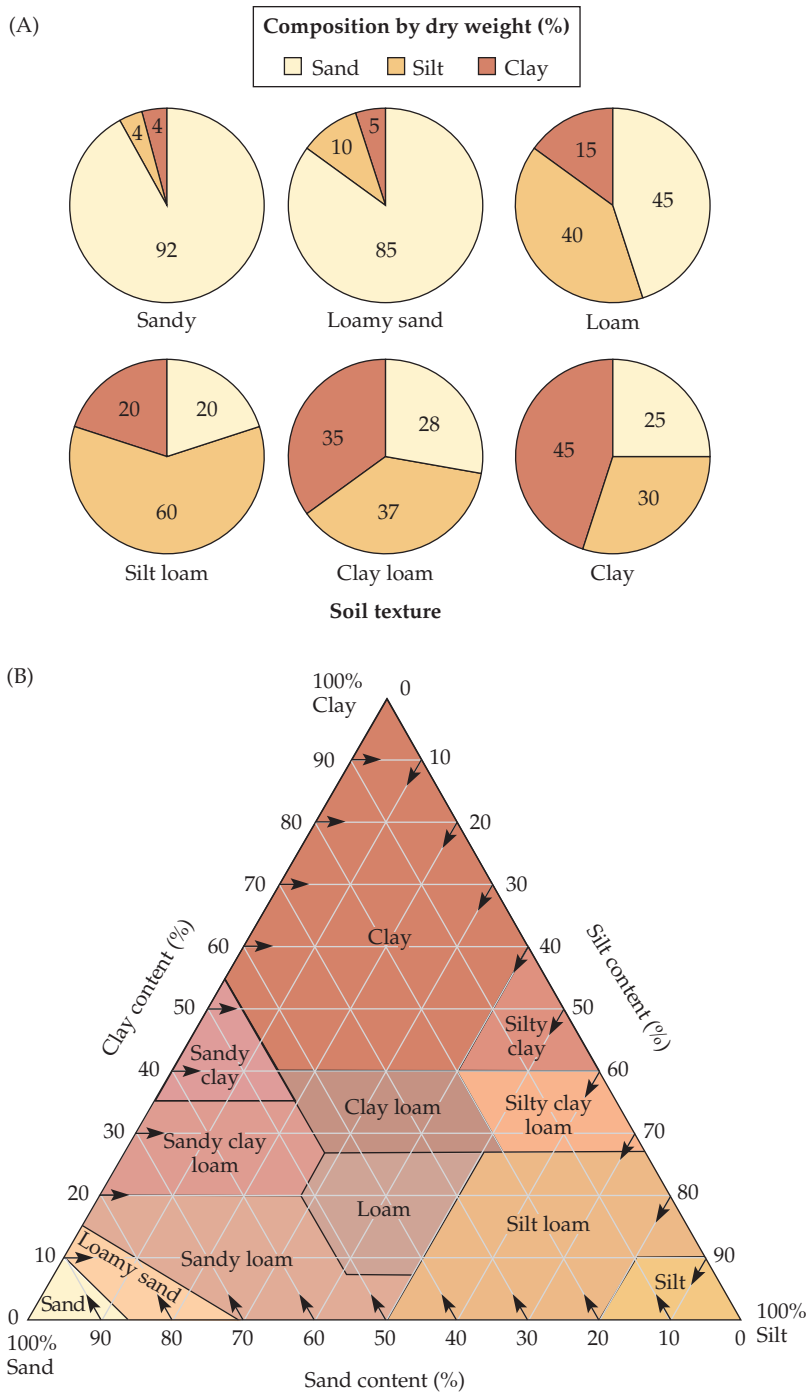



**Figure 4.4** Soil texture is determined by the relative proportions of different-sized particles making up the soil. (A) Examples of the soil particle size distributions of several different soil texture classes. (B) The soil texture triangle, which shows how soils are classified according to the percentage of sand, silt, and clay they contain (by dry weight). The grid lines representing values for clay are drawn parallel to the sand side of the triangle, those for silt are parallel to the clay side, and those for sand are parallel to the silt side. The points on the grid at which the three lines intersect define the type of soil; for example, the grid lines for a soil with 20% clay, 40% sand, and 40% silt intersect in the region that shows that a soil with this particular composition is a loam.

to chunky, spherical shapes (Figure 4.5A). In contrast, clay particles have a much more specific structure. These particles are made up of plates, whose size, shape, and arrangement depend on the minerals they contain and the conditions under which they were formed. A clay particle is made up of two or three flat crystalline plates, layered or laminated together (Figure 4.5B, C). These particles can be hexagonal in shape with distinct edges, or they may form irregularly shaped flakes or even rods. Soils may also contain particles similar in size to clays, but without the distinctive crystalline structure of clay particles. For example, an amorphous material called allophane, with particles about the same tiny size as clay particles, is prevalent in soils developed from volcanic ash.

Most of the clay particles in soil exist in a colloidal state (sand and silt particles are much too large to form colloids). In a **colloid**, fine particles of one substance are suspended and dispersed throughout a second substance. (Other examples of colloids are gelatin, fog, cytoplasm, blood, milk, and rubber.) Clay adds an enormous amount of external surface area because the particles have a large surface-to-volume ratio and because there are so many of them in a given volume of soil. In some kinds of clays, there is additional internal surface area between the plates (see Figure 4.5B, C). Thus, the huge surface area that characterizes clay particles in soil is due both to the fineness of the particles and to their platelike structure. The external surface area of a gram of fine clay is at least a thousand times that of a gram of coarse sand. If you spread out the surface area of the top 10 cm of the clay present in a half hectare of a clayey soil, it would cover the entire continental United States.

Because sand particles have a low surface-to-volume ratio, sandy soils have large, open pores between the mineral grains. Water drains easily from these pores because there is nothing to hold the water against the pull of gravity, and air penetrates them readily.

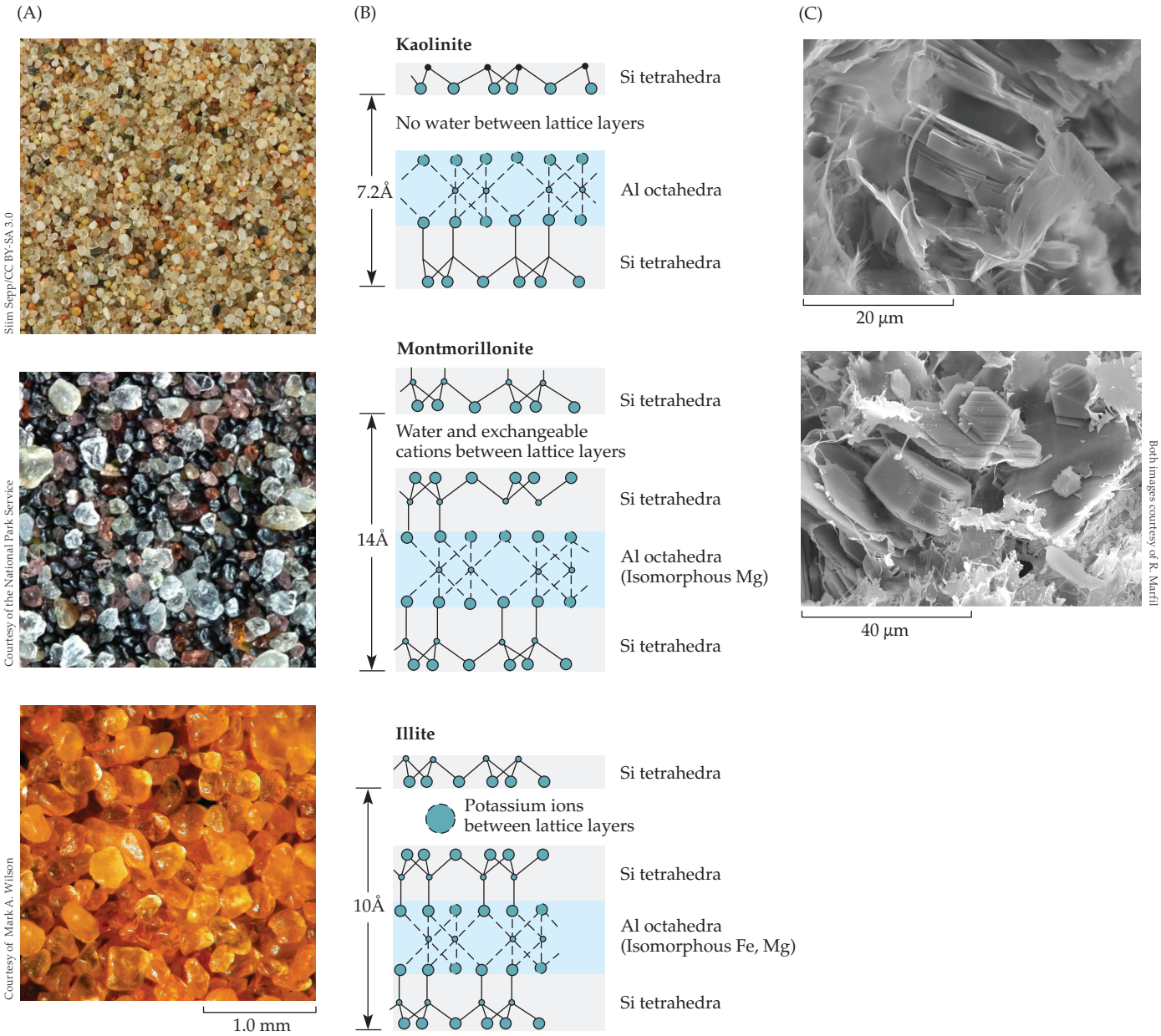



**Figure 4.5** (A) Sand grains are irregular in size and shape and are composed of different minerals depending on location and parent materials. Rounded and fine-grained eolian sand sample (top) from the Gobi desert (near Dalanzadgad in Mongolia). The width of the view is 10 mm. Sand grains from Fire Island National Seashore, New York (center), consisting of quartz with garnet, magnetite, and feldspar, tourmaline, shell fragments, and other mineral grains. Sand from Coral Pink Sand Dunes State Park, Utah (bottom). (B) Clay particles have a distinctive crystalline structure based on layers sandwiched together consisting of silicon (Si) and aluminum atoms (Al; arranged respectively in octahedral or tetrahedral shapes, represented by circles with lines indicating chemical bonds). The sandwiched sheets form lattice layers (plates) separated by spaces. The three major kinds of clay particles are shown here. Montmorillonite has the greatest cation exchange capacity (ability to hold cations) of the three because, in addition to binding cations on its surface, it has a large interior exchange surface available between the lattice layers. Contrast the highly structured form of these clay particles with the rather formless sand particles in part A. The clay particles are also far smaller than sand particles. (C) Scanning electron microscopy (SEM) images of clay particles showing (top) lath-shaped particles of illite, (bottom) pore-filling euhedral crystals of dickite displaying blocky morphology; from Khatatba sandstone, desert region of northwestern Egypt. (B after J. R. Etherington. 1982. *Environment and Plant Ecology* 2nd ed. John Wiley and Sons: New York, NY.)

Clay-dominated soils have smaller pores but a much larger number of pores than do sand-dominated soils. The total amount of pore space—the total proportion of the soil occupied by air and water—is greater in clayey soils (50%–60%) than in sandy soils (35%–50%), because of both the number and arrangement of the pores. Clay particles (as well as particles of organic matter) tend to cluster together, forming porous aggregates, resulting in much more pore space than in sands, in which the particles lie close together (Figure 4.6). Soils with native vegetation often have more pore space than those that have been cultivated, because tilling disrupts the aggregates and partially destroys the soil structure. Paving, building, trampling, and driving heavy equipment on and over soils compacts them, drastically and permanently changing their structure.

The greater pore space in clay-dominated soils is one factor contributing to the greater amount of water they can hold, but another factor is even more important. Unlike sand and silt particles, clay particles typically bear a strong negative electrochemical charge. Thus, they act as anions in the soil, attracting both **cations** (positively charged ions, which include the major plant nutrients; see Chapter 5) and water molecules to their surfaces (Figure 4.7). The role of clay particles in **adsorbing** nutrient cations—attracting and holding them to their surfaces—is one of their most important properties for plants.

Although many kinds of cations are attracted to clay particles, certain cations are most common and most important for plant growth. In humid regions, hydrogen and calcium ions ($H^+$ and $Ca^{2+}$) are usually most abundant in soils, followed by magnesium ($Mg^{2+}$), potassium ($K^+$), and sodium ($Na^+$) ions. In soils of arid regions, hydrogen ions move to last place in this list, and sodium ions become more important. These positively charged ions attract numerous molecules of water to their surfaces, adding to the overall capacity of the clay particles to retain moisture in the soil.

The cations adsorbed on clay particles are partially available to be taken up by plants. There is a continuous, dynamic interaction between the ions adsorbed on clay and other colloidal particles and those in the **soil solution**—the water in the soil and its associated dissolved minerals. Ions that are displaced from their positions on the surface of a particle enter the soil solution, from which they can be taken up by plants, **leached** (lost from the surface soil in the water that drains out of the soil into runoff and groundwater), or adsorbed on another particle. The behavior of these ions differs greatly among soils of different origins, textures, and chemical compositions and among regions differing in temperature and especially in rainfall.

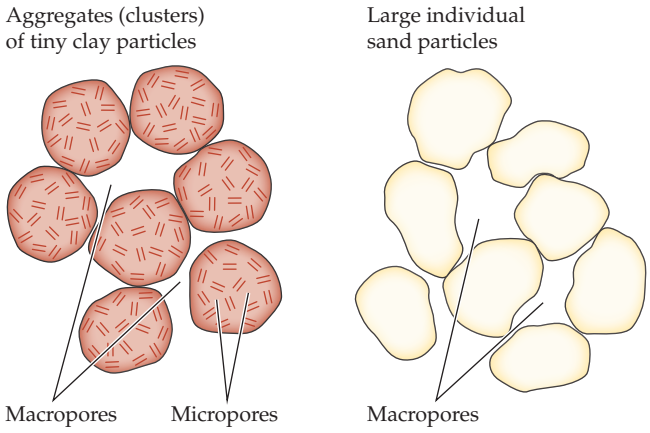



**Figure 4.6** Clay particles (left) are charged, and they aggregate (cluster together), with micropores between the particles and with macropores between the clusters. Sand particles (right) are much larger than clay particles and have no charge; macropores occur between the individual sand particles. (After N. C. Brady and R. R. Weil. 2008. *The Nature and Properties of Soils*, 14th ed. Pearson: New York, NY.)

## Soil pH has profound but indirect effects

The **pH** of soil—the negative logarithm of the concentration of $H^+$ ions in the soil solution—varies widely among different soils. The pH scale ranges from 1 to 14, where 7.0 is neutral (the pH of pure water); acidic soils have lower pH numbers and higher $H^+$ ion concentrations. Soils in the United States can range from less than 3.5 (e.g., in the pine barrens of Long Island, New York, and New Jersey) to as high as 10 (e.g., in arid grasslands of the southwestern United States).

For a soil to have a pH above 7 (neutral soil), it must be calcareous (containing calcium carbonate, $CaCO_3$), sodic (containing sodium carbonate, $Na_2CO_3$), or dolomitic (containing dolomite, $CaCO_3 \bullet MgCO_3$). Most crops grow best in slightly acidic soils, but native vegetation can be adapted to anything from very acidic to neutral to alkaline soils.

Soil pH has a great effect on plant growth and in determining what species can survive and grow in the soil. It mostly acts indirectly on plants by affecting the availability of mineral nutrients and toxic materials and by affecting the activity of some soil organisms (such as bacteria and fungi). These effects change the conditions for plant growth in a complex manner by altering the degree to which various nutrients are bound to soil particles, as well as other aspects of soil chemistry. For example, at very low soil pH, it is more difficult for plants to take up soil nitrogen and phosphorus, while the toxicity of aluminum is greater because it is more available to plants. In contrast, iron is readily available to plants in low pH soils but mostly unavailable at high pH.

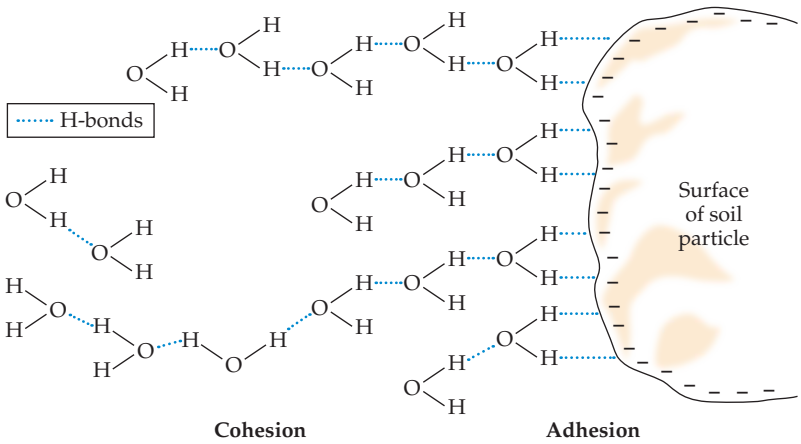



**Figure 4.7** Hydrogen bonds cause the adhesion of the hydrogen in water molecules attaching to the surface of negatively charged soil particles like clay and humus, and cohesion attaching the water molecules to one another. Hydrogen bonds are not that strong, though, and clusters of water molecules are constantly forming and dissociating. Water molecules, with two hydrogen atoms and one oxygen atom arranged in a V shape, are polar—the oxygen end has a negative charge, and the hydrogen end is positively charged. Because cations like $Na^+$, $Ca^{2+}$, and $K^+$ are positively charged, they are attracted to the oxgen end of water molecules. Water molecules stick to one another because the hydrogen of one water molecule is attracted to the oxygen of another one, forming a hydrogen bond. Many of the properties of water are a result of hydrogen bonds formed between water molecules, including its high specific heat, high boiling temperature, and high surface tension. (After N. C. Brady and R. R. Weil. 2008. *The Nature and Properties of Soils*, 14th ed. Pearson: New York, NY.)

Different plant species respond differently to soil pH because they are adapted to soils with different properties. Forest trees can grow over a range of soil pH but are especially tolerant of acidic soils. Many conifers and some other tree species are adapted to acidic soils and also tend to increase the acidity of the soils in which they are growing, primarily through the properties of the **litter** (needles, leaves, and other material) they shed. Grasslands tend to be found on relatively alkaline soils, although the relationship between grass species and soil pH may be indirect, because low rainfall results in soils with high pH (as we will see below) and also favors grass-dominated vegetation. There are other characteristic associations between soil pH and plants. Species in the Ericaceae, for example, such as heaths and blueberries, tend to grow only in very acidic soils, while other taxa, such as *Larrea tridentata* (creosote bush, Zygophyllaceae; see Figure 3.6A), are typically found in alkaline soils.

Soil pH strongly affects cation availability to plants. Cations bind loosely to clay particles, which are generally negatively charged. Under acidic conditions, the excess $H^+$ ions tend to bind more strongly to the clay particles than nutrient ions do, displacing these nutrients into the soil solution. As a result, mild acidity generally promotes nutrient availability. Under extreme acidity (whether natural or due to acid precipitation), however, the nutrient cations are so mobile that they are easily leached, and they are carried off in **groundwater** (water found underground in aquifers, rock crevices, and so on).

What determines the pH of a soil? This simple question can have a variety of complex answers because the physical and chemical properties of soils are both involved, as are biotic and external influences like climate. Recall that the more $H^+$ ions in a solution, the more acidic it is, and the lower the pH. The most common source of $H^+$ ions comes from the formation and dissociation of carbonic acid ($H_2CO_3$), formed when $CO_2$ in soil air pores joins with water and then dissociates into $HCO_3^- + H^+$. Hydrogen ions are continually added to the soil by decaying organic matter, roots, and various soil organisms. They are actively exchanged between the surfaces of colloidal particles in the soil and the soil solution, contributing directly to the acidity of the soil. Aluminum, which is a part of many soil minerals, is also important in determining soil pH. Aluminum ions ($Al^{3+}$) indirectly cause hydrogen ions to be released from colloidal particles by reacting with water to release $H^+$ and $OH^-$ ions. The $OH^-$ ions join the $Al^{3+}$ to form hydroxy aluminum ions, $AlOH^{2+}$, $Al(OH)_2^+$ and $Al(OH)_3$. The $H^+$ ions that were released lower the pH of the soil solution, increasing soil acidity. The hydroxy aluminum ions (the various positively charged ions of $Al(OH)_x^{y+}$) are attracted to and tightly adsorbed onto colloidal clay and humus (organic matter) particles, replacing the $H^+$ ions that had been adsorbed onto those particles. As these $H^+$ ions are released into the soil solution, this further increases soil acidity. High rainfall levels favor the predominance of aluminum and hydrogen ions, making the soil more acidic, in part because the other cations are more readily leached and thus lost from the soil. Aluminum ions are highly toxic to plants; also, because they occupy many of the sites otherwise available to attract cations onto the negatively charged clay and humus particles, they make the nutrient cations less available to plants when soil pH is low.

The **exchangeable bases** consist of most of the other cations (other than the hydrogen and aluminum ions). These positively charged ions, mostly $Ca^{2+}$, $Mg^{2+}$, $K^+$, and $Na^+$, attach to clay and organic particles and can be exchanged with each other on the particles and in the soil solution. They contribute to making the soil more

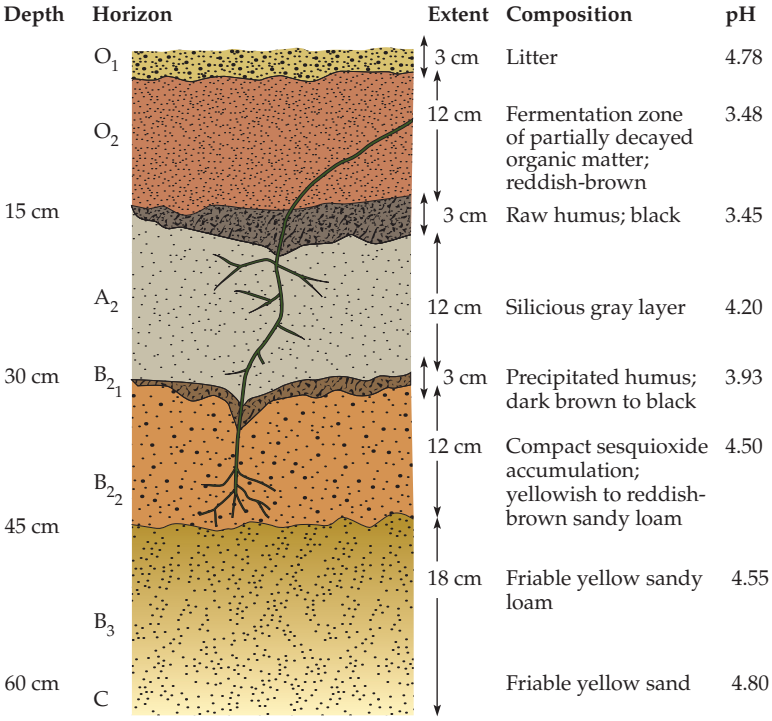
alkaline. The **cation exchange capacity** (**CEC**) of a soil is a measure of the total ability of the soil colloids to adsorb cations (in units of centimoles of positive charge per kilogram of dry soil, $cmol_c/kg$). The **percentage base saturation** is the proportion of the CEC that is occupied by exchangeable bases. In arid regions the bases are not leached out of the soil by rainfall, so the percentage base saturation is very high (90%–100%), $H^+$ concentration is low, and the soils tend to be alkaline. In areas with higher rainfall, the bases are leached more easily while $H^+$ and $Al^{3+}$ are retained, so the percentage base saturation is much lower (50%–70%), and the soils tend to be acidic.

Plants and soil organisms are an important source of soil acidity. When roots or soil organisms respire, they generate $CO_2$. In wet soils, this $CO_2$ goes into solution immediately, creating a weak acid, carbonic acid ($H_2CO_3$). This is also the reason why rainwater is naturally somewhat acidic: $CO_2$ from the atmosphere dissolves into the raindrops. In tropical rainforests, where conditions favor massive amounts of respiration and there are often few clay particles to bind cations, this added $CO_2$ can help make the soils quite acidic and promotes the mobility of cations. The consequence is rapid uptake of nutrients by plants in undisturbed rainforests, and rapid loss of those nutrients when forests are cleared.

In forests in hilly areas, soil pH tends to increase as you go from ridges downslope. This happens because bicarbonate ions ($HCO_3^-$), formed from $CO_2$ in the air dissolved into soil water, leach downslope, reacting with $H^+$ ions and removing them from the soil solution, making it less acidic. In areas with acidic forest soils, the pH near roads may be much higher than in the surrounding natural areas. A major component of paved roads is concrete, and the cement in concrete has a very high pH (around 11). The concrete leaches ions to the roadside soils, reducing the pH and strength of the concrete and greatly increasing the pH of the acidic soils, which can have large (but often overlooked) effects on roadside vegetation.

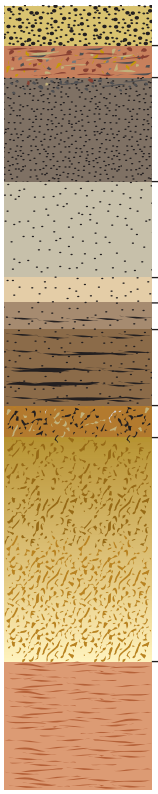
## Soils are characterized by horizons—layers with distinctive properties

Soils are not homogeneous; as Hans Jenny appreciated so well, they contain characteristic layers, or **horizons**, that differ among soil types, series, and even orders. If you look at roadcuts or excavations dug for construction, you can often see the horizons quite clearly. The sequence of horizons that characterizes a soil is called the **soil profile** (Figure 4.8). The horizons are grouped into four categories: O, A, B, and C. These are each divided into subcategories that are numbered according to their particular characteristics (Figure 4.9). The roots of species with different growth forms are generally concentrated in different parts of the soil profile; for example, annuals and small herbaceous perennials typically have shallower roots, while prairie grasses and woody species have roots that extend deeper into the soil profile (see Figure 3.19). The properties of the horizons differ and affect water and nutrient uptake by plant roots differently.



Courtesy of J. Gurevitch

**Figure 4.8** Profile of a forest soil at the edge of the Adirondack Mountains in northern New York State (in the Spodosol soil order). This soil is a stony loam, forested with birch, hemlock, and spruce, and is quite acidic throughout the profile. There is a layer of litter and a thick organic horizon at the top of the profile, and there are several distinct mineral horizons, each with particular properties and a characteristic appearance. Roots are shown reaching down to the top of the $B_3$ horizon at about 45 cm deep, although the deeper roots of large trees would certainly penetrate more deeply into this soil. (After H. O. Buckman and N. C. Brady. 1960. *The Nature and Properties of Soils*, 6th ed. Macmillan: New York, NY.) The inset shows a landscape in the Adirondack mountains.



| Horizon | Properties |
|---|---|
| $O_1$ | Plant litter and organic debris in various stages of decay |
| $O_2$ | Partially decomposed organic debris |
| $A_1$ | Mineral soil with organic material, deep brown to black |
| $A_2$ | Silicate clays and iron and aluminum oxides eluviated (removed, deposited in B horizon), may be called E |
| $A_3$ | Dominated by A properties, transitional horizon between A and B |
| $B_1$ | Dominated by B properties, transitional horizon between A and B |
| $B_2$ | Zone of illuviation (deposition) of silica clay particles, iron and aluminum oxides, organic matter |
| $B_3$ | Dominated by B properties, transitional horizon between B and C |
| C | Partially weathered rock, minor physical, chemical or color development; varies among different soils |
| R | Unweathered bedrock or other material, such as loess |

**Figure 4.9** An abstract, general soil profile showing the major horizons that might be present in particular soils. No one soil is likely to have all of the horizons shown, and particular soils may have greater development of subhorizons than what is shown. In the C horizon, only the upper part is considered to be part of the soil proper. The depth of soils varies tremendously with the location and nature of the soil, but to gain some perspective, readers might picture this illustrated profile as being about a meter deep to the top of the bedrock. (After H. O. Buckman and N. C. Brady. 1969. *The Nature and Properties of Soils*, 7th ed. Macmillan: New York, NY.)

The O horizons consist of organic material formed above the mineral soil, derived from decaying plant materials, microbial matter, and the remains and waste products of animals. The A horizons—the surface layer of mineral soil—represent the region of maximum leaching, or **eluviation**. The uppermost A horizon, the $A_1$, is often darker than the rest of the soil and may contain highly decayed organic matter. The B horizons, deeper in the soil, represent the region of maximum **illuviation**, or deposition of minerals and colloidal particles leached from above. Clays, iron, and aluminum oxides often accumulate in the $B_2$ horizon. The C horizon is the undeveloped mineral material deep in the soil; it may or may not be the same as the parent material from which the soil develops. Below that may be bedrock, or just deep accumulations of mineral material deposited by wind, water, or glaciers.

No one soil has all of the horizons pictured in Figure 4.9. Some horizons may be much more distinct and better developed than others. Plowing and the activity of earthworms may obscure the distinction between the upper horizons. The soil profile may not be fully developed, and some horizons may be absent or indistinct, if the soil is relatively young. Upper horizons may have been lost through erosion (Box 4B), leaving the deeper horizons (subsoil) exposed at the surface.

Soils greatly vary in depth, from thin layers of soil barely covering a rock substrate (e.g., in many alpine areas) to very deep soils of close to 2 m (e.g., in some well-developed prairie soils). Soil depth has a great deal of influence on vegetation and plant growth: the greater the soil depth, the more favorable the soil for plant growth, all other things being equal. Deeper soils can hold more water and nutrients and can retain water for a longer time without rainfall, and they allow greater development of plant root systems. Plant species able to live on very shallow soils have particular adaptations for survival under these conditions, and they often have very distinctive growth forms (Figure 4.10).

Soils in arid regions may accumulate thick layers of calcium carbonate ($CaCO_3$), called a **calcic horizon** and known in the desert southwest in the U.S. as caliche

Courtesy of J. Gurevitch

**Figure 4.10** Plants growing in shallow, rocky soils sometimes are dwarfed or have distinctive growth forms. *Pinus rigida* (pitch pines, Pinaceae) in the Shawangunk Mountains, west of the Hudson Valley, New York, U.S.A.

## BOX 4B Soil Conservation Is a Major Global Environmental Issue

**Soil erosion** is the removal of the terrestrial substrate by liquid water, ice, or wind, in the presence of gravity. Gravity itself can cause wet or dry soil and even rocks and boulders to move downhill when dislodged by earthquakes or other causes. Erosion is a natural process; it has always occurred for terrestrial substrates on our wet, windy planet where temperatures change, ice forms and thaws, and storms pound the ground. It is most severe on steep slopes and where the substrate is loose and not anchored in place. Accelerated or **anthropogenic erosion** (sometimes just called erosion) occurs on a much shorter time scale and is due directly or indirectly to human activities.

Erosion due to human activities is greatest on steep slopes where forests have been clear-cut, on soils that have been cultivated for a long time, especially when using poor farming practices, and on lands that have been severely overgrazed. The removal of perennial grasses or trees facilitates erosion, especially when followed by plowing or other processes like paving that prevent deep, long-lived roots from becoming reestablished. Removal of soil for urbanization and road building are major factors in soil erosion and destruction. When plant roots are no longer anchoring the soil, it is vulnerable to being removed rapidly by wind or water, and these effects are most noticeable and severe on slopes. Soil erosion has had huge economic and social consequences for human societies and has even precipitated the collapse of several Old and New World civilizations. It can result in the loss of the upper layers of soil, in the creation of deep rills and gullies in the soil, or even in the removal of all of the soil down to bedrock. It remains one of the most critical (and most widely ignored) environmental problems. It occurs in many different parts of the world and affects many natural and seminatural ecosystems. It compromises our ability to raise food crops and can devastate the sustainability of natural ecosystems. It can have major consequences for people's homes, livelihoods and communities. Rebuilding soils to replace losses to erosion can take tens of thousands of years.

Plants can affect water fluxes into the soil by intercepting precipitation, reducing the impact of rain hitting the ground, and decreasing runoff from the soil. If vegetation is removed from an area—for example, by clear-cutting a forest—runoff and soil erosion can be greatly increased, particularly in areas where the terrain is hilly or mountainous. This can sometimes result in flooding and mudslides, further damaging vegetation and threatening human lives and homes. Even in semiarid regions, removal of vegetation can result in reduced precipitation, increased soil warming, and the onset of desertification (Schlesinger et al. 1990; Dirmeyer and Shukla 1996).

An important component of the organisms living in and on soils, particularly in arid regions, are those composing soil crusts. Soil crusts can include mosses, cyanobacteria, lichens, algae, and fungi. They are critically important in ecosystem function and affect above- and belowground interactions with other organisms, as well as interaction with the physical structure of the soil. They are highly vulnerable to disturbance by people and domestic grazing animals and may never recover after trampling, plowing, or damage due to vehicles driven on them.

John M. Holsenbeck Jr./NWS

A large dust storm hitting Lubbock, Texas, in October 2011.

Hilda Smith, USGS

Semiarid grassland vegetation and biocrusts in Utah. The biocrusts are the darker-colored patches between the bunches of grasses and cacti.

(Schlesinger 1982; see Figure 4.1). The calcium carbonate comes from dust deposited on the soil surface and ultimately from rocks. In areas with higher rainfall, the calcium carbonate dissolves and is leached from the soil into the groundwater. In arid and semiarid regions, it is deposited at various depths and accumulates over

very long periods of time. These high pH layers have abundant calcium but limit availability of many other necessary nutrients, including phosphorus and iron. They can be soft and porous or hard, concrete-like layers, often white or gray, that inhibit root penetration. The carbon in the calcium carbonate in such soils constitutes a substantial pool of stored carbon worldwide (Schlesinger 1982; Schlesinger et al. 2009).

## Soils are the unique product of living organisms acting on soil parent material

So far we have emphasized the role of physical processes in soil development. We began this chapter, however, by emphasizing that soils are the unique product of living things acting on the physical environment (and in turn being affected by that environment). What are some of the ways in which organisms affect soils?

**Organic matter** is the decaying and decomposed material in soil that comes from living things. It includes substances secreted by plants, microorganisms, and animals, products of excretion by animals, parts shed by animals and plants, and dead organisms and parts of organisms. Once decomposed, it becomes humus, which is the part of the organic matter left in soils after plant and animal matter is fully decomposed. Microscopic organisms, while tiny individually, collectively contribute a great deal of organic matter to the soil. The activities of various kinds of microorganisms, including the recycling of mineral nutrients used by plants, affect soil properties in many additional ways (see Chapter 5).

Organic matter binds mineral particles together to stabilize soil aggregates, reducing compaction. This maintains soil porosity despite the physical action of wetting and drying, or freezing and thawing. The porosity of soil contributes to how easily roots or fungal hyphae can penetrate it, and how easily animals can burrow and tunnel through it. These actions of organisms moving through the soil help to create pores and alter soil structure.

Organic matter is amorphous (it does not have a well-defined structure, unlike clay particles), and its particles typically have very large surface areas. Humus particles are colloids of organic origin and are not crystalline. Humic acids often consist of the basic materials of which living organisms are made: complex chains and rings of carbon atoms with hydrogen, oxygen, and nitrogen bound to them (Figure 4.11). These particles have the ability to absorb and retain large amounts of water. Organic matter is also negatively charged, holding $H^+$ ions on its surface, buffering soil pH. Soil organic matter is important in plant nutrition because the negatively charged particles retain nutrient cations on their very large surface areas. Fresh and decayed organic matter also contains organic acids and other compounds that chemically alter essential plant mineral nutrients such as calcium (Ca), iron (Fe), manganese (Mn), copper (Cu), and zinc (Zn), making them more readily available to plants.

The organic matter content of soils varies geographically and is highest in wetland regions, grasslands, and tundra. The amount of organic matter in a soil depends especially on how much carbon enters the soil every year (Figure 4.12A). The **turnover time** for organic matter (the mass of organic matter divided by how quickly it moves in and out of the soil; see Box 5.1) depends most on evapotranspiration and also on precipitation and temperature (Figure 4.12B). As we will see in later chapters, soil organic matter and turnover in soil organic

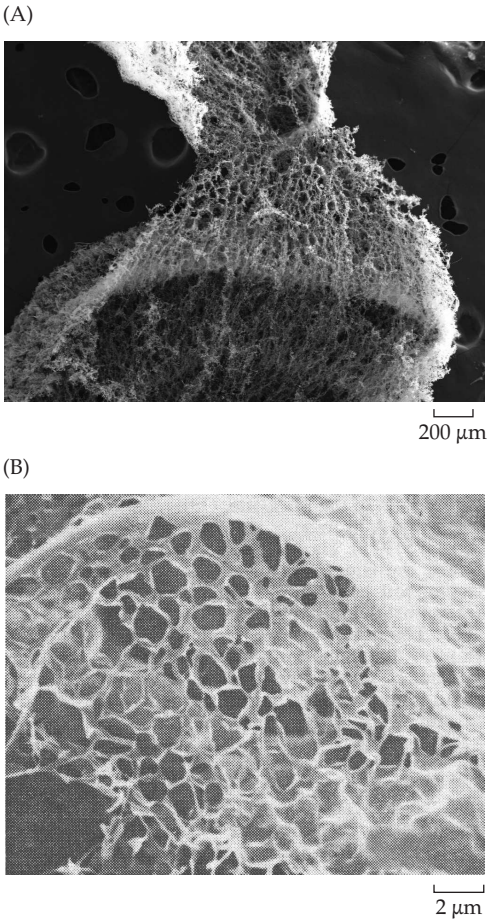



**Figure 4.11** SEMs of soil organic matter. (A) Humic acid isolated from lignite composed of nanotube membranes assemblies with their characteristic structural honeycomb pattern. (B) Fulvic acids isolated from a Cecil soil (Ultisols) in Georgia, U.S.A. showing a "fishnet network" or "chicken-wire" pattern of carbon nanotube assemblies in hexagonal structural arrangements. (From K. H. Tan. 2014. *Humic Matter in Soil and the Environment: Principles and Controversies*, 2nd ed. CRC Press: Boca Raton, FL. SEM micrographs © 2011 K.H. Tan, Professor Emeritus, University of GA, Athens, GA.)

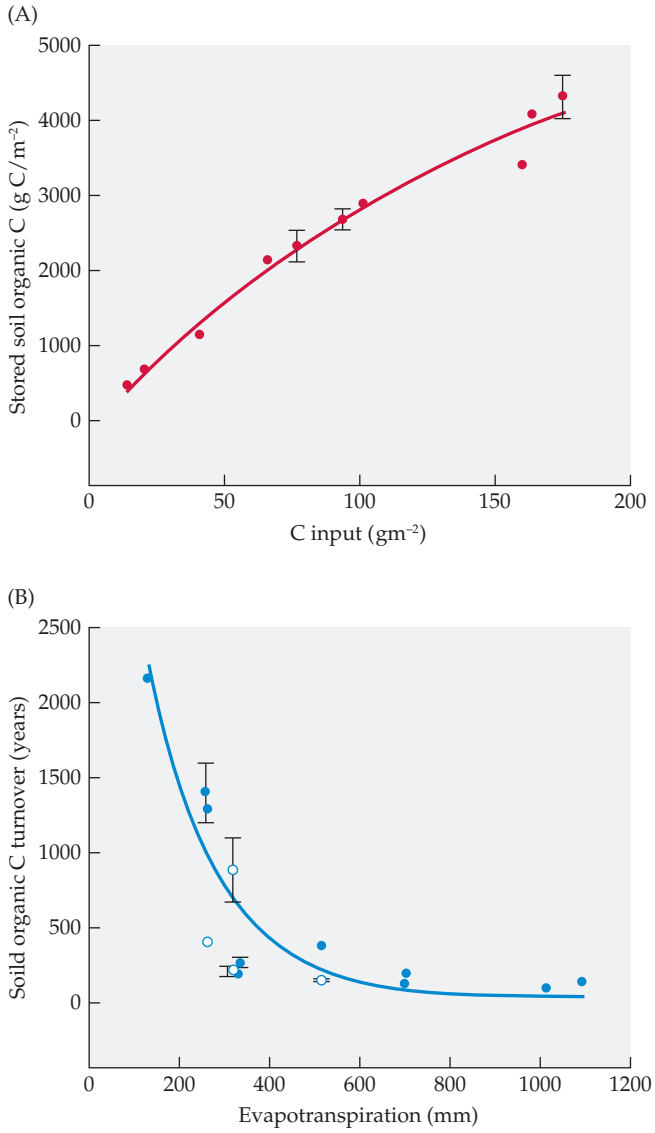



**Figure 4.12** (A) Soil organic carbon (in g C/m$^{-2}$) increases as the amount of carbon added to the 0–10 cm mineral soil carbon pool increases. (B) Soil organic turnover time varies with evapotranspiration; higher evapotranspiration results in more rapid turnover (shorter turnover times). (After D. A. Frank et al. 2012. *Ecosystems* 15: 604–615.)

matter have large implications for many ecological and environmental issues, including agricultural productivity, other ecosystem services, and global climate change (see Chapter 5 and Chapter 16).

## 4.2 The Rhizosphere Is a Unique Environment Created by Roots and Their Interaction with Microorganisms

Plant roots interact with the soil that immediately surrounds them to create a unique environment both for the roots themselves and for a rich community of other organisms, particularly microorganisms. This environment, defined as the region affected by plant roots and containing large associated microbial populations, is called the **rhizosphere**. Plant roots release many different chemicals—called **root exudates**—that attract, sustain, and interact in complex ways with the many different species of bacteria, fungi, nematodes, and other organisms living in that environment. The nature of these diverse communities can vary greatly depending on the plant species, the soil type, climate and other aspects of the environment, and of course, the presence and interactions of the soil organisms themselves. The physical and chemical interactions between the microorganisms and the plants transform the properties of the soil in the rhizosphere (Nihorimbere et al. 2011; Huang et al. 2014). Not only can plant exudates influence soil organisms, but also interactions between plant roots and the organisms in the rhizosphere can influence plant growth and even the interactions between aboveground herbivores and plant leaves (Badri et al. 2013). The end result is a complex set of feedbacks between the biotic and abiotic components of the soil.

Root exudates can include large, complex molecules, like polysaccharides and glycoproteins, and smaller molecules, such as amino acids and carbohydrates, as well as alkaloids, antibiotics, and hormones. These molecules not only support the microbial community in the soil by providing carbon and energy, but also mediate interactions between a plant's root and its associated microbial species, as well as among those species. For example, Changxun Fang and colleagues (2013) found that phenolic compounds produced by a strain of transgenic rice altered the bacterial and fungal community of its rhizosphere. Somewhat remarkably, plants can also recognize particular species or strains of microbes and communicate with them by producing signaling molecules that attract or repel particular mixes of microbial species (Haichar et al. 2014).

The microorganisms can have either harmful or beneficial effects, including protection from pathogens and enhanced nutrient uptake. Jos Raaijmakers and Mark Mazzola (2016) make the case that the disease protection aspect of this interaction is analogous to vertebrate immune systems. General suppression of pathogenic microbes appears to occur largely through interactions with the beneficial organisms. Suppression occurs when infected plants produce exudates that attract pathogen-suppressing microorganisms that in turn produce antimicrobial compounds or other means of suppressing the pathogens.

A great deal has been learned about the genomic and metabolic bases of these interactions in recent years. Although much of this has focused on agricultural crops rather than wild species, it is likely that similar processes occur in noncultivated species (Lagos et al. 2015). For

example, Serge Michalet and colleagues (2013) studied the plant-microbial interactions in a tropical tree from South America, *Eperua falcata* (Fabaceae). This species preferentially takes up nitrate rather than ammonium (see Chapter 5), and the authors hypothesized that root exudates might alter the activity or abundance of nitrifying and denitrifying bacteria in the soil, minimizing nitrate losses by inhibiting denitrifying bacteria in the rhizosphere. This appears to be the case, as they found that its root exudates caused selective metabolic inhibition of denitrifiers and that the abundance of denitrifying bacteria was suppressed in the rhizosphere.

## 4.3 Water Moves through the Soil to Reach Plants

Imagine a meadow or prairie in the summertime that has been without rain for some time. A thunderstorm sails in, offering a solid drenching. What happens within the soil? Pores that had been filled with air now fill with water, first in the upper layers of the soil profile, then deeper as more rain falls. Eventually the soil is **saturated** with water when almost all of the pores are filled. If rainfall continues, water will run off the soil surface. The amount of runoff depends on the vegetation, slope, and other factors (Figure 4.13). When a large amount of rain falls in a short time, the water cannot soak into the soil and drain to deeper layers quickly enough; runoff increases, sometimes resulting in severe flooding.

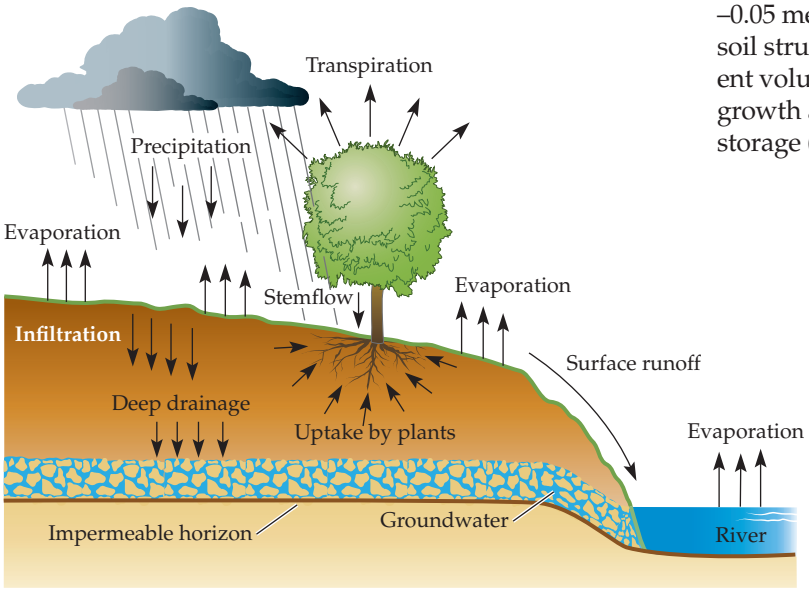



**Figure 4.13** The fate of water falling on the ground from a rainstorm. (After P. J. Kramer. 1983. *Water Relations of Plants*. Academic Press: New York, NY.)

The processes that take place when rain falls on dry ground are similar in many ways to those that take place in a suburban lawn, a forest, a tropical grassland, an agricultural field, a mediterranean chapparal, or even, to some extent, an urban vacant lot. One of the differences among these sites is what happens as the rain approaches the ground. In a forest, much of the water is intercepted by the leaves in the tree canopy, and it evaporates, runs down the tree trunks, or is scattered and falls to the ground with reduced force. In a modern agricultural field that has been pounded by repeated use of heavy farm equipment and compacted, some of the water will run off the ground rather than being absorbed. Runoff is even greater in urban settings where soil has been removed for roads, houses, and parking lots, and paved in impermeable surfaces, and along roadsides or other places where the ground is highly compacted and often has limited vegetation. The runoff water is likely to carry some soil with it, along with nutrients like nitrogen. This loss of soil is a serious conservation issue (see Box 4B). Soils on steep slopes are more vulnerable to soil erosion than those that are on level ground, but runoff can occur anywhere. Heavy and prolonged downpours are increasingly common as a result of climate change, making things worse (see Chapter 16).

Rainwater entering the soil immediately begins to drain downward due to the pull of gravity, eventually reaching the groundwater and draining into the streams, rivers, and lakes of the watershed. After about a day, this rapid downward movement of water slows, and many of the largest pores in the soil refill with air (Figure 4.14A). The smallest pores, however, remain filled with water, and the soil is at **field capacity** (Figure 4.14B). The water potential (see Chapter 3) of the soil is now at –0.01 to –0.05 megapascals (MPa). Depending on the soil texture, soil structure, and soil depth, soils may hold very different volumes of water, with major consequences for plant growth and survival and affecting water movement and storage (hydrology) at landscape scales (see Chapter 15).

Water is held in the soil largely by attraction to the surfaces of soil particles, particularly particles of clay and organic matter. It moves through small pores and the film of water surrounding mineral particles by **capillary action**, a process that occurs within narrow tubes or in a surface film of water (such as within soil pores, the xylem, and the substomatal chambers of leaves). Water is pulled upward (or horizontally) by the attraction of the water molecules to the charged surfaces and to one another. As the plants in our hypothetical meadow transpire, a water potential gradient (see Figure 3.1 and Figure 3.2) is established, and soil water moves toward the roots and is absorbed by them.

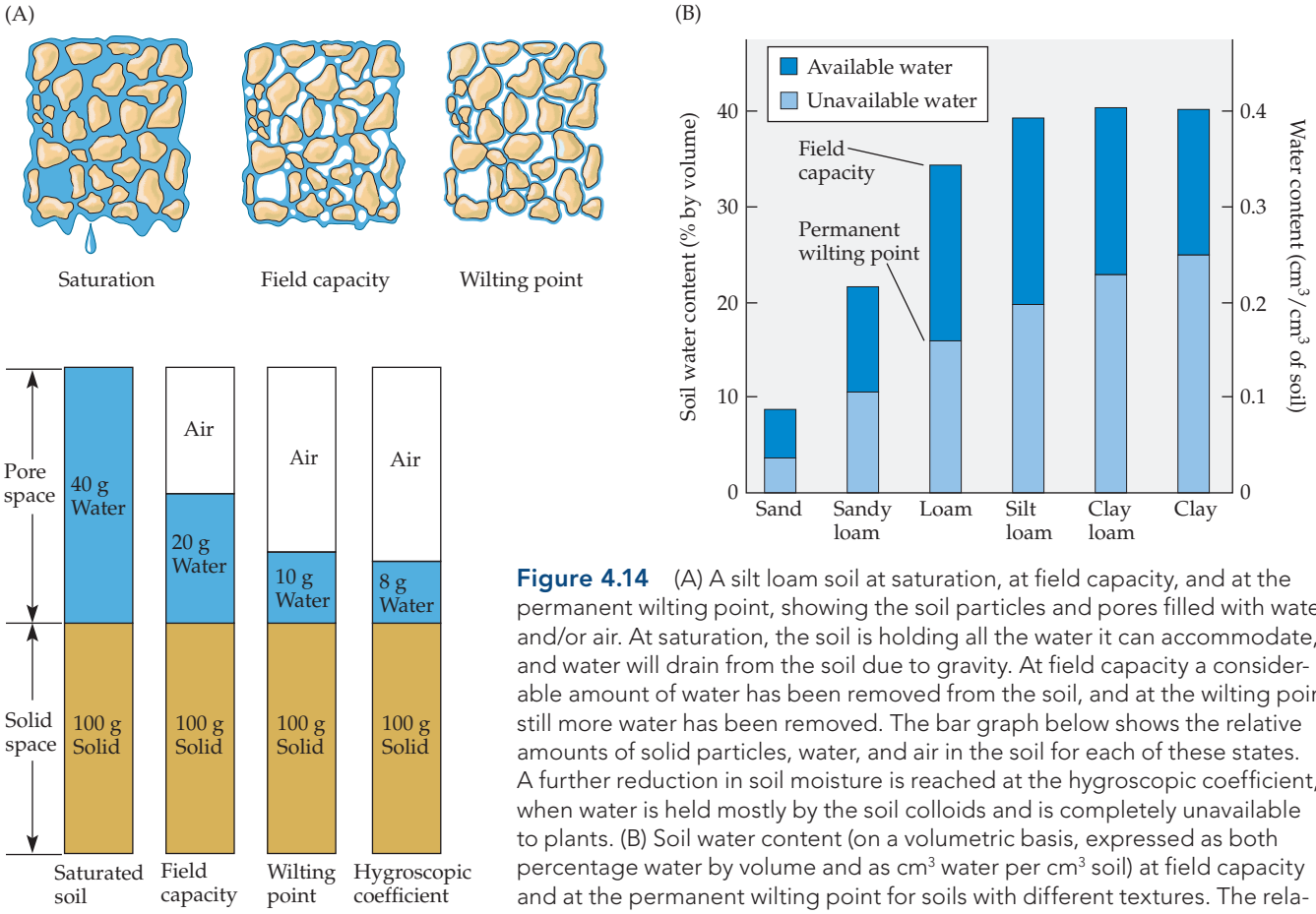



**Figure 4.14** (A) A silt loam soil at saturation, at field capacity, and at the permanent wilting point, showing the soil particles and pores filled with water and/or air. At saturation, the soil is holding all the water it can accommodate, and water will drain from the soil due to gravity. At field capacity a considerable amount of water has been removed from the soil, and at the wilting point still more water has been removed. The bar graph below shows the relative amounts of solid particles, water, and air in the soil for each of these states. A further reduction in soil moisture is reached at the hygroscopic coefficient, when water is held mostly by the soil colloids and is completely unavailable to plants. (B) Soil water content (on a volumetric basis, expressed as both percentage water by volume and as $cm^3$ water per $cm^3$ soil) at field capacity and at the permanent wilting point for soils with different textures. The relative amounts of available and unavailable water in soils of different textures are also shown. Sand has the lowest total water content and the least available water at field capacity, as well as the least unavailable water remaining at the permanent wilting point. At field capacity, clay soils have the greatest total amount of water and the greatest amount of unavailable water remaining at the permanent wilting point, all a consequence of the ability of the clay particles to attract water very effectively. Note that the "permanent wilting point" is where mesophytes (but not xerophytes) wilt and cannot recover, and "unavailable water" is unavailable to mesophytes such as crops but may be partially available to xerophytes (see Chapter 3). (A and B after H. O. Buckman and N. C. Brady. 1960. *The Nature and Properties of Soils*, 6th ed. Macmillan: New York, NY; A after L. R. Swarner. 1959. *Irrigation on Western Farms*. Agriculture Information Bulletin No. 199. U. S. Dept. of the Interior, Bureau of Reclamation and USDA, Soil Conservation Service. Washington D. C.)

As the soil continues to become drier, the small pores begin to be emptied of water, filling with air. The once continuous film of water is broken in many places, and water movement is greatly slowed. Water can continue to move through these empty pores as water vapor, but only small amounts of water can be transported in this way. Each day, as the plants transpire and the soil progressively dries, the water potential of the plants declines (becomes more negative), and each night, as stomata close, the water potential of the plants rises somewhat as the plants equilibrate with the soil (Figure 4.15).

When the soil reaches a water potential of about –1.5 MPa, most mesic plants no longer can extract water from the soil and exist in a permanently wilted condition, about to die. This state is called the **permanent wilting point** of the soil, and the soil moisture content at this point is called the wilting coefficient of the soil. However, xerophytes adapted to dry environments (see Chapter 3) can continue to remove water from soils that are much drier. The amount (by weight or volume) of water held in the soil at any point in this drying curve will vary greatly with the soil texture (Figure 4.16) and also with the proportion of organic matter in the soil.

Water moves upward in the soil by being pulled by the transpiration of plants combined with water cohesion and capillary action, and it moves downward primarily by gravity. The movement upward occurs mostly via the small pores and micropores in the soil, and in the film of water surrounding soil particles, while the downward movement occurs largely via the largest

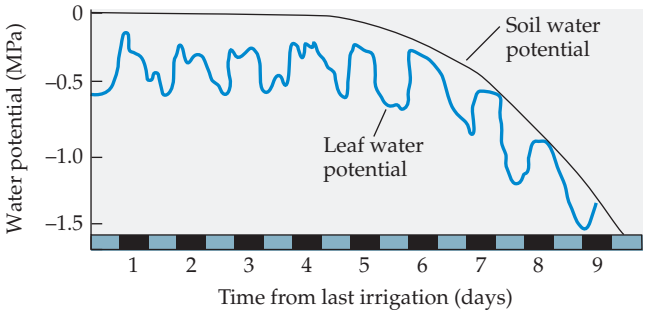



**Figure 4.15** Daily changes in soil and leaf water potential for a cultivated plant growing in a clay loam soil over the course of 9 days and nights after receiving irrigation. Black portions of the bar on the *x*-axis indicate nighttime conditions; blue portions indicate daytime. (After J. R. Etherington. 1975. *Environment and Plant Ecology*. John Wiley and Sons: New York, NY; based on W. R. Gardner and R. H. Nieman. 1964. *Science* 143: 1460–1462.)

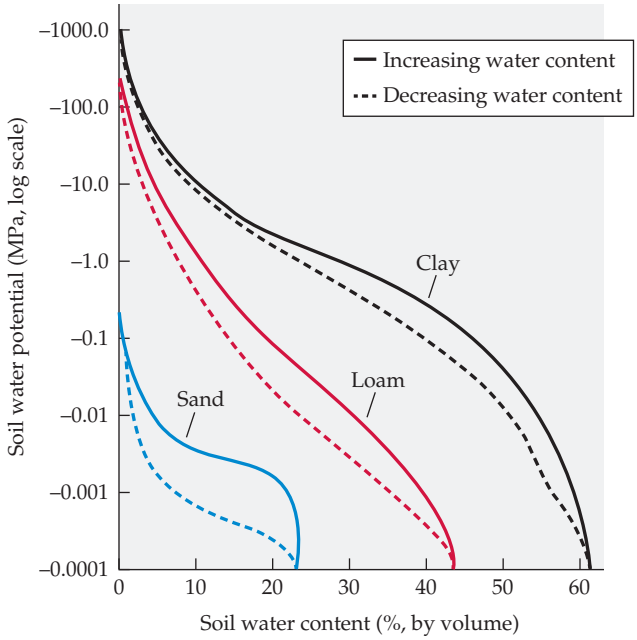



**Figure 4.16** Illustration of the general relationships between soil water content (percentage by volume) and soil water potential (MPa) for clay, loam, and sand. These curves differ depending on whether the soil was measured while water content was increasing or decreasing). (After V. V. Rendig and H. M. Taylor. 1989. *Principles of Soil-Plant Interrelationships*. McGraw Hill: New York, NY.)

pores between particles. As water evaporates from the soil surface, capillary action pulls water from deeper in the soil. The roots of transpiring plants also contribute to pulling water upward. Sometimes large amounts of water can be transferred from deep horizons to dry surface layers by deeply rooted plants. During the day, roots generally have more negative water potential than soils. At night, the dry upper layers of the soil may have more negative water potential than roots, and the soil may be capable of removing water from roots. In this way, deeply rooted plants may take up water from moist deep horizons, from which it travels up the root system and moves out of the upper part of the root system into the dry soil in the upper horizons. There, the water can become available to other, more shallowly rooted plants. This phenomenon, **hydraulic lift**, may in some cases allow the survival of plants with shallow roots during periods of drought or in arid environments. But water may be moved downward by roots as well. **Hydraulic redistribution** refers to rapid water redistribution in the soil in any direction by roots (Ryel et al. 2002). Hydraulic redistribution may alter the transpiration of the entire plant canopy substantially, particularly in arid environments, by making water available where it was not previously accessible (Ryel et al. 2003).

Soil provides the substrate in which most terrestrial plants are supported, the water that is essential for life (see Chapter 3), the environment in which roots interact with a complex network of organisms, and the mineral nutrients needed to build important compounds like proteins and nucleic acids. Remember, though, that soil does not provide the carbon and energy that make up most of plants' biomass—that comes from $CO_2$ in the air and energy from sunlight, captured in photosynthesis (see Chapter 2). We now take a look at exactly what the mineral nutrients are, and why plants need them.

## 4.4 The Basic Building Blocks of Plants are C, H, and O from Air and Water, and Macronutrients and Micronutrients from the Soil

What is a plant made of? A large tree is clearly many orders of magnitude larger than the seed from which it grew. Where did all of that "stuff" in the tree come from? A man or woman is much larger than a child, largely as a result of his or her cumulative food intake (of course, much of the food one eats is used to provide energy, which is used up along the way). But plants clearly do not eat.

Most of the material in a plant is made up of compounds of carbon, hydrogen, and oxygen, just like all other living organisms (Table 4.2). The carbon that makes up most of the biomass of a plant is obtained by fixing atmospheric $CO_2$ during photosynthesis. The oxygen also comes from the atmosphere, while the hydrogen comes from water taken up by the roots. With two important exceptions—the nitrogen fixed by symbionts of some species or absorbed through leaves of "carnivorous" species—all the rest of the elements in a plant are taken up from the soil solution either by the plant's roots or by mycorrhizal symbionts (see Chapter 5).

**TABLE 4.2** Elements essential for plant growth and survival

| Element | Symbol | Principal form in which absorbed | Concentration in plant: Average mass (g/kg plant dry weight) | Concentration in plant: Percentage of total mass (range) | Important functions |
|---|---|---|---|---|---|
| Found in all organic compounds | | | | | |
| Carbon | C | $CO_2$ | 450.0 | ~44% | Major component of all organic compounds, including cellulose in cell walls, which makes up much of the plant's dry weight; backbone of other carbohydrates (sugars, starches) and lipids, which store and transport the energy captured in photosynthesis |
| Oxygen | O | $H_2O$ or $O_2$ | 450.0 | ~44% | Component of all major organic compounds; cellular respiration |
| Hydrogen | H | $H_2O$ | 60.0 | ~6% | Component of all major organic compounds; essential to all major biochemical reactions and to acid/base balance |
| Macronutrients | | | | | |
| Nitrogen | N | $NO_3^-$ or $NH_4^+$ | 15.0 | 1%–4% | Essential component of nucleotides, nucleic acids, amino acids, proteins (including structural proteins and enzymes), and chlorophylls |
| Potassium | K | $K^+$ | 10.0 | 0.5%–6% | Involved in osmosis, ion balance, pH regulation, and opening and closing of stomata; activator of many enzymes; protein synthesis |
| Calcium | Ca | $Ca^{2+}$ | 5.0 | 0.2%–3.5% | Strengthens cell walls and some plant tissues; involved in cell division and cell elongation, membrane permeability, cation-anion balance; structural component of many molecules; second messenger in signal conduction between environment and plant growth and developmental responses |
| Magnesium | Mg | $Mg^{2+}$ | 2.0 | 0.1%–0.8% | Essential component of the chlorophyll molecule; activator of many enzymes; involved in cation-anion balance and regulation of cytoplasm pH |
| Phosphorus | P | $H_2PO_4^-$ or $HPO_4^{2-}$ | 2.0 | 0.1%–0.8% | Essential structural component of nucleic acids, proteins, ATP, and $NADP^+$; critical for energy transfer and storage |
| Sulfur | S | $SO_4^{2-}$ | 1.0 | 0.05%–1% | Component of some amino acids and proteins, coenzymes (including coenzyme A), and secondary metabolic products (including defensive compounds) |

Water is continually moving through plants from the soil into the air when stomata are open and plants are photosynthesizing (see Chapter 2). While most of the water entering a leaf exits through the stomata, leaves and herbaceous plants still consist of up to 90% water. The water content of the woody parts of trees and shrubs (branches, stems, trunk, roots) is lower; living wood is about 35%–60% water. Most of the "dry weight" of plants—the part that is not water—is constructed of materials that ultimately come from the air (C and O atoms, see Table 4.2). The fraction of the dry weight of plants that comes from the soil (the mineral nutrients such as nitrogen and phosphorus) ranges from under 1% to only a few percent of its weight but is critical for plant functions because these elements play a central role in molecules like proteins, DNA, and cholorophyll.

**TABLE 4.2** (*continued*)

| Element | Symbol | Principal form in which absorbed | Concentration in plant: Average mass (g/kg plant dry weight) | Concentration in plant: Percentage of total mass (range) | Important functions |
|---|---|---|---|---|---|
| Micronutrients | | | | | |
| Chlorine | Cl | $Cl^-$ | 0.1 | 100–10,000 ppm | Used in stomatal regulation, proton pumps, and osmoregulation; critical for splitting water molecules in photosystem |
| Iron | Fe | $Fe^{3+}$, $Fe^{2+}$ | 0.1 | 25–300 ppm | Component of heme proteins (such as cytochromes, essential molecules in light reactions, cellular respiration, and nitrate reduction; leghemoglobin, used in N fixation) and iron-sulfur proteins; necessary for chlorophyll synthesis and in the electron transport chain in light reactions |
| Manganese | Mn | $Mn^{2+}$ | 0.05 | 15–800 ppm | Present in several enzymes; needed for activation of many enzymes |
| Boron | B | $B(OH)_3$, $B(OH)_4^-$ | 0.02 | 5–75 ppm | Poorly understood; cell wall synthesis, nucleic acid synthesis, and plasma membrane integrity |
| Zinc | Zn | $Zn^{2+}$ | 0.02 | 15–100 ppm | Present in several enzymes; needed for activation of many enzymes; protein synthesis and carbohydrate metabolism; structural component of ribosomes; important but not well understood role in plant hormone metabolism |
| Copper | Cu | $Cu^{2+}$ | 0.006 | 4–30 ppm | Present in some proteins and in plastocyanin (necessary for light reactions); enzyme activation; pollen formation and ovule fertilization; lignification of secondary cell walls (in wood formation) |
| Nickel | Ni | $Ni^{2+}$ | ~0.0001 | ~0.1 ppm | Necessary for function of some enzymes; nitrogen metabolism |
| Molybdenum | Mo | $MoO_4^{2-}$ | 0.0001 | 0.1–5 ppm | Enzyme cofactor for N fixation and some other reactions; pollen formation and seed dormancy |
| Essential to some plants | | | | | |
| Sodium | Na | $Na^-$ | Trace | | Osmotic and ionic balance, especially in some desert and salt marsh species |
| Cobalt | Co | $Co^{2+}$ | Trace | | Required by N-fixing microorganisms |
| Silicon | Si | $Si(OH)_4$ | Variable | | Not well understood; appears to have a role in disease resistance; important structural component of cells in grasses, *Equisetum* spp. (horsetails), and other plants; may reduce leaf water loss |

*Sources*: Compiled in part from data in E. C. Miller. 1938. *Plant Physiology*. McGraw-Hill: New York, NY; S. Epstein. 1965. In *Plant Biochemistry*, J. Bonner and J. E. Varner (Eds.), pp. 438–466. Academic Press: London; P. R. Stout. 1961. *Proc 9th Annu Calif Fertilizer Conf*, pp. 21–23; P. H. Raven et al. 1981. *Biology of Plants*, 3rd ed. Worth Publishers: New York, NY; P. H. Brown et al. 1987. *J Plant Nutr* 10: 2125–2135.

How do plants use the elements that they take up from the atmosphere and the soil? There are two major roles that an element can play in an organism: it can be part of the material that makes up the organism's structure, or it can be essential for the operation of the organism's metabolism. The chemical elements plants need to live, grow, and reproduce (except for the major components, C, O, and H) are called the **essential mineral nutrients**. Table 4.2 lists some of the main functions of each of these elements. The essential nutrients needed in larger quantities are called **macronutrients**; those needed in very small quantities are

known as **micronutrients**. Another group of minerals is the **beneficial mineral elements**; these minerals either are essential only for certain plant species, or are not essential but do stimulate growth. The beneficial mineral elements include sodium, silicon, and cobalt.

While plants cannot grow and function without the essential mineral nutrients, more is not necessarily better. Most of these minerals are necessary in small to moderate amounts, but toxic in large quantities. Some species have evolved tolerance of high concentrations of minerals that are toxic to most other species (see Figure 9.6).

## The stoichiometry of elements in plants and soils regulates many ecological processes

The two most important essential mineral nutrients for plants are nitrogen and phosphorus. These nutrients are found in a number of common forms and in different proportions in the soil solution in different environments. Variation in the ecological **stoichiometry**—the relative proportions—of different elements in the soil can have important consequences for the stoichiometry of nutrients in plants, thereby affecting their growth and function (Sterner and Elser 2002) (Figure 4.17). For example, the ratio of carbon to phosphous is important in determining the maximum growth rate of an individual plant (Thompson et al. 1997; Ågren 2004). Individual plants can compensate for imbalances in available resources by changing their allocations to different tissues and different physiological functions (Chapin 1980; Bazzaz 1997). For example, when nitrogen is limited relative to carbon—the ratio of nitrogen to carbon is low—a plant might produce more roots relative to aboveground tissues, resulting in an increase in nitrogen uptake and a decrease in photosynthetic carbon gain. The end result is that the ratio of nitrogen to carbon is increased, thus restoring the stoichiometric balance. Soil nutrients, water, and light interact in their effects on plant function, and a plant responds to these multiple limiting factors in a coordinated fashion (Chapin et al. 1987; Aerts and Chapin 1999). Plant stoichiometry changes with age and over a plant's life history, as it grows from a seedling to maturity and becomes reproductive.

Moving up in conceptual and spatial scale, the stoichiometry of the soil environment is an important factor in ecosystem processes (see Chapter 5), with major consequences for species composition and species interactions in different habitats. Nutrient stoichiometry, particularly the ratio of carbon to nitrogen, is a critical component of the nutritional value of plant tissue to herbivores (see Chapter 11) and may also affect plant competitive interactions and interactions with various microorganisms, including pathogens. Plant stoichiometry varies seasonally, among species, among habitats, and at larger scales across latitudinal and altitudinal gradients (Körner et al. 1989; Méndez and Karlsson 2005).

## Nitrogen is often the limiting nutrient for plant growth

Plants need large amounts of nitrogen to carry out the major functions of life, including photosynthesizing (see Chapter 2), growing, and reproducing (see Chapter 6), and the amount of nitrogen available in the soil frequently

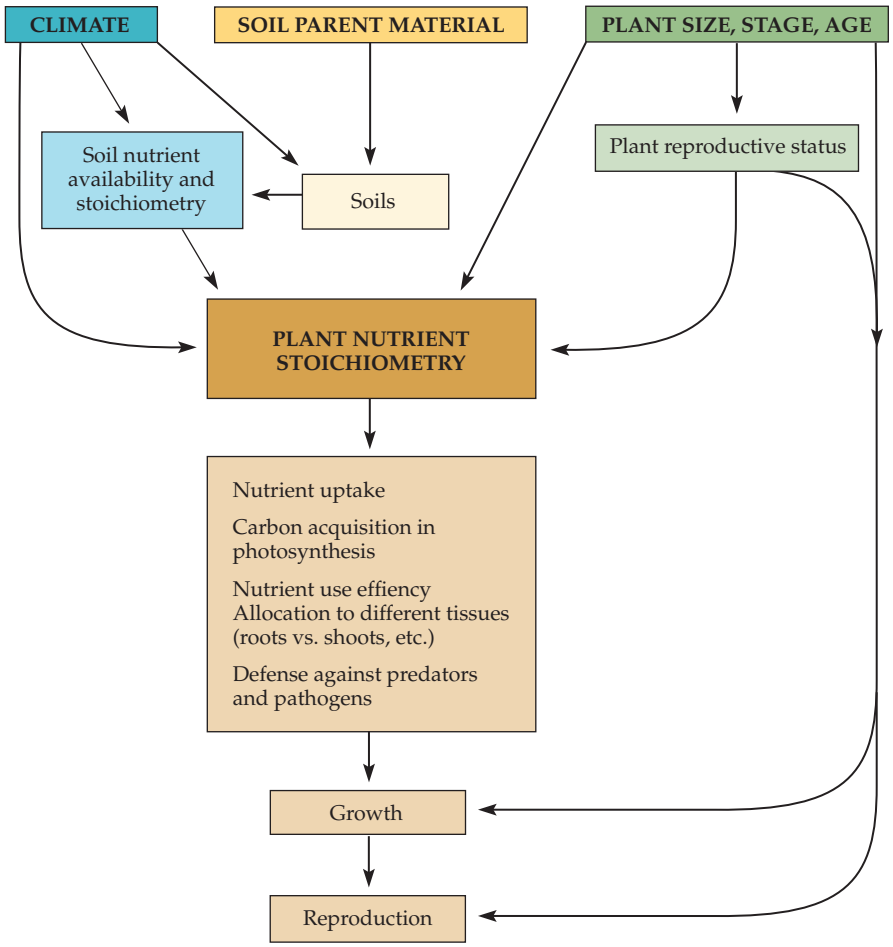



**Figure 4.17** A conceptual model of the relationships between factors influencing plant nutrient stoichiometry and the effects of plant nutrient stoichiometry on other plant functions. (After M. Méndez and P. S. Karlsson. 2005. *Ecology* 86: 982–991.)

limits the ability of the plant to carry out these functions. Nitrogen, which is essential for the manufacture of proteins and other major components in plant cells, is a major limitation to plant growth and function in most soils. The Green Revolution, which greatly increased human food supplies in the mid-twentieth century, was in part a result of the discovery of artificial **nitrogen fixation** by the Haber process (see Chapter 5, especially Box 5C) and its use in synthetic commercial fertilizers, along with the breeding and selection of crop plants that could respond with strong gains in productivity to nitrogen fertilizer.

Plant form and development are also influenced by nitrogen availability. Increased nitrogen availability may delay the senescence of leaves or of the entire plant, increase the proportion of biomass allocated to shoots relative to roots, and increase the proportion allocated to vegetative growth relative to sexual reproduction. Nitrogen availability may affect the plant's biochemical composition in a complex manner and may alter the concentrations and types of defensive compounds present (see Figure 11.13). The ratio of carbon to nitrogen in plant tissues is affected by nitrogen supply rates that in turn affect many other aspects of plant life, including susceptibility to herbivores and litter decomposition rates.

Nitrogen exists in the soil in inorganic and organic forms. The two major inorganic forms that plants take up from the soil are nitrate ($NO_3^-$) and ammonium ($NH_4^+$), but nitrate is generally more abundant and more readily taken up. Some plants are able to take up ammonium, particularly when they are young, and plants adapted to acidic soils or to soils with limited oxygen potential (such as waterlogged soils) preferentially take up ammonium (see Chapter 5). The costs and benefits of nitrate versus ammonium uptake are complex and depend on their concentrations in the soil as well as other factors. Nitrate must be reduced to ammonia before it can be incorporated into organic compounds in the plant, but nitrate is highly mobile and can be transported readily in dissolved form in the xylem and stored in cell vacuoles. Ammonium is largely incorporated into organic compounds in the roots directly after it is taken up. Ammonium can be toxic to plants, particularly when converted to ammonia ($NH_3$). A small number of plant species are able to take up organic nitrogen in the form of amino acids directly from the soil. Otherwise, plants can use these amino acids in the soil only if they are first converted by bacteria to $NH_4^+$.

### In some plants nitrogen comes from fixation by symbiotes

Nitrogen is needed in large quantities by plants, and its availability in the soil is almost always low. Ironically, plants are surrounded by an ocean of nitrogen in the atmosphere in the form of elemental $N_2$, which they cannot tap directly. You might wonder why plants cannot use this form of nitrogen. It turns out that some plants are able to do so—not directly, but by forming symbioses (Box 4C) with nitrogen-fixing prokaryotes.

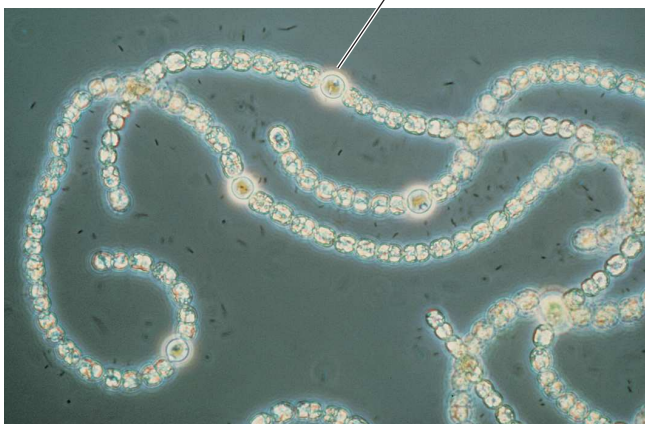
Only bacteria can directly use $N_2$ and "fix" it into biologically usable forms. While the nitrogen comes from the atmosphere, the bacteria obtain it mainly from air in soil pores underground. These nitrogen-fixing organisms are remarkably diverse and encompass 19 families, including 8 families of cyanobacteria (the blue-green bacteria, which are photosynthetic). Some nitrogen fixers are free-living,

## BOX 4C Symbioses and Mutualisms

The word symbiosis is rooted in two Greek words, bios ("life") and sym ("together"). Symbioses are associations between members of two different species that live in intimate contact with each other. There are many different kinds of symbiotic associations among organisms.

A symbiotic association may benefit members of both species involved, or it may benefit only one member, with no effect or a negative effect on the other. If both members benefit, the symbiosis is a **mutualism**. If one member benefits while the other is unaffected, the relationship is called a **commensalism**. And if one member benefits while the other is harmed (sometimes even destroyed), it is **parasitism**. Any of these three types of symbioses may be either **obligate**—necessary for the survival of one or both members—or **facultative**—existing only under some conditions. The nature of a particular symbiosis may even change over time or as environmental conditions change.

Two important symbioses between plants and other organisms are nitrogen fixation by certain species of bacteria living in or on plant roots, and the formation of mycorrhizae by plants and certain species of fungi (see Chapter 11). Both of these symbioses are often, but not always, mutualistic. Other common symbioses include interactions between some plants and animals that act as pollinators or seed dispersers (see Chapter 7). The nature of the interactions in these different symbioses differ greatly; because they have little to do with one another, we consider them in the context of their particular effects on plants (e.g., plant-pollinator associations affect plant reproduction). However, the benefits conferred by all of these associations can vary greatly and can range from absolute dependence on the association on the part of both members, to more opportunistic benefits to one or the other species, to a situation in which one benefits and the other is "taken advantage of" by the association.




© J. R. Waaland/Biological Photo Service

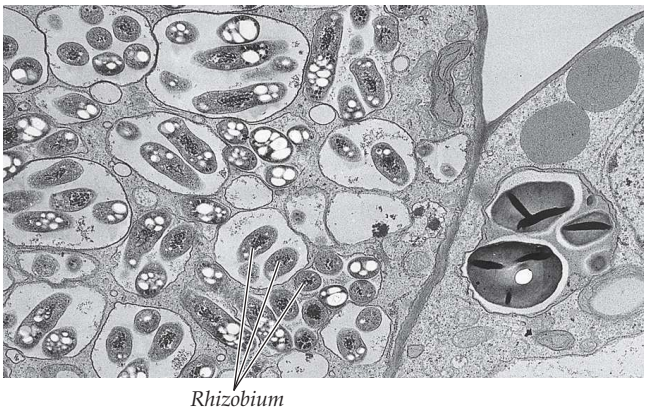
**Figure 4.18** Free-living nitrogen-fixing cyanobacteria such as *Anabaena azollae* fix nitrogen in specialized cells called heterocysts.

some are symbiotic with terrestrial plants, and some are associated with the root surfaces and intercellular spaces of plants. The biological reduction of $N_2$ to $NH_3$ is catalyzed by an enzyme complex, called nitrogenase, that is found in all $N_2$-fixing prokaryotic microorganisms and nowhere else in the living world. The reaction requires a great deal of energy, and the source of that energy is the key to understanding the different ecological roles played by different nitrogen fixers.

Free-living nitrogen-fixing bacteria are found in soil surfaces, in aquatic environments, and in anoxic environments such as mud. They are either photosynthetic cyanobacteria, such as *Anabaena* (Figure 4.18), or are heterotrophic and obtain their energy from the decomposition of plant residues in the soil. The volume of nitrogen fixation by the heterotrophs is generally low because they are limited by the amount of carbon and energy they can extract from dead material. In contrast, the photosynthetic free-living nitrogen fixers can provide substantial amounts of nitrogen to the systems in which they live because they use the far more abundant energy of sunlight. An ecologically particularly important group of free-living nitrogen fixers are the photosynthetic cyanobacteria. They are abundant as free-living organisms in the oceans, in freshwater, on the surface of snow, and on soil surfaces (including soil crusts).

Symbiotic nitrogen fixation is more important in most ecosystems than fixation by free-living organisms. The types of symbiotic associations between nitrogen-fixing bacteria and terrestrial plants are many and various. There is generally a strong specificity in most of these symbioses, in which particular bacterial species preferentially infect particular plant species. One of the most important ecological and agricultural nitrogen-fixing symbioses is the one among bacteria in the genera *Rhizobium* (that grow quickly) and *Bradyrhizobium* (that grow slowly) and plants in the family Fabaceae (legumes such as peas, beans, clover, and acacia trees; Figure 4.19A). The bacteria live in nodules on the legumes' roots (Figure 4.19B), providing the plant with a source of $NH_3$ and receiving carbohydrates (which provide carbon and energy) from the host plant. The nodules provide the anaerobic conditions that the

(A)

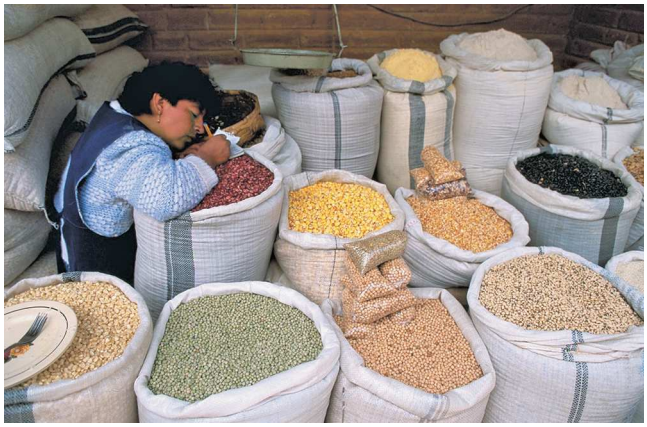



© E. Newcomb and S. Tandon/Biological Photo Service

(B)

© K. Wagner/Visuals Unlimited

**Figure 4.19** Symbiotic nitrogen-fixing bacteria of the genus *Rhizobium* cannot fix nitrogen until they form an association with a plant's root cells. (A) The cell on the right has not been "infected" with *Rhizobium*. Bacteria have entered the cell on the left, where they take the nitrogen-fixing form called bacteroid. Root cells with bacteroids form nodules. (B) Root nodules on *Glycine max* (soybean, Fabaceae). The anaerobic conditions inside the nodules provide the necessary environment for *Rhizobium* to fix nitrogen.



© Royalty-free/Corbis

**Figure 4.20** A wide variety of dried beans are for sale in this market stall in Ecuador. Legumes such as beans, members of the family Fabaceae, are crucial sources of protein for people throughout the world as well as for many other animal species.

bacteria require to fix nitrogen. The Fabaceae is one of the largest and most ecologically diverse plant families, and many of its members form symbioses with nitrogen-fixing bacteria. Many important crop species are members of this family, and its crops are major protein sources for civilizations worldwide (e.g., soybeans in eastern Asia, lentils in southern Asia, pigeon peas and peanuts in Africa, pintos and other beans in the Americas, and fava beans and garbanzos in the Middle East; Figure 4.20). Many tropical trees and shrubs (such as *Acacia* spp.) belong to the Fabaceae as well; these plants can play a critical role in ecosystem function by providing nitrogen to soils and high-protein food (particularly seeds and fruits) to wildlife.

Another common nitrogen-fixing symbiosis is that between *Frankia* (an actinomycete) and plants in the genera *Alnus* (alders) and *Ceanothus* and a few other tree and shrub species. In these symbioses, the bacteria live in nodules on the plants' roots and obtain carbon and energy from the plants while providing them with $NH_3$. There are about 200 plant species in the tropics and subtropics that form such associations, but they also occur in some temperate systems and are common early colonizing trees in many riparian communities; the nitrogen that is fixed by this symbiotic association can be very important in these successional ecosystems (see Chapter 12).

Cyanobacteria are also important partners in various nitrogen-fixing symbioses, in addition to occurring as free-living organisms. One example of such a symbiosis is that between *Azolla*, a tiny aquatic fern, and *Anabaena* (or other cyanobacteria) in rice paddies and other freshwater tropical systems (Figure 4.21). This *Azolla*-cyanobacterium symbiosis can be a major source of nitrogen in the ecosystems where it occurs, including traditional rice production (where it is deliberately maintained by farmers). In these types of plant-bacteria associations, there may be high specificity between the host plant and the bacterial species, but the bacteria remain external to the root cells. The bacteria obtain energy from root exudates from the plant, but the cyanobacteria provide little nitrogen to the plant until they die. Nitrogen fixation also occurs in the non-plant symbiosis in lichens between a fungal species and cyanobacteria. Lichens are symbiotic associations between a photosynthetic species and a fungal species; lichens associated with nitrogen-fixing cyanobacteria can fix nitrogen, while those associated with eukaryotic Chlorophyta (green algae) cannot.

External energy is required for nitrogen fixation because elemental nitrogen ($N_2$) has less energy than the forms usable by plants. The symbiotic bacteria obtain this energy from photosynthate supplied by the plant host, which is assumed to be diverted from other

Courtesy of J. Gurevitch

**Figure 4.21** Traditional terraced rice paddy in Madagascar. Rice is central to cultural life and is an essential part of the diet in Madagascar. An important component of the nitrogen economy of paddy agriculture is the symbiosis between the freshwater fern *Azolla* and the nitrogen-fixing cyanobacterium *Anabaena* (see Figure 4.18).

plant functions such as growth or reproduction. Factors that increase photosynthetic rates tend to increase nitrogen fixation rates, and factors (such as light and water limitation) that reduce photosynthetic rates tend to decrease nitrogen fixation.

The presence of large amounts of mineral nitrogen in the soil inhibits nitrogen fixation by suppressing nitrogenase activity and severely reducing the number of nodules formed. As nitrogen becomes more readily available in the soil, however, the uptake of soil nitrogen becomes energetically cheaper than providing carbohydrates to bacterial symbionts. From the perspective of plant fitness, the switch from using symbiotically fixed nitrogen to direct uptake from the soil should occur at this point.

### Phosphorus is limiting for plant growth in many environments

Like nitrogen, phosphorus is needed by plants in larger quantities than are usually available in soils. Next to nitrogen, phosphorus is likely to be the most limiting mineral nutrient for plant growth. Phosphorus is also needed to support symbiotic nitrogen fixation.

Soils vary greatly in the amount of phosphorus they contain. In the United States, total soil phosphorus tends to be greatest in the Pacific Northwest, extremely low in the Southeast, and low to moderate in the Northeast, Midwest, and Southwest. Soils derived from high-phosphate limestones are generally highest in phosphorus, and those derived from sandstones and acidic igneous rocks lowest. Calcareous soils in arid regions are usually high in phosphorus because substantial leaching has not occurred. Very old soils, particularly in the tropics and subtropics, such as those in Australia and parts of Africa, tend to be exceptionally low in phosphorus due to steady, slow losses by leaching over an extended period of time.

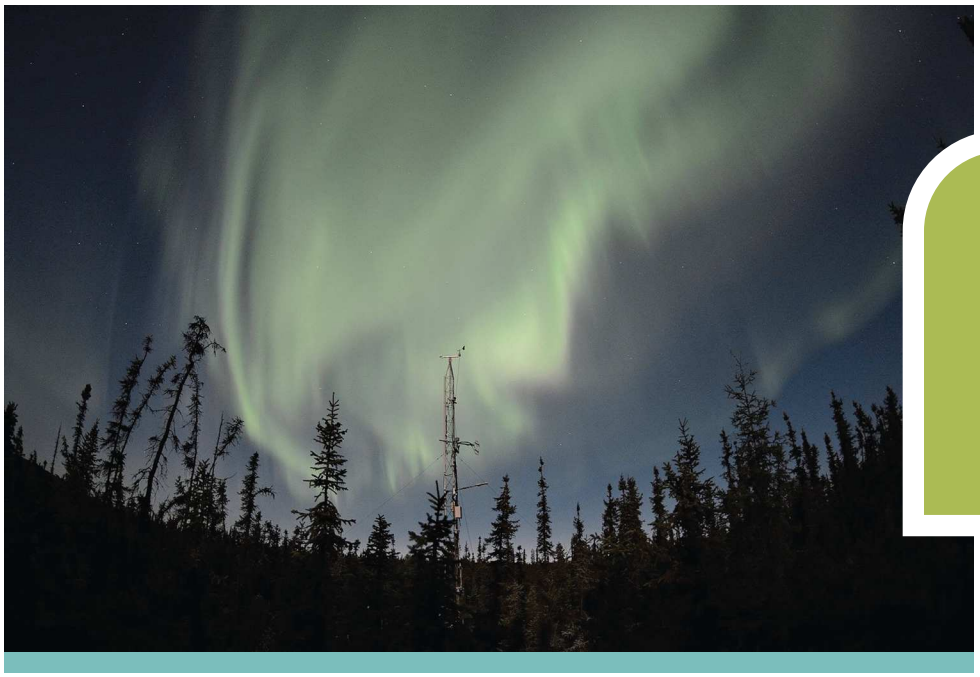
This may seem surprising, but the amount of phosphorus available to plants is not directly related to the total amount of phosphorus present in the soil. Unlike nitrogen, most phosphorus in the soil is bound in ways that make it completely inaccessible to plants. Soil phosphorus chemistry is exceedingly complex and depends on pH and many other factors. Phosphorus in the soil exists in various organic and inorganic forms, and both organic and inorganic phosphates can be tightly bound in ways that make them unavailable to plants. Phosphorus can be tightly bound to clay particles, or complexed with calcium, iron, aluminum, and silicates, or bound up in organic matter. Thus, phosphorus leaches much less readily and more slowly than nitrogen. Microbial activity removes available forms of phosphorus from the soil solution and locks them up in forms unavailable to plants (that is, it "immobilizes" the phosphorus), but it also can transform it back into available forms and release them back into the soil solution, where they can be taken up by plants once again (see Figure 5.23).

Phosphorus in soil is derived from apatite minerals in the parent material. In natural systems, most of the phosphorus in plants is recycled by microbial decomposition. While some systems have considerable reserves of phosphorus in the soil, in others, such as many grasslands and some tropical rainforests, essentially all of the available phosphorus exists in living plant tissue, litter, and decomposing organic matter, from which it is rapidly recycled or lost. Most of the phosphorus taken up by plants comes from the rhizosphere.

It is sometimes possible for plants to directly influence the availability of nutrients in the rhizosphere. Some plants, for example, can create their own locally acidic environment by secreting organic acids that promote the uptake of cations. Plants of the family Proteaceae—a prominent group on the old Gondwanan continents (Australia, South Africa, and South America; see Chapter 17)—which often grow in very phosphorus-poor soils, may form specialized **proteoid roots**. These dense clusters of short, fine lateral rootlets produce organic acids or other chelating agents, which act to release phosphorus from insoluble complexes with calcium, iron, aluminum. This ability allows these plants to acquire phosphorus (and sometimes other mineral nutrients) in soils where other plants are not able to remove it from the soil. Some unrelated plants—such as a *Lupinus* sp. (lupine, Fabaceae)—have also been observed to form proteoid roots.

## Summary

- Soil is a complex product of the interaction between living organisms and their terrestrial substrate.
- There is a tremendous diversity of soil properties and characteristics across different parts of the world, and even within local areas. These differences are an important factor in determining the kind of vegetation that grows in a site as well as the growth of individual plants.
- Soil texture describes the relative proportions of clay, silt, and sand mineral particles in the soil. These particles differ in size, shape, charge, and composition, and they impart different characteristics to soils. Soil organic matter is a critical component in determining soil structure.
- Plants depend on soil to obtain water. Different soils hold different amounts of water in the larger and smaller pores of the soil structure.
- The rhizosphere is the area immediately surrounding plant roots and affected by the complex interactions between the root and microorganisms living there.
- Soil conservation is a critical environmental issue globally.
- Plants depend on soil for the mineral nutrients that are essential to their survival and growth. Nitrogen and phosphorus are generally the most limiting mineral nutrients for plants.
- Symbiotic nitrogen fixation is an important source of nitrogen in some plants.



Courtesy of Masahito Ueyama/Ameriflux

5

# Ecosystem Processes

Ecosystems are all of the living organisms, the remains or discarded parts of organisms, and the nonliving environment with which organisms interact and which surrounds and pervades them. Sunlight energy is captured by living organisms—autotrophs—and stored in the chemical bonds of molecules in living and once-living organisms and is transferred among organisms and between living and nonliving parts of the ecosystem. The process of respiration releases the energy from the chemical bonds to drive the processes of life. Materials, including carbon, nitrogen, phosphorus, and water among others, are also stored and transferred among living and nonliving parts of ecosystems. Quantifying these pathways, the size and nature of the pools of materials, the rate at which fluxes of materials and energy are transferred, and the factors controlling these pools and fluxes are the focus of ecosystem ecology. While a single ecosystem is generally of limited size that typically corresponds to a local community (see Chapter 12), the processes of material and energy flow can extend from that local community to the entire globe. How those processes link across scales has taken on increased importance as we cope with global change.

Above: Ameriflux eddy covariance flux tower in Fairbanks, Alaska, U.S.A. under the Northern Lights.

## 5.1 Ecosystem Processes Set the Stage for Life in a Salt Marsh

Imagine that you are standing on the edge of a great coastal marsh, such as those that occur—or once occurred—along much of the east coast of North America. You might find similar coastal wetlands—salt marshes—along the northwest coast of North America, the southern part of South America, much of northern and western Europe, the southernmost and northernmost parts of Africa, parts of coastal east Asia, and Australia (Figure 5.1). These habitats include **estuaries**, highly productive and ecologically important habitats formed where rivers and streams flow into the ocean, creating a mixture of salty and fresh (nonsalty) water. In tropical regions, mangrove swamps are common along coastlines and estuaries. These habitats, although globally distributed, do not dominate large areas because they are limited to coastlines. They are highly threatened by urban development and rising sea levels in many parts of the world (see Chapter 19).

You are at the boundary between the marine and terrestrial worlds. In front of you are seemingly endless extents of waving grasses and other plants uniquely adapted to the conditions of the salt marsh: regular inundation by salt water, which is toxic to most terrestrial plants, scouring by waves, and in some places scouring by ice. Grasses like *Spartina patens* (saltmeadow cordgrass, Poaceae) and *Spartina alterniflora* (smooth cordgrass, Poaceae) (Figure 5.2), and salt-tolerant *Suaeda maritima* (herbaceous seepweed, Amaranthaceae) and *Salicornia maritima* (slender glasswort, Amaranthaceae), are abundant in the salt marshes of the east coast of North America. Here, the marsh is broad, low, and flat and rises gradually away from the ocean to upland terrestrial habitats. Fresh water washes through channels from rivers or streams and enters the marsh from rain, and mineral nutrients are often abundant. Because the soil pores are often flooded with water, the soil is likely to be anaerobic, or low in oxygen. Few or no trees are present, at least close to the shore, so there is little shade, and temperatures may range from very hot to very cold.

Yet, there is an abundance of life in this environment. You might startle large flocks of birds nesting in the low-growing vegetation or notice others searching for food along the shore. Depending on where you are, you might see red-winged blackbirds, osprey, blue herons, and different species of terns, as well as shore birds such as godwits, sandpipers, and the amazing little red knot, which can migrate from the southernmost shores of South America to northern Alaska. Hermit crabs and fiddler crabs scuttle along the sand, and clams, mussels, snails, insect larvae, amphipods, and isopods are hidden below the mud, often with little holes for air giving them away. Turtles of various species may be present, either as hatchlings emerging from the sand, or as terrestrial species in the higher parts of the marsh. Salt marshes have some of the highest rates of **productivity**—the capture of sunlight energy and its storage in carbon bonds—of any of the world's ecosystems. What ecosystem processes

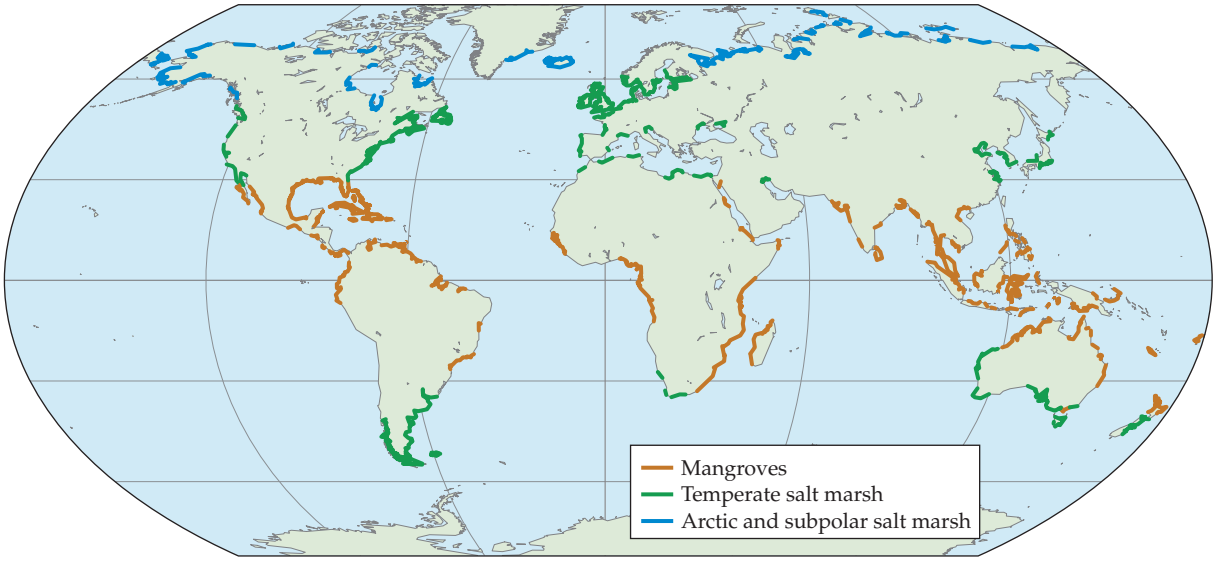
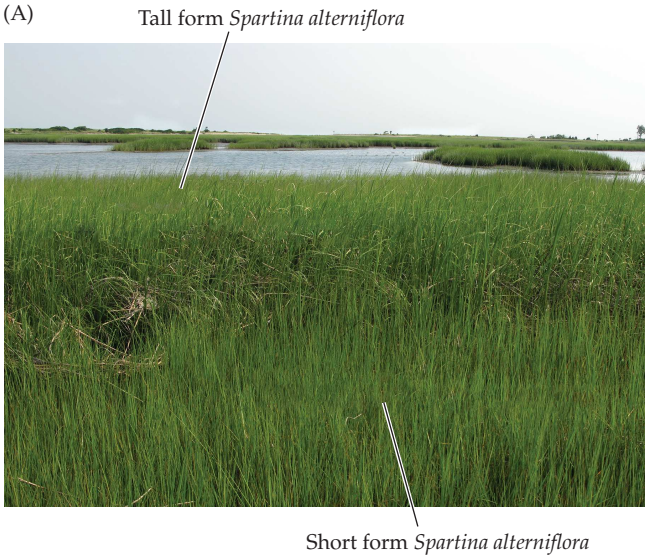



**Figure 5.1** Global distribution of temperate salt marshes, mangrove forest, and arctic salt marshes. Salt marshes and mangrove forests, the intertidal wetlands of the world's coastlines, provide key ecosystem services across the globe. (After D. B. Scott et al. 2014. *Coastal Wetlands of the World: Geology, Ecology, Distribution, and Applications*. Cambridge University Press: Cambridge.)

(A)




Both photos courtesy of Sandy Richard

(B)

**Figure 5.2** (A) A salt marsh on the north shore of Long Island, New York, showing the tall form of *Spartina alterniflora* (smooth cordgrass, Poaceae), which is more common in the low marsh areas where it receives tidal flooding daily, and the short form of the same species in the high marsh areas of the salt marsh. (B) High marsh in the same area showing a salt panne, a depression containing water within a salt marsh where salt accumulates over time.

are occurring in the marsh, and what controls them? What benefits does the salt marsh provide to people, and what role does it play through its ecosystem processes in the natural world? We will return to this example of an ecosystem later in this chapter as we examine general concepts and information about ecosystem processes.

The word *ecosystem* was originally defined by the plant ecologist Sir Arthur Tansley (1935). He used the word to encourage ecologists to think about the entire system of living organisms in the context of the physical factors on which they depend and with which they are interconnected. An **ecosystem** is defined as all of the organisms in an area and all of the abiotic materials and energy with which they interact. Together, all of the Earth's living organisms and their physical environments make up the **biosphere**. The **ecosystem concept** states that the biological and physical components of the environment form a single interactive system. Tansley also noticed that ecosystems exist over a range of spatial extents, which is a very contemporary concept. Tansley wrote, "the [eco] systems we isolate mentally are not only included as parts of larger ones, but they also overlap, interlock and interact with one another." Ecosystem ecologists study the roles of plants as conduits for and transformers of energy and materials in ecosystems, as well as the effects of the supplies and flows of energy and materials on plants.

What determines plant productivity, and why is it so different in different places on Earth? Understanding the things that control the flow of energy and materials in ecosystems is a critical aspect of ecology. Ecosystem research has revealed the most important chemical elements affecting plant growth and the processes of decomposition in ecosystems, but we may want to learn more about how the relative amounts of these nutrients affect these processes in different systems. In many cases, much remains controversial about how ecosystem processes interact with communities' characteristics, including species diversity and plant-animal and plant-microorganism interactions. Inversely, how are the ecosystem processes affected by community characteristics?

We begin by considering one aspect of how a young plant grows—where the materials in its body come from—widening our focus to the ecosystem from there. A seed, with its embryo and other associated tissues, contains very limited quantities of energy and materials. Seeds are generally very small—orders of magnitude smaller than the plants into which they will grow. The spore that will develop into a mature fern is even smaller than most seeds. Where does the material in a mature plant come from—what produces the wood in the trunk, branches, and roots of a huge forest tree? As the newly germinated seedling grows into a mature plant, it transforms inorganic forms of carbon with other materials into complex organic molecules, and it stores the energy captured from photons of light in the form of chemical bonds in these molecules (see Chapter 2). Most of the material in the towering forest tree consists of carbon compounds, initially captured in the process of photosynthesis from carbon dioxide in the air. Nitrogen, phosphorus, and other essential elements taken up from soil contribute to fundamentally important molecules such as proteins and nucleic acids (see Chapter 4). The oxygen (O) available to aboveground parts of plants is

almost never limiting to terrestrial plant growth; but as we will see, oxygen can be limiting below ground, particularly in waterlogged soils. Carbon (C), in the form of carbon dioxide ($CO_2$), is generally available everywhere on the Earth, although its availability depends on altitude; as altitude increases, carbon becomes less available to plants as air pressure and the concentration of $CO_2$ decline. Yet, even at sea level, the production of plant biomass varies enormously over the Earth's surface (see Table 18.2). What are the causes of this tremendous variation? Climate, ecosystem age, soil nutrients, and disturbance are among the factors affecting productivity. In this chapter we consider some of the major factors controlling the ability of plants to obtain essential nutrients, as well as the roles of plants as conduits for the fluxes of those nutrients, as agents controlling those fluxes, and as reservoirs for the storage of those nutrients.

Not only do plants produce the materials in their own bodies, but also, in terrestrial systems, the energy and material used by **heterotrophs**—animals, fungi, bacteria, and protists that consume other organisms for food—is transformed by plants from inorganic building blocks that animals cannot use into useful organic compounds. Photosynthetic bacteria, diatoms, and unicellular and multicellular marine photosynthetic organisms belonging to various phyla (divisions) play a similar role in aquatic systems and sometimes in soils. The story is more complicated than that, however, because some of the materials needed by terrestrial plants must first be transformed by microorganisms into forms the plants can use (see Chapter 4).

## 5.2 Ecosystem Pools and Fluxes Form Cycles of Nutrients and Energy

A major goal of ecosystem ecology is to understand what regulates the **pools** (quantities stored) and **fluxes** (flows) of materials and energy in the various abiotic and biotic components of ecosystems (Box 5A). The major fluxes and pools of a material in a system are collectively called a **cycle**; for example, we might want to quantify the nitrogen cycle or phosphorus cycle of a wetland ecosystem such as the saltmarsh described above. At larger spatial scales, we can also quantify regional or global ecosystem cycles (see Chapter 16). Most of the nutrients that are important to plants move through ecosystems in **biogeochemical cycles**—both biological and chemical reactions are involved. Geology, too, plays a role. Energy transformations are closely tied to the biogeochemical cycling of materials.

The cycles of water and different elements have different time scales, varying over orders of magnitude in turnover times and retention times. Organisms, however, have a fairly fixed ecological stoichiometry (the ratio of nutrients in their tissues; see Chapter 4), even though supply rates differ greatly among different nutrients. Differences in stoichiometry among plants, microorganisms, herbivores, and the physical environment interact with nutrient turnover times to regulate the dynamics of nutrient cycles. Living organisms both regulate and are regulated by the fluxes of particular materials through the biosphere (Redfield 1958).

### BOX 5A Biogeochemical Cycles: Quantifying Pools and Fluxes

Pools and fluxes are related to one another by **turnover time**. Turnover time, or **mean residence time**, is a measure of how rapidly materials move through a system. It can be used to describe systems at any scale—from an individual plant to an ecosystem to the atmosphere. Turnover time is equal to the total mass of a component divided by its flux in or out of the system. The fraction lost each year, $-k$, is equal to 1/turnover time in years, and the amount remaining after $t$ years as a fraction of the starting amount is $e^{-kt}$ (assuming that the flux rate is constant). The turnover time for different materials differs greatly. Some materials, like sugars, decay very rapidly, while others, like lignin, are retained in the system for a very long time. There are also large differences globally in the turnover times for the same materials as they pass through different components of the biosphere—the living organisms and their physical environment on Earth. The **retention time** for nutrients within a component of an ecosystem or in the biosphere is inversely related to turnover time. Retention time is the average length of time a material remains in an ecosystem component. Retention times provide insight into how nutrients move through components of the biosphere, and the dynamics by which they affect living organisms.

A basic approach that is often used in accounting for the magnitude of pools and fluxes is the **mass balance** approach to constructing nutrient budgets. The mass balance approach states that inputs – outputs = change in storage. In other words, if we can measure everything that goes into a system and everything that comes out, the difference between the two quantities must be reflected by a change in the material stored in the system. The mass balance approach allows ecosystem ecologists to account for difficult-to-measure quantities by subtraction from accurate measurements of the other terms (Vitousek and Reiners 1975).

Plants have a major role in the local, regional, and global cycles of water and some chemical elements. The large part that plants—and humans—play in global biogeochemistry also has major implications for global change, as we will see in Chapter 16. Nitrogen (N) and phosphorus (P) are essential nutrients needed by plants in relatively large quantities. Both are major constituents of essential organic molecules (see Table 4.2), but their supply is limited in most soils. Consequently, the availability of N and P controls primary productivity in many ecosystems. In contrast, other nutrients such as sulfur (S) and potassium (K) are available in larger quantities, and therefore primary productivity can determine the rate at which they cycle in the ecosystem (Schlesinger 1997). In both cases, living organisms have a major effect on the geochemistry—the pools and fluxes—of these major components of living things. In contrast, the cycling of elements that are not major constituents of living things, such as sodium (Na) or aluminum (Al), is relatively independent of the actions of living organisms (Schlesinger 1997).

Cycles of energy and materials are often illustrated for convenience by diagrams that show the fluxes and pools of a single element. It is important to recognize that the cycles of various materials can be highly interconnected and that they interact to affect plant growth. The carbon cycle, for example, both depends on and strongly affects the nitrogen and phosphorus cycles (Figure 5.3). These cycles are interconnected by their mutual effects on plant growth, tissue composition, leaf longevity, litter production, and litter decomposition rates. These interconnections determine the flux and pool sizes of C, N, and P. All three cycles are also simultaneously affected by climatic and local environmental factors such as temperature, precipitation, and light availability.

There are large differences in the magnitudes of pools and fluxes among different materials and ecosystems. Fluxes of water and carbon, for example, are vastly greater than fluxes of phosphorus. Pools of many nutrients are much larger in living plants in tropical forests than they are in tundra plants, although the pools of carbon and other nutrients may be far greater in many tundra soils than in tropical soils. There are generally huge pools of stored carbon in wetland soils, substantial (but not as much) stored carbon in grassland soils, and very little stored carbon in desert soils.

The ultimate sources of the different elements that make up living tissues also differ. The atmosphere is the ultimate source of N, O, and C, although plants obtain N from the air only indirectly. The atmosphere contains about 78% N and about 21% O. The concentration of $CO_2$ in the atmosphere, while rising (see Figure 16.23 and Figure 16.24), is surprisingly low—about 4/100 of 1% on average, or about 407 parts per million by volume (ppmv) in 2018. Nevertheless, because the volume of the atmosphere is so large, the total pool of carbon in the atmosphere

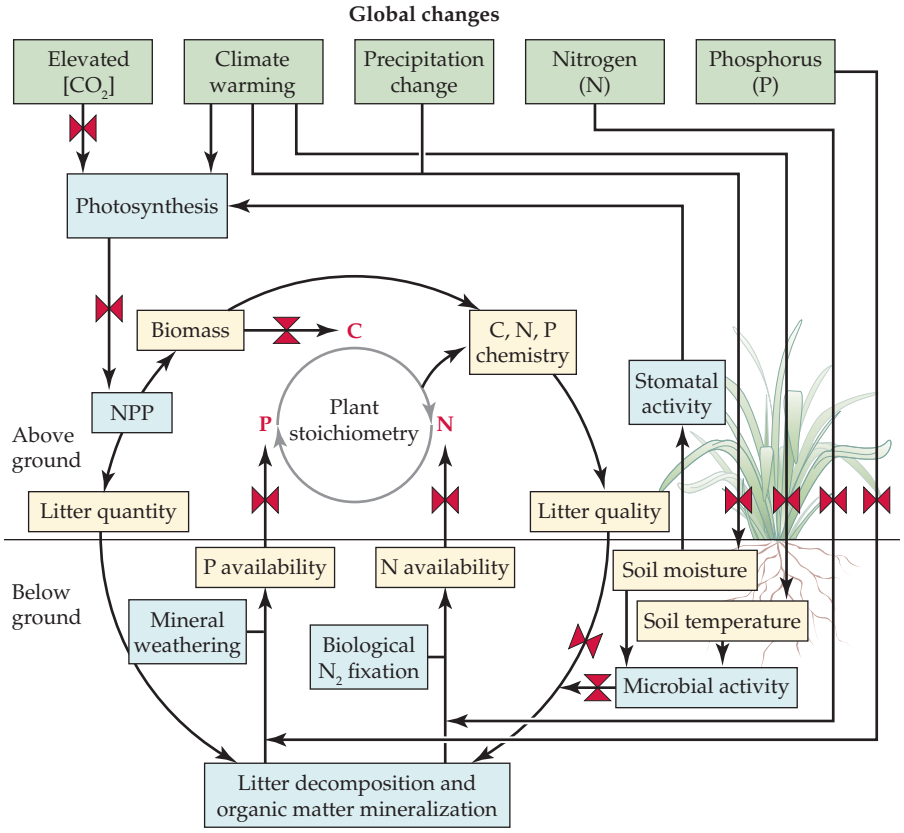



**Figure 5.3** Global changes in atmospheric $CO_2$ concentration, nutrient inputs, temperature, and precipitation affect the coupled nitrogen, carbon, and phosphorus cycles, altering plant stoichiometry and causing decoupling of some components. Ivory rectangles are ecosystem pools and characteristics and blue rectangles are processes. and the red triangles indicate points where there are controls that act on 'valves' altering plant nitrogen, carbon, and phosphorus. Stoichiometric coupling can affect individual plant performance and plant community dynamics over both short and long time scales. (After Z. Y. Yuan and H. Y. H. Chen. 2015. *Nat Climate Change* 5: 465–469.)

is large. Large amounts of carbon are also stored as carbonate ions ($CO_3^{2-}$) in rock and dissolved in seawater; additional carbon is contained in living biomass and in soils. Liquid water is the source of hydrogen (H), and water molecules are the ultimate source of the O released to the atmosphere in the light reactions in photosynthesis (see Figure 2.2). Rock weathering is the major source for most of the other elements needed by plants—such as calcium (Ca), magnesium (Mg), iron (Fe), and K and P—while S comes both from atmospheric deposition and rock weathering (Schlesinger 1997). In addition to new inputs from the atmosphere and the weathering of rock, plants rely on ecosystem recycling of elements. One of the most fascinating and unexpected examples of the range of such recycling resulting from erosion and deposition is the contribution of Sahara sand blowing west across the South Atlantic Ocean and providing nutrients for the rainforests in the Amazon Basin. This sand is particularly important for the transportation of phosphorus, which is lacking in the ancient soils of the Amazon basin (see Chapter 4) and is critical for plant growth (Yu et al. 2015).

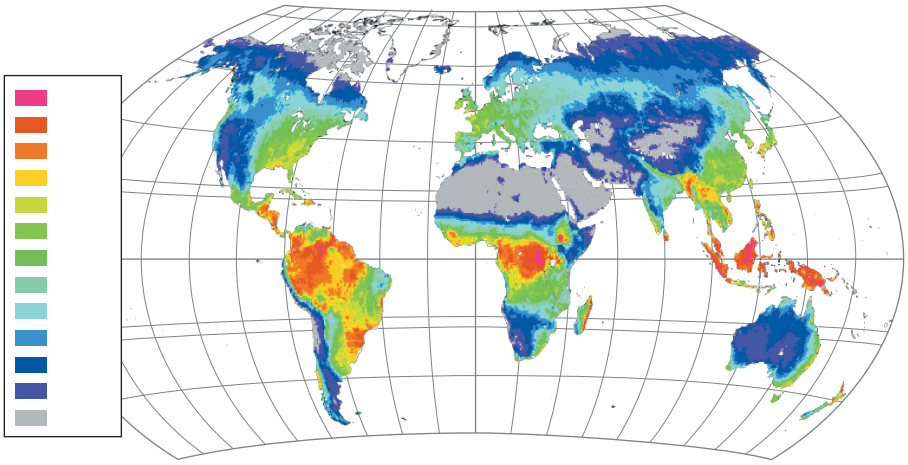
## 5.3 Carbon Is the Foundation of Life on Earth

### Productivity measures how carbon moves between living things and their nonliving environment

Carbon is the backbone of organic molecules (see Chapter 2). Because of its importance to both living and nonliving components of the Earth, ecologists have been interested for a long time in understanding and quantifying where it is stored and how it moves between the different ecosystem components (Lindeman 1942; Odum 1960; Woodwell and Whittaker 1968). **Productivity** is the rate of carbon or energy transformation from one trophic level to the next, per unit area per unit time. **Primary productivity** is the rate of transfer of carbon from oxidized carbon ($CO_2$), an inorganic form in the atmosphere, into reduced forms as organic carbon compounds in living organisms by photosynthesis. Productivity is often measured in terms of the amount of carbon transferred, because energy is captured with carbon in photosynthesis, and this energy is stored in carbon compounds. Photosynthesis links the water and carbon cycles because water leaves plants in transpiration through the same stomata through which carbon enters the leaves (see Chapter 3).

The total energy (or carbon) fixed by producers in an ecosystem is called **gross photosynthetic production**, or **gross primary production** (**GPP**). Since photosynthesizing organisms use some of the total energy fixed, **net primary production** (**NPP**) is equal to the total energy captured (GPP) minus the losses to respiration by primary producers. The term *primary* indicates that we are concerned with the productivity of the first trophic level in the system—that is, of the autorophic organisms that capture solar energy and incorporate it with atmospheric $CO_2$ into organic compounds—in contrast to the productivity of herbivores, carnivores, or detritivores. In practice, GPP and respiratory losses by primary producers are difficult to measure directly on an ecosystem basis, and most data on productivity are collected as NPP. The quantity of material transformed into organic forms by an individual plant can range from fractions of a gram to metric tons (1 metric ton = 1000 kg), depending on the size of the plant and the substance being transformed. At the ecosystem scale, this can add up to many metric tons of material. Terrestrial primary productivity is typically on the order of 5 to 10 metric tons per hectare. On the global scale, fluxes of carbon, for example, are on the order of $10^{15}$ g (gigatons, Gt, or picograms, Pg) per year (see Figure 16.23).

NPP varies enormously over Earth's surface (Figure 5.4). The greatest amounts of carbon are fixed in warm, moist ecosystems; much less is fixed in arctic systems, and the least in deserts (see Table 18.2). Across a variety



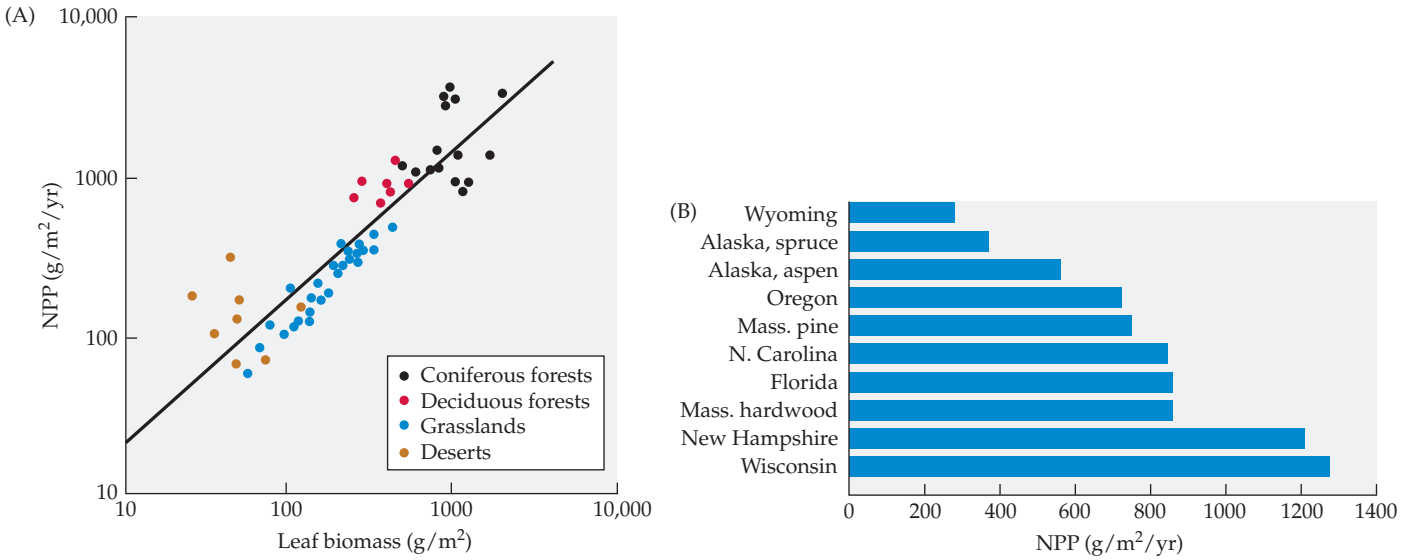



**Figure 5.5** Net primary production varies greatly on a continental scale. (A) NPP (above ground, in g/m²/year) across a wide range of ecosystems in North America, from deserts to grasslands to deciduous and coniferous forests. NPP is closely related to leaf biomass. The scale for both axes is logarithmic, so these variables have a very large range. (B) NPP (above ground, in g/m²/year) in ten North American forests. Aboveground NPP in these forests ranges from <300 g/m²/year to almost 1300 g/m²/year, with a median of about 750 g/m²/year. (A after W. L. Webb et al. 1983. *Ecology* 64: 134–151; B after J. D. Aber and A. Federer. 1992. *Oecologia* 92: 463–475.)

of ecosystems within the North American continent, for instance, aboveground NPP varies over more than two orders of magnitude (Figure 5.5A). Even if we consider only forests within this continent, the range in aboveground NPP is surprisingly large (Figure 5.5B). At the continental and larger scales, NPP depends mostly on climate (particularly as influenced by latitude and elevation), and soils (Figure 5.6).

NPP in terrestrial ecosystems is broadly limited by climate, particularly by sunlight duration and intensity, inadequate temperatures and insufficient moisture for growth, which together determine the length of the growing season (Figure 5.7A, B). NPP is also limited either by a lack of essential nutrients or by an excess of certain other materials (e.g., toxic levels of salt or heavy metal ions) in the soil. Soils determine and regulate plant mineral nutrition and water availability in concert with precipitation (see Chapter 4). Climatic or soil factors can influence properties of the vegetation,

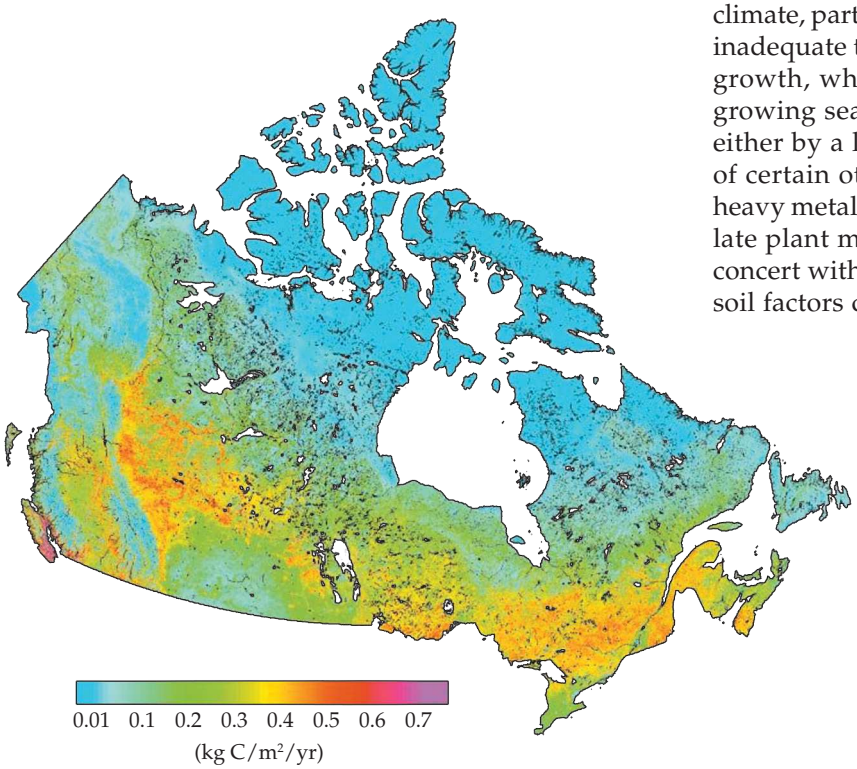
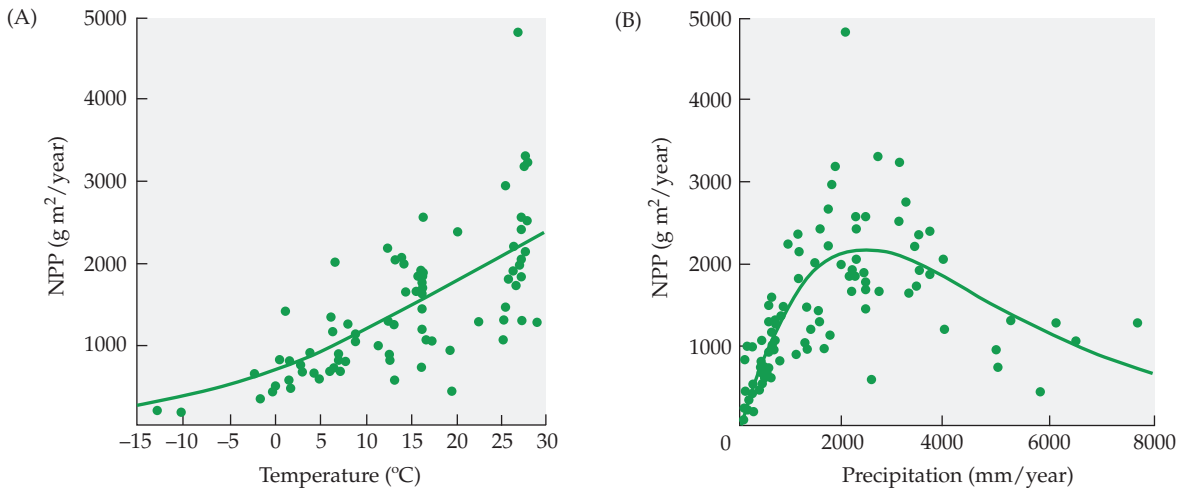



**Figure 5.6** Temperature and precipitation play a primary role in determining net primary production. NPP (above ground, in kg C/m²/year) across Canada, based on a simulation model with input for leaf area index and vegetation cover from satellite data, plus meteorological and soil data. (After J. Liu et al. 2002. *Global Ecol Biogeogr* 11: 115–129.)



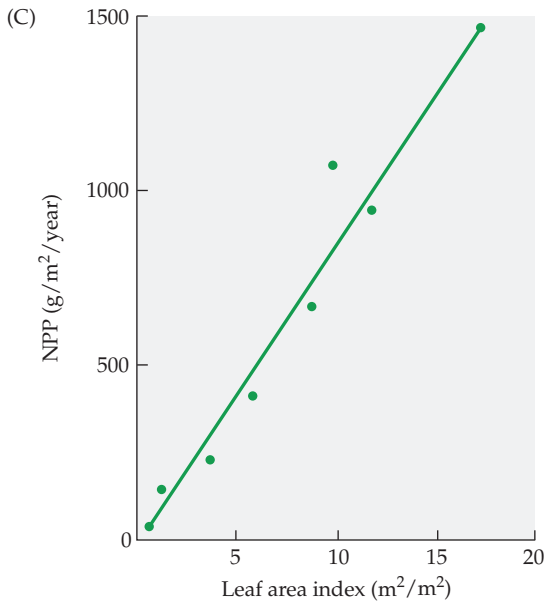




**Figure 5.7** NPP (above ground, in g/m²/year) for sites across several continents, as related to (A) mean annual temperature and (B) precipitation at or near each site. NPP is lowest in cold and dry ecosystems (tundra, desert) and greatest where temperatures are high and moisture is generally available to plants throughout the year (tropical rainforests). NPP declines at extremely high precipitation levels due to lack of oxygen in the soil and excessive leaching of soil nutrients. (C) Relationship between aboveground net primary production and leaf area index (defined as the total area of one side of all the leaves per unit of ground area) for forests in the northeastern United States. (A, B after F. S. Chapin et al. 2002. *Principles of Terrestrial Ecosystem Ecology*. Springer: New York; C after W. H. Schlesinger. 1997. *Biogeochemistry: An Analysis of Global Change*, 2nd ed. Academic Press: New York; based on H. L. Gholz. 1982. *Ecology* 63: 469–481.)

such as the leaf area index (see below) and total leaf area and biomass, which then affect NPP (Figure 5.7C). NPP can also vary within a given system as the species composition changes and plants grow bigger. For example, in forests, NPP typically rises over the course of succession (see Chapter 13) as leaf area index and biomass increase, reaching a plateau or even declining as the canopy closes and the proportion of the biomass in nonphotosynthetic woody tissues becomes greater.

Ecosystem models quantify current carbon cycles at local to global scales and may predict future scenarios (Figure 5.8). These models vary from being highly detailed to very coarse-scaled and general, require different inputs of data, and have different goals. Waring and Running (1998) compare some of the major carbon models used by ecosystem ecologists. One widely

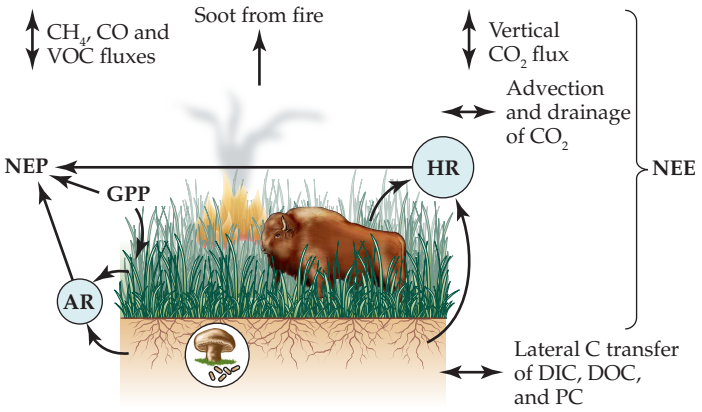



**Figure 5.8** Net ecosystem carbon balance (NECB) is the long-term net balance of all C inputs and losses of an ecosystem, including all components shown in the diagram. NECB includes net ecosystem exchange (NEE) plus methane ($CH_4$), carbon monoxide (CO), and volatile organic C (VOC); as well as fluxes of organic and inorganic carbon (DOC and DIC), and of particulate carbon (PC). NEE is part of NECB and includes the net exchange of $CO_2$ between the ecosystem and the atmosphere. Net ecosystem production (NEP) is part of NEE and includes gross primary production (GPP), autotrophic (plant) respiration (AR), and heterotrophic respiration (HR) from animals, plants, bacteria, and fungi. Particulate carbon is exchanged by animal movement, soot emissions from fires, water and wind deposition and erosion, and transport or harvest by people. (After F. S. Chapin et al. 2006. *Ecosystems* 9: 1041–1050.)

used model of soil carbon fluxes and pools is the CENTURY model; it has been updated as the Millennial model, a conceptual model that emphasizes measurable carbon pools (Parton et al. 1988, 1993; Abramoff et al. 2018) (Figure 5.9). The modeling of carbon fluxes is critical to our understanding of, and our ability to predict, human effects on the global carbon cycle (see Chapter 16). More recently, the development of terrestrial biosphere models (TBMs) integrates approaches from biogeography, vegetation dynamics models, biochemistry, and biophysics to better understand and predict vegetation responses to environmental factors, including global responses and feedback to anthropogenic climate change (Fisher et al. 2014) (Figure 5.10). TBMs can be used to understand systems at a range of scales, from leaf structure to landscapes, regions, and global scales. One goal of TBMs is to provide more accurate and detailed input to global climate models (GCMs; see Box 16B).

## Carbon is stored in the living and nonliving components of ecosystems

Over time, ecosystems typically accumulate and store carbon. Our vast underground stores of coal, for example, are the result of geological periods over which ecosystem carbon accumulation was high. **Net ecosystem production** (**NEP**) was initially defined by George Woodwell and Robert Whittaker (1968) as either the net accumulation of carbon per year by an ecosystem, or the GPP minus total ecosystem respiration (the sum of all autotrophic and heterotrophic respiration). The problem with this dual definition is that carbon can enter or leave ecosystems in ways that are not accounted for by the difference between GPP and total ecosystem respiration, such as the emission of methane ($CH_4$) and volatile organic carbon compounds (VOCs), and the leaching loss or deposition of dissolved organic carbon (DOC) in streams and groundwater. **Net ecosystem carbon balance** (**NECB**; see Figure 5.8) is the rate of carbon accumulation or loss in ecosystems. NECB is important for accounting for carbon gains and losses in models of climate change (Chapin et al. 2006) (see Chapter 16). During the growing season, NEP and NECB are usually positive because plant photosynthesis is far greater than heterotrophic respiration. Fires, logging, erosion, or other factors that remove carbon from an ecosystem can result in negative NECB. When plants are dormant in seasonal climates (winter or dry season) and photosynthesis is minimal, NEP is negative and NECB is also likely to be negative (Figure 5.11).

Both NECB and NEP generally depend on the length of time that an ecosystem has remained undisturbed,

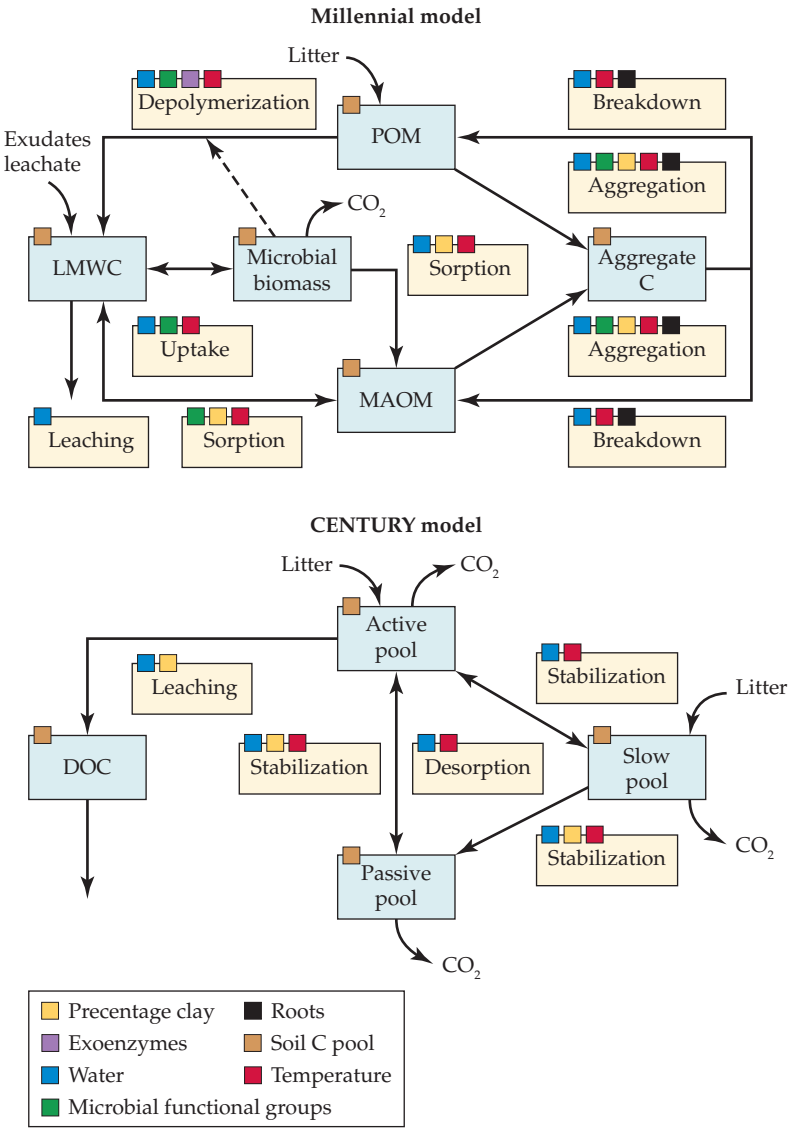



**Figure 5.9** Conceptual model diagram of the Millennial (top) and CENTURY models (bottom). The blue boxes are carbon pools, and the ivory boxes are fluxes. Solid arrows indicate the direction of each flux. The color legend indicates edaphic, biological, and climatic factors that may modify the rate of a given flux. Dashed lines indicate controls (i.e., microbial biomass regulates the depolymerization rate). The Millennial model has five measurable soil C pools: particulate organic matter (free fragments of plant detritus, POM); low-molecular-weight C (root exudates and the by-products of exoenzyme activity, LMWC); aggregate C; mineral-associated organic matter (MAOM); and microbial biomass carbon. (After R. Abramoff et al. 2018. *Biogeochemistry* 137: 51–71.)

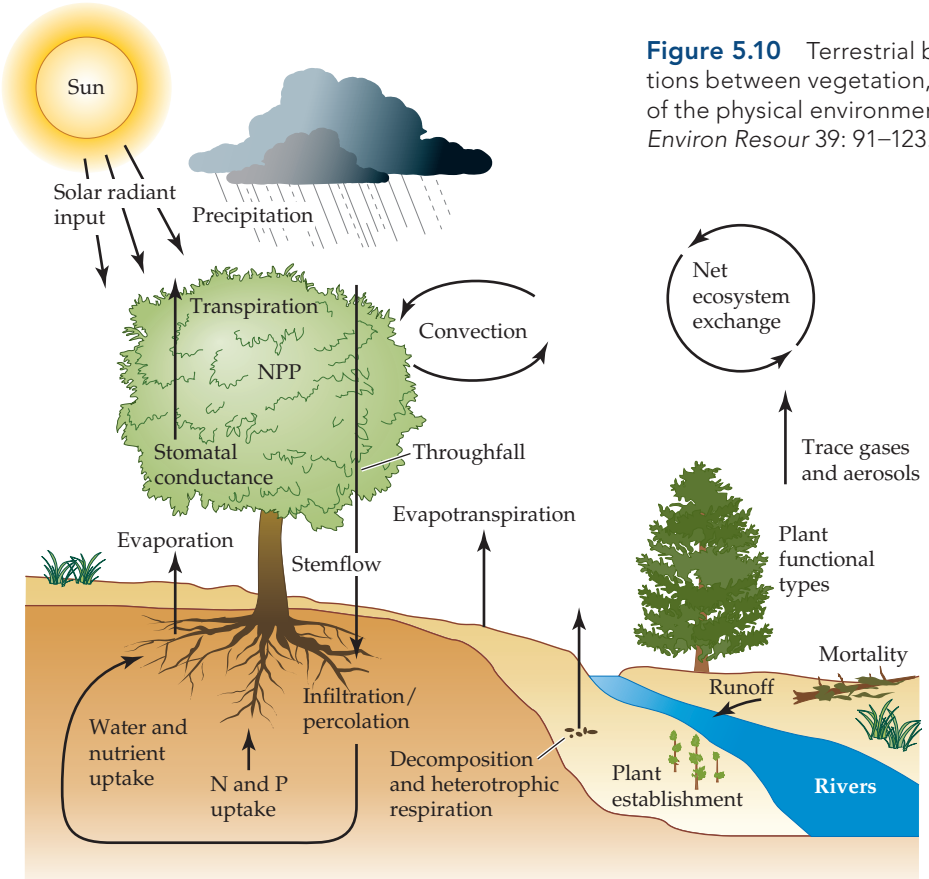



**Figure 5.10** Terrestrial biosphere models represent the interactions between vegetation, microbial processes and many aspects of the physical environment. (After J. B. Fisher et al. 2014. *Annu Rev Environ Resour* 39: 91–123.)

rather than on temperature or moisture. Undisturbed ecosystems typically show small positive accumulations of carbon each year. Disturbances typically remove large amounts of carbon and energy from ecosystems. Over the course of ecosystem development following a major disturbance, the biomass of woody tissue in long-lived trees and shrubs increases, as does soil organic matter in many ecosystems. In grasslands and other communities dominated by nonwoody herbaceous perennials, carbon accumulates in underground storage tissues and plant litter (discarded and dead leaves and other tissues). The sum of all organic matter in living vegetation (including the nonliving structural material in wood)—the **standing biomass**—may become very high, particularly in older forests, where carbon gradually accumulates in long-lived trees over many centuries. Eventually, the allocation of carbon to new woody tissue is balanced by losses due to the deaths of individuals and plant parts, and forest ecosystems no longer gain additional standing biomass (Odum 1969). The amount of carbon in the ecosystem may still continue to rise after that time, though, as soil organic matter and the carbon it contains continue to accumulate.

The relative proportions and total amounts of carbon and mineral nutrients stored in pools in different ecosystem components vary greatly among ecosystems. Flux rates and turnover among the components within ecosystems also vary. Pools and fluxes are related to one another by turnover time, or mean residence time, which measures how rapidly materials move through a system (see Chapter 4). For example, the pools, fluxes and turnover of biomass (mostly carbon) and major mineral nutrients in a California chaparral are shown in Table 5.1. Information about pools, fluxes and turnover times in various ecosystem components can reveal a great deal about the dynamics of

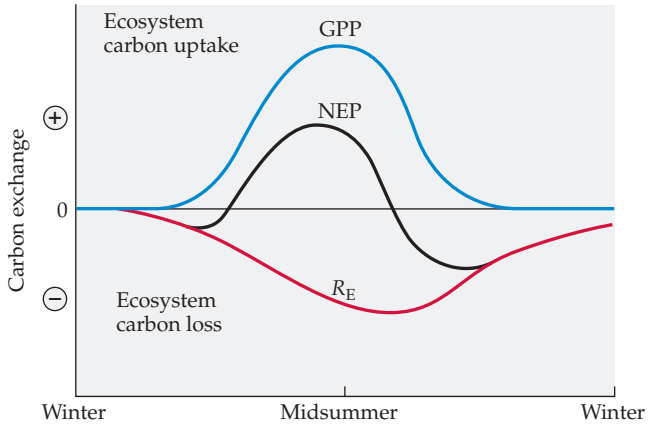



**Figure 5.11** Net ecosystem production (NEP) is the net accumulation of carbon per year by the entire ecosystem, including both living organisms and organic matter in the soil. In a typical temperate environment, NEP increases during spring and peaks in midsummer, when photosynthesis by plants is greater than total ecosystem respiration. Ecosystem respiration ($R_E$) is minimal during winter. It increases in spring, causing larger ecosystem carbon losses as soil temperatures warm. The largest carbon losses (most negative values) occur in late summer or early autumn. As GPP declines in autumn, NEP also declines, even though losses due to respiration are also being reduced (becoming less negative). (After F. S. Chapin et al. 2002. *Principles of Terrestrial Ecosystem Ecology*. Springer: New York.)

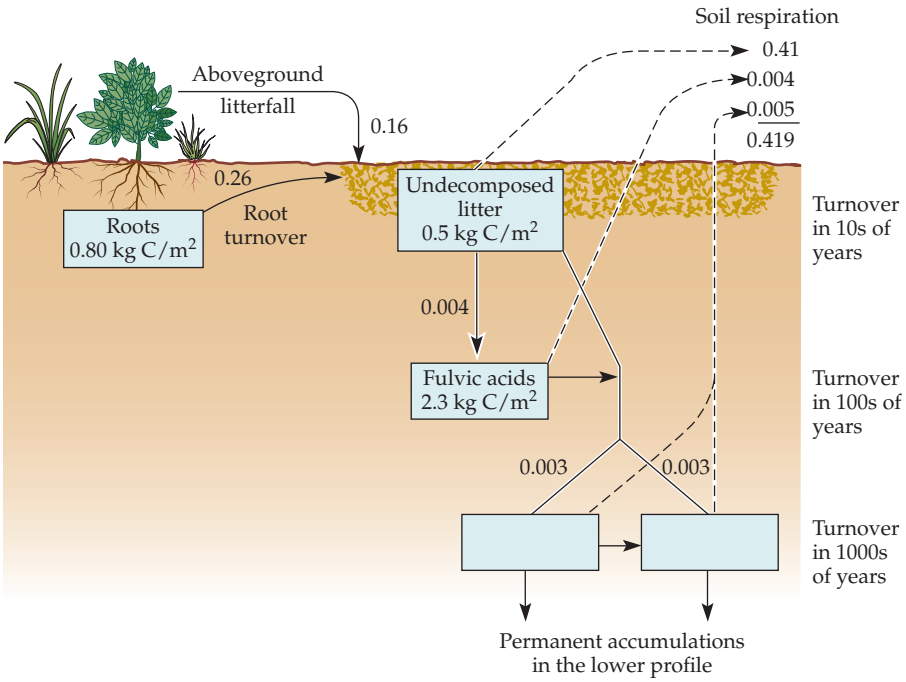
**TABLE 5.1** Nutrient cycling in a 22-year-old stand of the chaparral scrub *Ceanothus megacarpus* (bigpod ceanothus, Rhamnaceae) near Santa Barbara, California

| | Biomass | N | P | K | Ca | Mg |
|---|---|---|---|---|---|---|
| Atmospheric input (g/m²/year) | | | | | | |
| Deposition | — | 0.15 | — | 0.06 | 0.19 | 0.10 |
| Nitrogen fixation | — | 0.11 | — | — | — | — |
| Total input | — | 0.26 | — | 0.06 | 0.19 | 0.10 |
| Pools (g/m²) | | | | | | |
| Foliage | 553 | 8.20 | 0.38 | 2.07 | 4.50 | 0.98 |
| Live wood | 5929 | 32.60 | 2.43 | 13.93 | 28.99 | 3.20 |
| Reproductive tissues | 81 | 0.92 | 0.08 | 0.47 | 0.32 | 0.06 |
| Total live | 6563 | 41.72 | 2.89 | 16.47 | 33.81 | 4.24 |
| Dead wood | 1142 | 6.28 | 0.46 | 2.68 | 5.58 | 0.61 |
| Surface litter | 2027 | 20.5 | 0.6 | 4.7 | 26.1 | 6.7 |
| Annual flux (g/m²/year) | | | | | | |
| Requirement for production | | | | | | |
| Foliage | 553 | 9.35 | 0.48 | 2.81 | 4.89 | 1.04 |
| New twigs | 120 | 1.18 | 0.06 | 0.62 | 0.71 | 0.11 |
| Wood increment | 302 | 1.66 | 0.12 | 0.71 | 1.47 | 0.16 |
| Reproductive tissues | 81 | 0.92 | 0.08 | 1.47 | 0.32 | 0.07 |
| Total in production | 1056 | 13.11 | 0.74 | 4.61 | 7.39 | 1.38 |
| Reabsorption before abscission | — | 4.15 | 0.29 | 0 | 0 | 0 |
| Return to soil | | | | | | |
| Litterfall | 727 | 6.65 | 0.32 | 2.10 | 8.01 | 1.41 |
| Branch mortality | 74 | 0.22 | 0.01 | 0.15 | 0.44 | 0.02 |
| Throughflow | — | 0.19 | 0 | 0.94 | 0.31 | 0.09 |
| Stemflow | — | 0.24 | 0 | 0.87 | 0.78 | 0.25 |
| Total return | 801 | 7.30 | 0.33 | 4.06 | 9.54 | 1.77 |
| Uptake (= increment – return) | — | 8.96 | 0.45 | 4.77 | 11.01 | 1.93 |
| Stream water loss (g/m²/year) | — | 0.03 | 0.01 | 0.06 | 0.09 | 0.06 |
| Comparisons of turnover and flux | | | | | | |
| Foliage requirement/total requirement (%) | — | 71.3 | 64.9 | 61.0 | 66.2 | 75.4 |
| Litterfall/total return (%) | — | 91.1 | 97.0 | 51.7 | 84.0 | 79.7 |
| Uptake/total live pool (%) | — | 21.4 | 15.6 | 29.0 | 32.6 | 45.5 |
| Return/uptake (%) | — | 81.4 | 73.3 | 85.1 | 86.6 | 91.7 |
| Reabsorption/requirement (%) | — | 31.7 | 39.0 | 0 | 0 | 0 |
| Surface litter/litterfall (year) | 2.8 | 3.1 | 1.9 | 1.2 | 3.3 | 4.8 |

*Sources*: W. H. Schlesinger. 1997. *Biogeochemistry: An Analysis of Global Change*. Academic Press: New York; modified from J. T. Gray. 1983. *J Ecol* 71: 21–41 and W. H. Schlesinger et al. 1982. *Water Resour Res* 18: 623–629.

a system. In a study of a grassland system, the turnover rate for undecomposed litter was two orders of magnitude greater than that for humic acids (part of the soil organic matter; see Chapter 4), while the amount of carbon stored in the humic acids was much greater than in the litter (Figure 5.12).

NPP tends to be closely tied to soil microbial respiration because the carbon and energy fixed by plants is the substrate for microbial respiration, and because microbial respiration releases nutrients that plants need in order to capture carbon by photosynthesis (as well as to survive and grow). In some cases, however, there is




We return to our example of salt marsh ecosystems to look at productivity. Maximum rates of net primary productivity in salt marshes can be among the highest of any habitat, rivaling NPP of tropical rainforests at up to 2000 g C/$m^2$/year or more (Dame et al. 2000), but NPP in salt marshes can also be much lower, depending on geography as well as local conditions (Tobias and Neubauer 2019). Much of the biomass in a salt marsh is below ground, in the form of roots and rhizomes. Plants experience an environment that is both stressful and advantageous, with high sun, high nutrients, occasional to regular flooding, low oxygen, and moderately to highly saline soils resulting in low osmotic potentials (see Chapter 3). The "brown food web" or soil food web of microbial decomposers (see below) is central to productivity in this system: particulate organic carbon (shredded bits of plants and other organisms), dissolved organic carbon, and dissolved inorganic carbon move into the water and the soil and move from seagrasses and algae in the water back into the terrestrial environment.

no strong link between NPP and soil microbial activity, and this can affect NEP. For example, in peat bogs, the anaerobic soil environment severely limits microbial activity but does not strongly limit plant productivity, resulting in high carbon accumulation in the soils and high NEP and NECB. Factors that remove living plant biomass or soil organic matter result in losses of carbon from the ecosystem and negative NECB and can alter subsequent NPP in complex ways.

## Soil food webs are the recycling engine of terrestrial ecosystems

From the shortest-lived microorganisms to the longest-lived trees (Figure 5.13), everything that lives eventually dies. If organic **detritus**—the waste products and dead bodies of living and once-living organisms—were not decomposed, most ecosystem processes would come to a screeching halt, and needless to say, the piled-up wastes and dead bodies would not be a pretty sight. Sooner or later, much of the carbon and other materials and some of the energy released by the process of decomposition are taken up by living organisms and used again. How are materials that were once part of living organisms recycled?

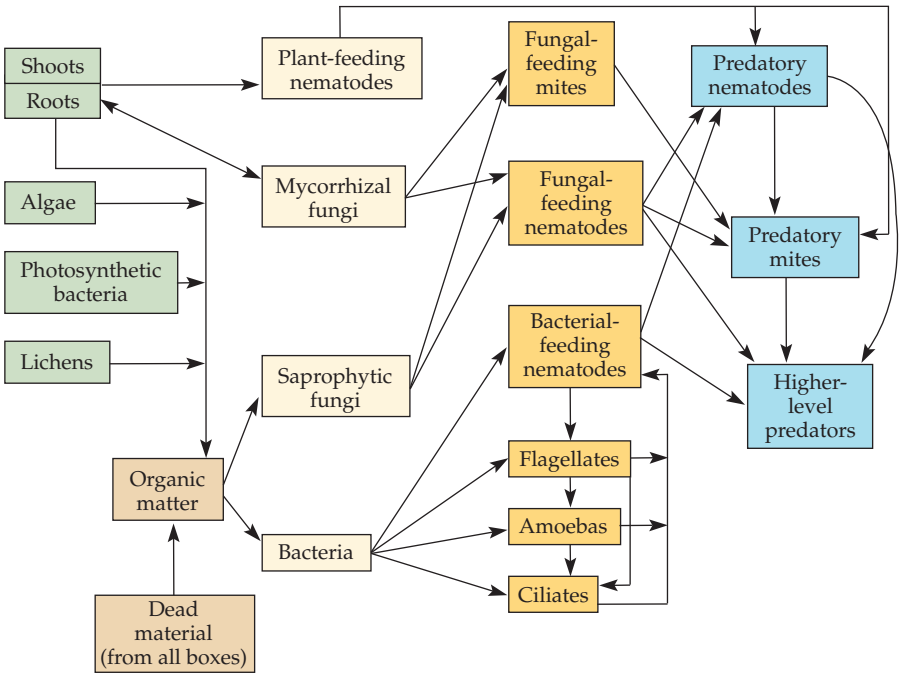



**Figure 5.14** Diagram of a soil food web (brown food web). Autotrophs (green boxes) at the soil surface and in upper layers of the soil include plants, with photosynthetic parts above ground and roots below ground, photosynthetic single-celled protists (including diatoms and green algae), and cyanobacteria (photosynthetic bacteria). Heterotrophs include many different kinds of bacteria, single-celled animals (amoebas, ciliates, flagellates), mycorrhizal and saprophytic (decomposing) fungi, nematodes, soil microarthropods and macroarthropods, including mites, and a variety of vertebrates as well. Arrows indicate trophic interactions, including consumption, mutualism (mycorrhizae and plants) and decomposition. (Courtesy of E. R. Ingham, at http://www.soilfoodweb.com, modified from H. W. Hunt et al. 1987. *Biol Fert Soil* 3: 57–68.)

Decomposition transforms dead organisms, their shed or removed parts (e.g., leaves, bark, branches, and roots), and animal feces and urine into soil organic matter and ultimately into inorganic nutrients, $CO_2$, water, and energy. The initial stages of decomposition transform detritus into soil organic matter. Further decomposition of the organic matter occurs via a set of processes collectively called **mineralization**, in which soil microorganisms use the energy in the chemical bonds of the organic matter, releasing the carbon in respiration as $CO_2$ and converting the nutrients to inorganic forms. The nutrients may then be taken up by plants directly, or they may be taken up after further microbial conversions. From the functional perspective of the ecosystem, this process is a reversal of photosynthetic primary production, although it involves very different biochemical processes.

In terrestrial ecosystems, decomposition involves both physical and chemical alteration of the original materials by a complex soil food web (also called a brown food web, in contrast to the green food web based on primary producers), which exists mostly (but not entirely) on or just under the surface of the soil (Figure 5.14). Decomposition is largely an aerobic process (one requiring oxygen), and it occurs slowly, if at all, in waterlogged soils or other anaerobic environments such as those in salt marshes and other wetlands. Warm, moist conditions favor the highest rates of decomposition. Cold, anaerobic (oxygen-free) conditions inhibit decomposition, resulting in the buildup of organic material. Decomposition rates also depend on other factors, especially the physical and chemical characteristics of the material being decomposed (Figure 5.15).

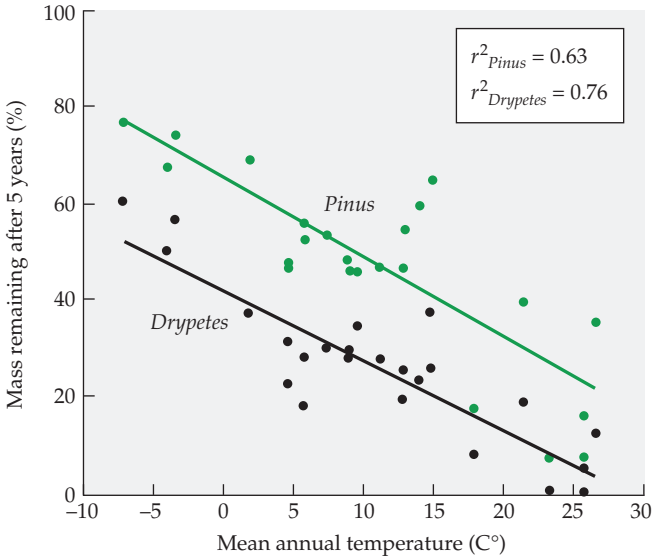



**Figure 5.15** Decomposition rates of root material of two species of trees with highly contrasting properties: *Drypetes glauca* (Euphorbiaceae), a tropical hardwood, and *Pinus elliottii* (slash pine, Pinaceae), a temperate softwood. Mesh bags containing root material were buried in 28 sites in North and Central America. The roots of both species decomposed faster at sites with higher annual temperatures. *Drypetes* roots decomposed faster than *Pinus* roots at all sites. (After H. L. Gholz et al. 2000. *Global Change Biol* 6: 751–765.)

Litter decomposition in both nutrient-poor and arid soils is typically slow. For instance, because conifer needles and oak leaves decompose very slowly and tend to acidify soils under moist conditions, the presence of large numbers of these trees can result in positive feedbacks that favor plants like conifers and oaks with leaf traits that reduce nutrient loss (see Chapter 10). Thus, on nutrient-poor soils, the evergreen habit tends to be self-reinforcing (Aerts 1995). Indeed, the dominance of conifers in boreal forests plays a role in the acidification and nutrient depletion of these soils.

Scavenger animals begin the process of decomposition of materials of animal origin by consuming animal bodies. These scavengers include large animals (such as hyenas and vultures) as well as insects (such as burying beetles and the dipteran larvae known as maggots). Dung beetles (in the family Scarabaeidae, the scarabs of ancient Egypt, which occur globally) are important and astonishingly effective consumers of mammal feces in tropical grasslands and in many other ecosystems; their diversification is closely correlated with the rise of the ecological dominance of angiosperms in the lower Cretaceous. Nicole Gunter and colleagues (2016) hypothesize that as angiosperms became prolific, the feces of herbaceous dinosaurs eating them provided broad new niches for scarab beetles.

The most important source of material for decomposition is **plant litter**, which consists of a variety of materials—from dead roots and shed leaves, needles, or bark to dead tree trunks. Litter must be broken up into small pieces of organic matter before it can be effectively decomposed by microorganisms. This process of litter fragmentation is largely accomplished by animals and abiotic processes such as freezing and thawing. Larger animals, from deer and bears to gophers and voles, break litter apart as they search for food in the soil and on its surface, as well as consuming both living and sometimes nonliving plant material.

At an intermediate scale, earthworms, where they are present, ingest and process huge volumes of soil, causing dramatic physical and chemical alteration of the soil and the organic matter contained in it. Termites consume living and dead plant material, particularly decomposition-resistant wood, digesting cellulose and lignin by means of mutualistic protists and bacteria that inhabit their hindguts. These insects can be major contributors to tropical ecosystem dynamics. The effects of these soil animals can interact with the effects of plants to determine decomposition rates. In an experiment in a Swiss hardwood forest, Stephan Hättenschwiler and Patrick Gasser (2005) reported that pill millipedes (Figure 5.16A) and earthworms greatly speeded up

(A)

Hans Hillewaert/CC-BY-SA 4.0

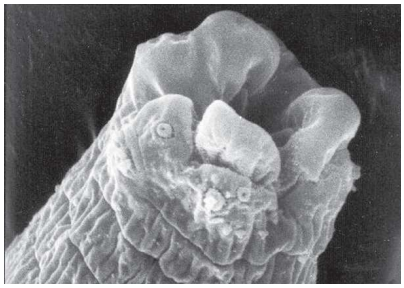
(B)

Phillipe Garcelon/CC BY 2.0

(C)

© Scenics & Science/Alamy Stock Photo

(D)



Courtesy of Sven Boström

**Figure 5.16** Some invertebrates that are important in soil food webs. (A) A common European pill millipede, *Glomeris marginata*, from France; adults are 8–20 mm long. (B) *Sminthurinus elegans* (a springtail; Collembola). Springtails are very common tiny soil insects that consume fungi. The organism pictured is about 3 mm long. (C) A soil mite. These tiny (about 2 mm) arachnids consume fungal hyphae and spores, as well as algae, lichens, and bacteria. Soil mites are common worldwide. (D) A nematode, *Acrobeloides nanus*, showing its front (anterior) end. This nematode species is probably one of the most common invertebrates in the world. It lives on all continents, including the subantarctic islands, and in favorable sites may have populations of over a million individuals per square meter. Its entire length is about 0.2 mm.

decomposition of the leaves of different common forest tree species, and decomposition tended to be much larger when different species of leaves were mixed together than for leaves of individual species.

The smallest multicellular animals in soil food webs are microscopic insects and arachnids known as soil microarthropods, or soil microfauna and mesofauna. These organisms are particularly important in the decomposition process, especially in forests. Among them are the Collembola (springtails, Figure 5.16B), which are insects that eat fungi, and arachnids (particularly mites, Figure 5.16C) that directly consume litter as well as eating soil fungi and bacteria. Nematodes (Figure 5.16D) are spectacularly abundant and critical in soil food webs as herbivores that eat plant roots, carnivores that eat soil animals, and consumers of soil-dwelling bacteria and fungi. Unicellular protists, including ciliates and amoebas, live in the film of soil water between and surrounding soil particles and prey on soil bacteria. By their voracious consumption of soil bacteria and fungi, soil microarthropods, nematodes, and protists are responsible for converting large amounts of microbial nitrogen and phosphorus into forms available to plants, as well as releasing microbial carbon as respired $CO_2$.

**Saprophytic** fungi rely on nonliving organic material for their carbon and energy. They are major decomposers of dead leaves and other plant litter and account for approximately half of the microbial biomass in grassland soils and about 90% in temperate forest soils. These fungi are heterotrophs that externally secrete powerful digestive enzymes, decomposing their substrate and ingesting the products of the decomposed material. The extensive multicellular mycelia that make up a fungus can obtain carbon from one source and nitrogen from a different source at some distance, further enhancing the ability of fungi to use dead plant material as a source of nutrients and energy.The fine hyphae of the fungi can penetrate plant organs and cells, piercing the defensive cuticle to gain access to the interior and speeding decomposition. However, both plant structural defenses, such as lignin, and various aspects of plant defensive chemistry, including toxins and pH levels, can inhibit their effectiveness. Such differences in plant chemistry and physical structure contribute to large differences in decomposition rates among different plant species (see Figure 5.15) and different plant parts (e.g. leaves versus wood).

### Soil organic matter revisited: Bacteria are an essential component of soil organic matter

Bacteria play an essential role in soil food webs and nutrient cycling. Bacteria can rupture and break down living and dead cells of plants, animals, fungi, and other bacteria. Their presence in the soil is highly variable in space and time; they are found in greatest abundances at warm temperatures and when water and nutrients are available. They are found largely in soil pores, in the rhizosphere, and coating aggregates of soil particles (see Chapter 4). They move through the soil passively, carried by water or animals.

Soil microorganisms are an essential component of soil organic matter both because dead bacteria are a major component of soil organic matter and because they produce materials that contribute to it (Kallenbach et al. 2016; Kästner and Miltner 2018). Soil organic matter consists of continually decomposing organic compounds that interact dynamically with clay particles and the bacteria in the soil (Lehmann and Kleber 2015). Plant biomass input to soil organic matter is less from aboveground material such as discarded leaves and more from decaying roots and root exudates (Jackson et al. 2017; Sokol et al. 2019), and much of the decomposition of this material occurs just under the soil surface where roots proliferate and conditions favor decomposition. Like most things about soils, much more is known about soils in agricultural than ecological systems. We return to the roles of various functional groups of bacteria below, because they play a critical role in nitrogen as well as carbon cycling.

### Primary productivity can be measured or estimated in a variety of ways

Measuring productivity and being able to predict its magnitude has great commercial value (for instance, in determining timber production). It is also tremendously important for understanding local ecosystems and as a basis for understanding and predicting components of global carbon balance. While the concept of productivity may seem simple, actually measuring it raises various questions and involves solving some challenging problems. While on-the-ground measurements have been used for many years, at larger spatial scales, important tools for estimating NPP include eddy flux towers and **remote sensing** methods (Box 5B, see chapter heading photo).

Ecologists might need to measure productivity in a local area, such as a forest preserve, or a larger area, such as a national park. If they want to know how productivity affects community structure, they might measure the potential productivity of a site—how much carbon can be captured at the site under optimal conditions—or the actual productivity—the amount actually fixed by the vegetation in a given time period. These will give different answers. The actual productivity will always be less than potential productivity because many factors limit plant growth, including competition, drought, heat, and cold. The direct loss to herbivores would not count as a reduction in productivity (after all, the carbon was captured by the plants before being eaten, and this carbon contributes to the secondary productivity of the herbivores), but the missing plant tissue would result in a reduction in productivity in the future. In general, we are more concerned with actual productivity, and we

## BOX 5B Using Remote Sensing and Eddy Covariance Methods to Estimate NPP

Remote sensing methods can be used to estimate NPP at larger scales than other methods. Images that record the different wavelengths of light reflected by the ground surface can be analyzed to provide an estimate of *greenness*, or the amount of chlorophyll present, because chlorophyll absorbs (and reflects) light at specific wavelengths (see Chapter 2 and Box 15A). Since the amount of chlorophyll is related to the amount of live photosynthetic material, this information can be used to estimate live biomass at a site or in a region. Successive estimates of biomass can be used to obtain an estimate of productivity.

The National Ecological Observatory Network (NEON; www.neonscience.org) employs data-intensive measurements to better understand the status of Earth's terrestrial and aquatic ecosystems; it has extensive and ongoing collections of spatial data, biogeochemistry, and many other kinds of ecological data, including data on ecosystem productivity and other ecosystem processes. To determine NPP, imaging spectrometers are deployed on satellites or planes to quantify the wavelengths useful for detecting vegetation and monitoring its condition. Multispectral and hyperspectral resolution imaging spectrometers make highly accurate measurements of NPP and many other characteristics and properties of plants at larger spatial scales. More recently, drones are being used as well to gather information at more local scales.

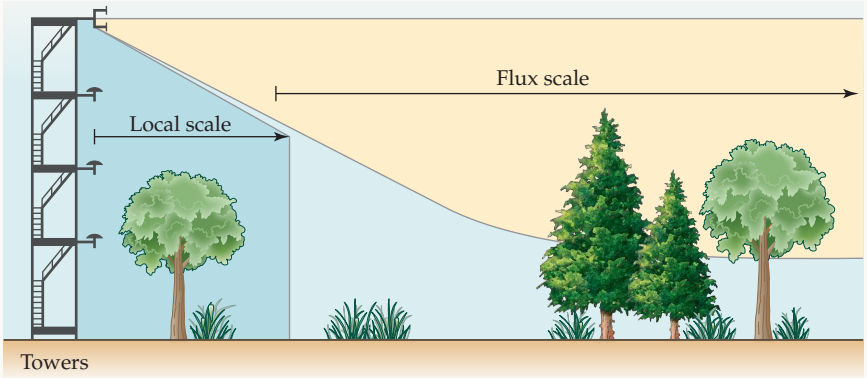
Images from satellite and aerial photographs are used to calculate the **normalized difference vegetation index** (**NDVI**) which measures standing biomass based on the reflectance at all the near-infrared (NIR) and visible (VIS) wavelengths:

$$NDVI = \frac{(NIR - VIS)}{(NIR + VIS)}$$

where the reflectance is integrated over all the relevant wavelengths. Regressions relating NDVI to NPP have been worked out by relating the reflectance values to those from actual harvests of vegetation. NDVI has been widely used and provides good estimates of relative NPP in many systems. Accuracy has been greatly increased by technological developments and by simple approaches such as measuring NDVI at different times of the year. If an absolute measure of NPP is needed, it may be necessary to calibrate the measurements using on-the-ground harvests. New methods have been developed to take images that reveal vegetation layers under canopies even for sites with multiple vegetation layers.

Another approach to the indirect estimation of productivity is the use of techniques based on micrometeorology that measure concentrations of $CO_2$ in the air. Networks of micrometeorological stations across the world integrated through FLUXNET (https://fluxnet.fluxdata.org/about/) continuously measure fluxes of gases between vegetation and the atmosphere, as well as biologically important climate and weather features including temperature, humidity, wind speed, rainfall, and atmospheric carbon dioxide. FLUXNET is perhaps the most extensive environmental and ecosystem network in the world. These measurements provide critical input not only for biogeochemistry and ecosystem ecology, but also for meteorology and climatology. NEON is coordinating the use of large flux towers in many different sites across the U.S. to conduct such measurements. These techniques have become increasingly important in measuring productivity (and other plant-atmosphere interactions) since the 1980s. The eddy covariance (or eddy correlation) method is the most widely used of these approaches. This technique measures **net ecosystem exchange** (**NEE**), which is the net transfer of $CO_2$ into or out of an ecosystem.

Eddies are parcels of air that move essentially as discrete units and within which gas concentrations are approximately uniform. Eddies carry $CO_2$ into and out of an ecosystem, which consists of a complex mix of carbon sources (such as respiring organisms) and sinks (primarily photosynthesizing plants). During the daytime, by far the most important sink for $CO_2$ is its uptake by the photosynthesizing leaves in the ecosystem. As an eddy comes into contact with a surface, such as a leaf, it slows down because of friction. At that point, $CO_2$ can diffuse out of the eddy and into the boundary layer of the leaf. When conditions are favorable, there will be a net transfer of $CO_2$ into the ecosystem from the surrounding air due to photosynthetic uptake by the canopy (that is, gross uptake minus losses due to respiration in the ecosystem). During the night, there will be a net transfer of $CO_2$ out of the ecosystem into the air, due primarily to respiration by the plants and soil organisms in the ecosystem.




Sensors on the tower collect flux scale data and microclimate data. (After L. Goldman and C. Sturtevant. 2015. *The power of measuring vertical wind profiles using NEON towers.* National Ecological Observatory Network (NEON). https://www.neonscience.org/observatory/observatory-blog/)

focus on that here, although making this distinction is important for sorting through theories about how productivity determines species diversity (see Box 19B).

NPP can be estimated in a variety of ways. At local scales we need to sample the area to estimate productivity (see Chapter 12). In the simplest cases, we would start the year with a patch of bare ground. At the end of the year, we would measure the standing biomass by harvesting all of the plants, drying them, and weighing them. In perennial herbaceous ecosystems, standing biomass is harvested from matched plots at the beginning and at the end of the growing season, with the difference in mass providing an estimate of the amount of material produced. Losses to herbivory and other factors must be estimated to obtain an accurate assessment of NPP. One way to estimate the magnitude of such losses is to exclude herbivores from other matched plots by fencing, insecticides, or other means.

Harvesting plant material to measure productivity has two important limitations. First, harvesting standing biomass is destructive. This method would not be very useful, for instance, in a site that we wanted to study over many years. One way around this limitation is to measure productivity in a number of harvested plots while other measurements are made in adjacent, undisturbed plots. Second, harvesting only aboveground plant parts ignores belowground productivity, which is important in many ecosystems. Measuring belowground productivity, however, is difficult. Not only is there a great deal of labor involved in digging up roots, but it creates a substantial disturbance in a plant community under study. Furthermore, much of the root material may be contained in very fine root hairs that are hard to separate from the soil. For these reasons, belowground productivity is often (if regretfully) ignored. Various methods for estimating root growth and productivity have been developed to overcome these limitations, including small, specialized cameras that can be inserted into the soil to observe and quantify belowground root growth and interactions (Cheng et al. 1990). The ratio of belowground to aboveground NPP can vary greatly among ecosystems. Kristina Vogt and colleagues (1982), for example, compared aboveground and belowground NPP in 23-year-old and 180-year-old *Abies amabilis* (silver fir, Pinaceae) forest stands in the Cascade Range of Washington State. In the young forest, belowground NPP was 45% of the total; in the old forest, the belowground component was 75% of total NPP.

To get around some of the limitations in measuring NPP and to extend measurements to larger spatial extents, ecologists have increasingly employed more technologically advanced approaches. Some of the most important of these approaches are the use of large flux towers that measure gas exchange using eddy covariance approaches. **Eddy covariance** methods are a micrometeorological approach that has become widely used to quantify $CO_2$ exchange as well as fluxes of energy, water vapor and other gases between plant canopies and the atmosphere. They depend on advanced instrumentation in many different kinds of natural vegetation and crops to quantify NPP and other relationships between vegetation and the atmosphere. The approach works by determining the covariance (a statistical measurement of how changes in one variable relate to changes in another variable) of fluctuations in wind velocity and in $CO_2$ or other substances being measured. At even larger spatial extents, aerial and satellite images are analyzed using remote sensing approaches (see Box 5B).

A widely used practical approach at more local scales depends on developing formulas relating changes in plant size to changes in biomass. For many tree species, especially economically important ones, the statistical relationship between average annual increases in stem diameter and average annual biomass accumulation in stem wood, leaves, branches, and large-diameter roots has been determined by cutting and measuring trees. These relationships are then used to develop formulas to estimate NPP and stores of carbon and nutrients in the different tree components at different stages of forest development. These formulas are widely used by forest ecologists and timber companies.

A very simple, indirect estimate of productivity can be obtained by extrapolating from observed correlations between actual evapotranspiration (AET, see below) and productivity (Rosenzweig 1968). While it can provide fairly good estimates of productivity in some climates, it is a poor predictor of productivity in tropical regions, and so it is not very useful for predicting global productivity.

A different approach to estimating long-term NPP based on analyses of soil organic matter can provide an estimate of NPP over a much longer period than is usually possible (Jenkinson et al. 1999). It estimates NPP by calculating how much organic matter must enter a soil annually to maintain the existing pool of soil organic matter at a steady state. The measurements used are the soil mass in each soil layer, the organic carbon content of each soil layer, soil microbial biomass and the radiocarbon age of each soil layer based on carbon-14 measurements. While this method has certain limitations and assumptions, under the right circumstances, it can be a valuable complement to other approaches to measuring NPP.

## 5.4 The Nitrogen Cycle Is an Essential Component of Ecosystems

Living organisms other than plants, particularly bacteria, are essential components of the nitrogen cycle. Nitrogen is essential for life because it is a fundamental component

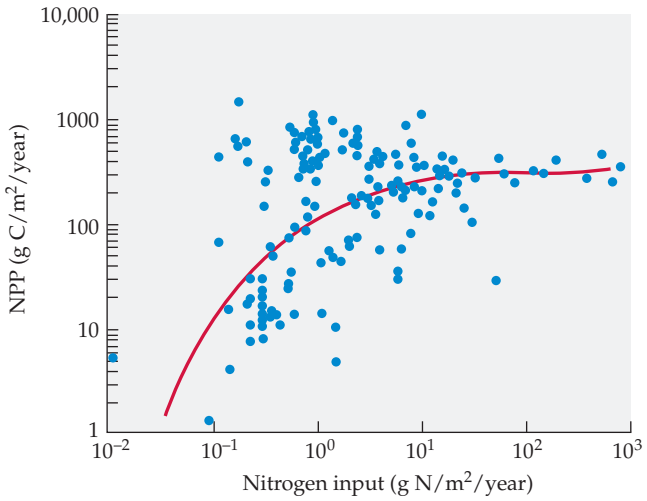



**Figure 5.17** Net primary production versus annual nitrogen input for a wide range of terrestrial, freshwater, and marine ecosystems. NPP levels off at about 10 g N/m²/year. (After W. H. Schlesinger. 1997. *Biogeochemistry: An Analysis of Global Change*, 2nd ed. Academic Press: New York; based on J. R. Kelly and S. A. Levin. 1986. In: *The Role of Oceans as a Waste Disposal Option*, G. Kullenberg [Ed.] NATO ASI Series, vol 172, pp 165–203. Springer: Dordrecht.)

of amino acids, nucleic acids, plant pigments such as chlorophyll, and other essential molecules (see Chapter 4). It is almost always limiting to individual plant growth and to community and ecosystem productivity (Figure 5.17).

Nitrogen moves quite rapidly through living organisms but is coupled to a large global pool of elemental $N_2$ in the atmosphere with slow turnover (Figure 5.18). Nitrogen enters the biosphere from the atmosphere through the process of organic and inorganic (e.g., via lightning) nitrogen fixation (see Chapter 4). Eventually it is returned to the atmosphere by the process of **denitrification** (and by other means, including volatilization in fires). The balance between nitrogen fixation and denitrification maintains the concentration of $N_2$ in the atmosphere at equilibrium (Schlesinger 1997). Nitrogen is recycled among living things at a relatively rapid rate, so primary productivity does not depend solely on nitrogen fixation. We focus here on the processes by which nitrogen is stored and moves between plants and other ecosystem components.

Organisms can interact in surprising ways to alter nutrient cycling and ecosystem functioning, with major consequences at least at local scales. For example, along the Pacific coast in northwestern North America, millions of salmon return each year to the streams and

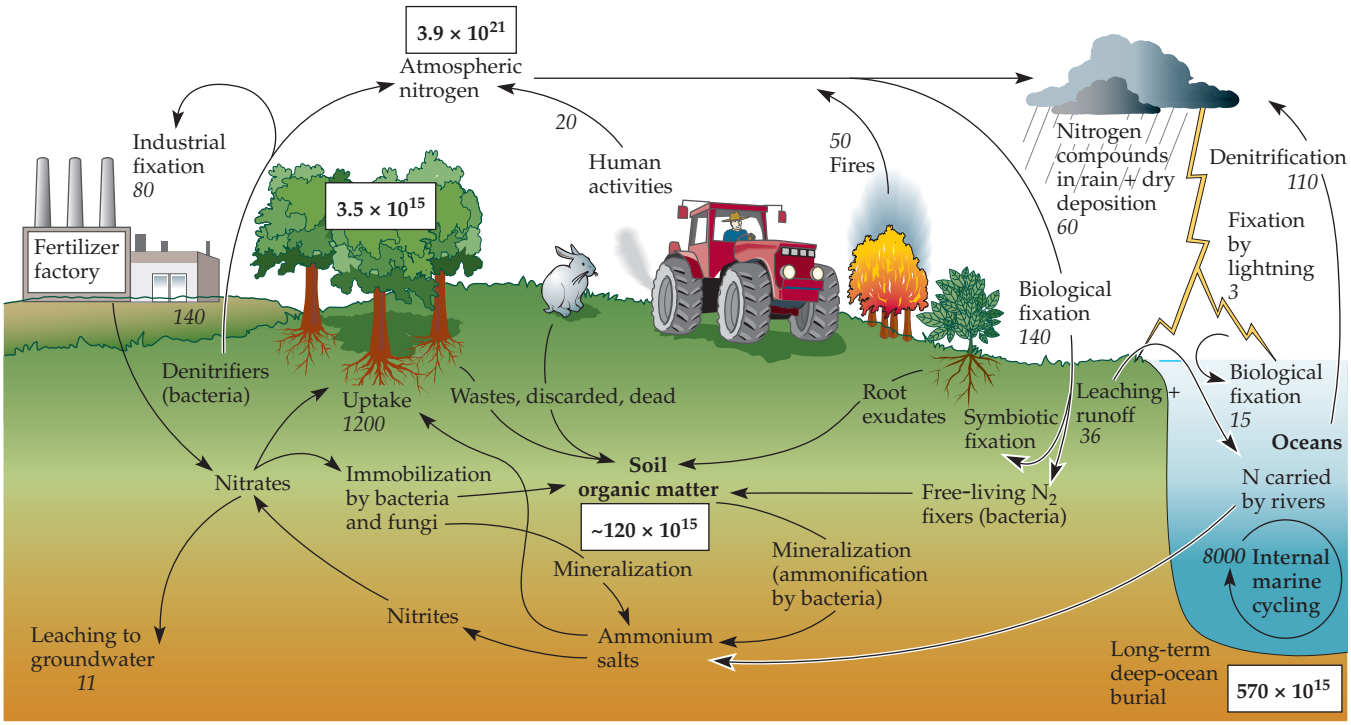



**Figure 5.18** The global nitrogen cycle, showing some of the major fluxes (in $10^{12}$ g N/year, values in italic type) and pools (shown in boxes; units as indicated). The atmosphere is by far the largest pool of nitrogen, but living plants, soil organic matter, and long-term ocean burial are also substantial. (Data from W. H. Schlesinger. 1997. *Biogeochemistry: An Analysis of Global Change*, 2nd ed. Academic Press: New York; L. Taiz and E. Zeiger. 1998. *Plant Physiology*, 2nd ed. Oxford University Press/Sinauer: Sunderland, MA.)

rivers in which they were hatched, swimming upstream to reproduce and die. In doing so, they carry large quantities of nutrients in their bodies, transferring these from marine environments to freshwater ecosystems. Some of these nutrients are also transferred to adjacent terrestrial systems; for example, capture of salmon by bears can result in substantial local increases in soil nitrogen and in plant productivity from bear feces and urine (Cederholm et al. 1999; Schindler et al. 2003).

## Bacteria mineralize organic N to inorganic forms taken up by plants

Most of the nitrogen used by plants comes from the recycling of N from decaying biomass (Figure 5.19 and see Figure 5.18). Biological nitrogen fixation from free-living and symbiotic bacteria accounts for a much smaller component of nitrogen uptake by plants globally (see Chapter 4). The two most common forms of nitrogen taken up by plants are nitrate ($NO_3^-$) and ammonium ($NH_4^+$). Most (but not all) plants preferentially take up $NO_3^-$ rather than $NH_4^+$. The $NH_4^+$ ions in soil are adsorbed onto clay and organic matter particles, due to the negative charge of those particles. If there are sufficient amounts of these particles in a soil, the $NH_4^+$ may be retained and build up in soil. In contrast, $NO_3^-$ ions are repelled by the negative charge in clay and organic matter, and they are readily leached (see Chapter 4). Some plants are able to take up ammonium, particularly when they are young, and plants adapted to acidic soils or soils with limited oxygen potential (such as waterlogged soils) preferentially take up ammonium. The costs and benefits of nitrate versus ammonium uptake are complex and depend on their concentrations in the soil, as well as other factors. Nitrate must be reduced to ammonia before it can be incorporated into organic compounds in the plant, but as nitrate it is highly mobile and can be transported readily in dissolved form in the xylem and stored in cell vacuoles. Ammonium is largely incorporated into organic compounds in the roots directly after it is taken up. Ammonium can be toxic to plants, particularly when converted to ammonia ($NH_3$). A small number of plant species are able to take up organic nitrogen in the form of amino acids directly from the soil. Otherwise, plants can use these amino acids in the soil only if they are first converted by bacteria to ammonium ions ($NH_4^+$).

The process of N mineralization begins with **ammonification**, by which nitrogen in organic compounds, such as proteins and nucleic acids, is released in the form of $NH_4^+$. Ammonification can be carried out by a variety of soil-dwelling microorganisms. In ammonification, heterotrophic microorganisms release N from organic matter by breaking the chemical bonds in carbon compounds to release fully reduced $NH_4^+$ in a series of enzymatically controlled steps. Various things can happen to the $NH_4^+$ released by the nitrifying bacteria, including being used by the same or other microorganisms or by plants. Ammonium ions do not move readily in the soil and are likely to be taken up only by roots in close proximity. The rate at which ammonification occurs is determined by a variety of environmental factors, particularly temperature, moisture, and the chemical characteristics of the soil (including pH).

The $NH_4^+$ ions may be directly taken up by some plants, remain in the soil, or undergo nitrification (see Figure 5.19). **Nitrification** is the oxidation of $NH_4^+$ to nitrite ($NO_2^-$) and nitrate ($NO_3^-$) by specialized **nitrifying bacteria** (also called nitrifiers) and a few other organisms. Bacteria in the genera *Nitrosomonas* (which convert $NH_4^+$ to $NO_2^-$) and *Nitrobacter* (which oxidize $NO_2^-$ to $NO_3^-$) are the two most important nitrifiers. These bacteria are chemoautotrophs that obtain energy from the oxidation of $NH_4^+$ or $NO_2^-$ and fix carbon from atmospheric $CO_2$. They

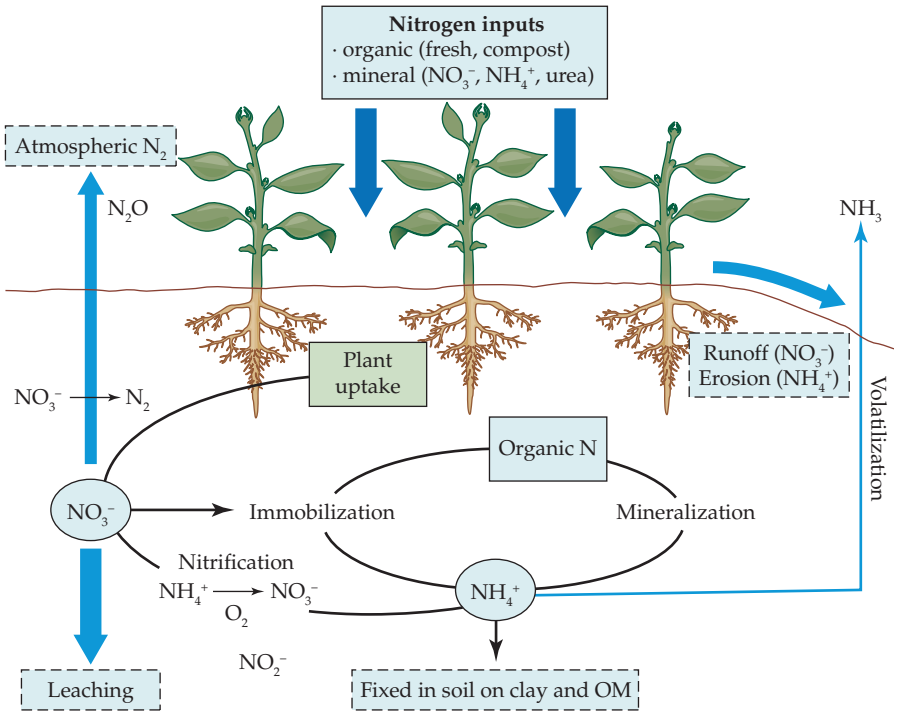



**Figure 5.19** The soil nitrogen cycle in a horticultural system, with the main transformations and pathways for loss. Dashed boxes indicate removal of N from the local ecosystem. (After M. D. R. Cameira and M. Mota. 2017. *Horticulturae* 3: 25.)

require oxygen (they are obligate aerobic organisms). The two groups of bacteria are generally found together in soils, so high levels of $NO_2^-$ do not often accumulate. The few other organisms that carry out nitrification belong to a variety of heterotrophic groups, including bacteria and fungi. They utilize $NH_4^+$ to produce $NO_3^-$ or $NO_2^-$, obtaining energy from the breakdown of carbon compounds in soil organic matter. These organisms may be particularly important in soils that are low in nitrogen.

The balance between $NO_3^-$ and $NH_4^+$ in soils differs among habitats, largely as a result of environmental effects on nitrification. While most plant species are better able to take up $NO_3^-$, plants adapted to environments where soil N is largely in the form of $NH_4^+$ (that is, where nitrification is inhibited, as in cold, waterlogged soils) may have a preference for that form. In most plants, both forms of nitrogen are converted in the roots to amino groups ($NH_2^-$), which are then attached to various organic compounds ($NH_4^+$ is toxic if it accumulates in plant tissues).

The conversion of $NH_4^+$ to $NH_2^-$ requires much less energy than conversion of $NO_3^-$ to amino groups. $NO_3^-$ is transported to plant roots readily by mass flow in the soil solution, but the lower mobility of $NH_4^+$ in the soil may require greater energy expenditure in the form of root proliferation (Schlesinger 1997). Nitrogen uptake as either $NO_3^-$ or $NH_4^+$ is aided by enzymes that expend energy to transport the ions across cell membranes in the roots. Active root growth and root proliferation into areas of the soil where N is more available also requires energy. In environments where N is less available, plants tend to allocate more biomass to roots, and root:shoot ratios are generally high, compromising the plant's capacity to capture carbon.

Uptake of positively charged ions, such as $NH_4^+$ and many other nutrients, such as $K^+$, from the soil could potentially lead to a charge imbalance in plant tissues. To counteract the positive charges, plant roots release hydrogen ions ($H^+$) into the soil solution, acidifying the soil. When plants take up $NO_3^-$, they release negatively charged ions such as organic acids and bicarbonate ($HCO_3^-$) to maintain their charge balance. The release of these ions in turn affects soil chemistry and the availability of nutrients to plant roots and soil microorganisms. Some ions (e.g., $Ca^{2+}$ and $Na^+$) are actively excluded by the roots of some plants in environments where their availability is too high.

In wetland soils, such as those of the salt marsh we have been considering (as well as freshwater marshes, bogs, fens, and others), frequent inundation of the soil with water results in anaerobic conditions in the soil (Figure 5.20). Because the nitrifying bacteria require oxygen to convert $NH_4^+$ to $NO_2^-$ and $NO_3^-$, little or $NO_2^-$ and $NO_3^-$ are produced in these anaerobic conditions, and wetland plants are specialists in being able to take up and use $NH_4^+$. Some unusual "carnivorous" plants living in nitrogen-poor bogs can obtain nitrogen from sources other than $NO_3^-$ and $NH_4^+$ in the soil by absorbing nitrogen from decomposing insects. *Sarracenia purpurea* (purple pitcher plant, Sarraceniaceae) obtains nitrogen from insects it captures and drowns in water in its pitcher-shaped leaves (Figure 5.21).

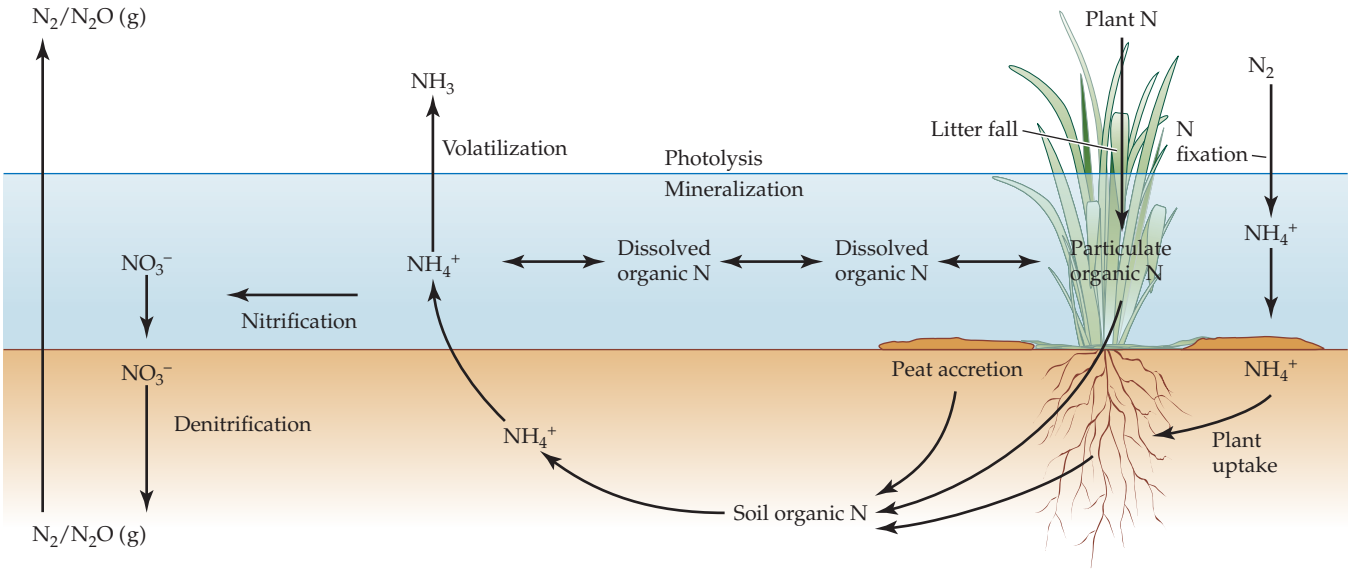



**Figure 5.20** Below ground processes in the nitrogen cycle in a wetland system. (After R. D. DeLaune and K. R. Reddy. 2008. *Biogeochemistry of Wetlands: Science and Applications*. CRC Press: Boca Raton, FL.)

(A) (B)

©Barry Rice/sarracenia.com

Rachel Korn/CC BY

**Figure 5.21** (A) *Sarracenia purpurea* (purple pitcher plant, Sarraceniaceae), a northern bog plant. Bogs are very low in certain nutrients, particularly nitrogen. (B) The pitcher plant absorbs substantial amounts of nitrogen from decomposing insects captured in its pitcher-shaped leaves. The insects fall into the "pitchers" but cannot crawl out, because of downward-pointing trichomes (hairs) on the leaf surface, and drown. If you look closely you can see the stiff downward facing trichomes; in the photo, a spider is waiting to snatch an insect before it falls into the "pitcher." Decomposition is carried out by a community of bacteria and invertebrates living in the water that collects in the "pitcher."

The $NH_4^+$ may also accumulate in the soil. If ammonification is also slowed, organic matter along with nitrogen in microbial tissues may also build up, in some soils leading to the formation of deep beds of organic matter. Vast reserves of organic matter in the form of peat are held in some wetland soils, containing very large amounts of both N and C. Harvesting peat for gardening and farming can disrupt wetland ecosystems. Anthropogenic release of methane from organic material stored in wetland soils threatens to become a major factor in accelerating climate change (see Chapter 16).

The most important immediate factor limiting nitrification rates is the amount of available $NH_4^+$. Nitrification is limited by anything that affects either the production of ammonium, or its immobilization (see below), as well as uptake by plants and capture of $NH_4^+$ on clay particles and organic matter. Nitrification rates are also affected by factors that control the population size and physiological activity of nitrifying bacteria, particularly soil water content, but also temperature, pH, and characteristics of the soil and the plant materials. Soil water content is important because very dry soils limit the growth and activity of the bacterial populations, while overly wet and waterlogged soils limit these aerobic microorganisms by excluding oxygen.

## Nitrogen is lost from ecosystems through denitrification and leaching

Unlike positively charged ammonium ions, negatively charged nitrate ions do not readily bind to soil minerals and organic matter, and they are therefore highly mobile in the soil solution. Nitrate ions in the soil solution can be transported some distance to soil roots both by diffusion and by mass flow. Because they are so mobile, they can also be easily leached from the soil into groundwater and runoff and thus are readily lost. (In some tropical soils, however, anions can be adsorbed onto positively charged aluminum or iron oxides and thus prevented from leaching.) Leaching of nitrate from fertilizer and animal waste can be an important source of water pollution.

Nitrate is also lost from ecosystems by the process of microbial denitrification, which is the reduction of $NO_2^-$ and $NO_3^-$ to various gaseous nitrogen compounds, including $N_2$ and nitrous oxide ($N_2O$). The latter is an important greenhouse gas that can contribute to global warming (see Chapter 16). The composition of the denitrifying bacterial community can differ strikingly between disturbed and undisturbed soils, and these differences can be reflected in very different responses to environmental conditions and in different rates of denitrification and production of $N_2O$ (Cavigelli and Robertson 2000).

Fire can be a major cause of nitrogen loss in some ecosystems. Most of the N released into the atmosphere by **volatilization** during a fire comes from the vegetation itself, but substantial amounts can come from the burning of accumulated litter, and very hot fires can burn the soil organic matter in the uppermost part of the profile, volatilizing nitrogen from the soil.

### Decomposition can immobilize soil nitrogen when $NO_3^-$ or $NH_4^+$ are sequestered in bacterial biomass

As soil organic matter is decomposed by microorganisms in the process of mineralization, much of the N they released is quickly taken up and reincorporated into the living microbial biomass (see Figure 5.18). The ratio of C to N in plant tissues ranges from about 25:1 to 150:1, but the ratios are much lower in fungi (between 4:1 and 15:1) and decomposer bacteria (3:1 to 5:1). As decomposition occurs, soil C:N (as well as C:P) ratios decline as the biomass in decomposing plant material is replaced by biomass in the living fungi and bacteria growing on it. When nutrients become immobilized by being incorporated into living soil microorganisms as organic molecules, these nutrients become less available for uptake by plants. This biological **immobilization** of nutrients by microorganisms not only makes the nutrients unavailable to plants, but alters their stoichiometry. These molecules are thus converted from soil organic matter back into organic form by microorganisms and can be recycled one or more times through them before being ultimately released to be taken up by plants, bound to soil minerals or soil organic matter, volatilized and lost, or oxidized.

Eventually the microorganisms die, and some of the nutrients that were immobilized again become available for uptake by plants. However, nitrogen immobilization by microorganisms may be a very important short-term limitation on plant nutrient uptake and growth. This effect can be particularly strong when the plant tissues that are decomposing have relatively high C:N ratios or are rich in plant materials (such as lignin) that resist decomposition. High C:N ratios and the presence of lignin and similar materials result in slower decomposition rates and a longer period of N and P immobilization. Plant species differ greatly in their ability to resorb nitrogen from leaves prior to leaf senescence (see Chapter 4). For instance, as a consequence of retaining their leaves for long periods of time and having low tissue nutrient concentrations, evergreen plants tend to have low nutrient loss rates, particularly for nitrogen. Evergreen leaves and leaves from plants native to low-nutrient environments typically decompose more slowly than those from deciduous plants and plants from fertile sites. The leaves of many *Quercus* (oak, Fagaceae) species, which are high in materials that physically and chemically resist decomposition, are incorporated into soil organic matter much more slowly than, for instance, *Acer* (maple, Sapindaceae) leaves from the same site. Leaves of nitrogen-fixing species decompose more quickly than those of co-occurring nonfixers, with consequently less immobilization of soil nutrients. Decomposing animal tissues have much lower C:N stoichiometric ratios, and tend to decompose much more rapidly than plant tissues.

## 5.5 Nitrogen Deposition and Acid Precipitation Can Alter Ecosystems

Two other related anthropogenic phenomena affecting plants are acid precipitation and nitrogen deposition. Like rising $CO_2$ levels, these anthropogenic changes in biogeochemical cycles affect plants, but they have effects mainly on a regional scale. However, they affect many different regions, with profound effects. Nitrogen pollution released to the atmosphere returns to the ground by both **wet deposition** (acid precipitation) and **dry deposition** (dryfall). Acid precipitation has been a cause for environmental concern since the mid-twentieth century, while nitrogen deposition has gained widespread attention only since the 1990s.

**Acid precipitation** refers to rainfall, snowfall, or fog that is more acidic than the historic average for an area. Normally, precipitation is slightly acidic because $CO_2$ dissolved in water droplets in the atmosphere forms carbonic acid. Acid precipitation is caused primarily by two common components of air pollution, sulfur dioxide ($SO_2$) and $NO_x$ (which includes nitric oxide [NO] and nitrogen dioxide [$NO_2$]), that are produced by industrial and automotive emissions. These gases undergo chemical reactions in the atmosphere to form sulfuric acid ($H_2SO_4$) and nitric acid ($HNO_3$), respectively. In solution, these acids produce highly acidified (very low pH) rain and snow.

Acid precipitation harms vegetation directly by damaging foliage and other plant parts. Indirect effects caused by acidification of the soil and surface waters have even greater negative consequences for plants. The release of ions harmful to plant growth in acidified soil may cause toxic reactions. More broadly, as soils become more acidic, leaching of cations can reduce nutrient availability or cause imbalances in the ratios of different nutrients, with detrimental effects for many plant species (see Chapter 4). Die-offs of *Acer saccharum* (sugar maple, Sapindaceae) in Quebec, Canada, have been attributed to deficiencies of soil K and Mg caused by acid rain. Forest declines in central Europe have been attributed to acid rain depleting levels of soil Mg relative to N.

Plants can also absorb $SO_2$ and $NO_2$ gases directly from the atmosphere. Damage to conifer foliage in Germany seems to be a result of direct uptake of anthropogenic $NO_2$ from the atmosphere, which results in leaching of $Mg^+$ and other cations within foliage tissues. $SO_2$ and $NO_2$ can both be deposited as dry deposition on leaves, soil, and other surfaces, where they may later react with rainfall to produce acids. High levels of $NH_4^+$ and $NO_3^-$ in fog may also directly harm trees.

Nitrogen occurs in suspension in the atmosphere in a wide variety of forms, including $NO_3^-$, $NH_4^+$, organic N, and gaseous forms ($NH_3$, $HNO_3$, or $NO_2$). Some of

these forms can enter terrestrial ecosystems and directly or indirectly be taken up by plants. The major forms of N deposition are wet deposition of nitrogen dissolved in rainwater, dry deposition of nitrogen in particles and dust, and fog or cloud deposition of nitrogen dissolved in water droplets onto plant surfaces (see Figure 5.18). Much of this nitrogen deposition results from human activities, often as $NO_x$ (NO and $NO_2$).

While recent efforts to curb air pollution have been quite successful in limiting $SO_2$, nitrogen-containing pollution has increased, becoming a greater problem. Nitrogen from anthropogenic sources enters the atmosphere in various forms, primarily $NO_x$, but also $N_2O_5$ (dinitrogen pentoxide), $HNO_3$, and $NH_3$. Human activities have doubled the rate at which nitrogen enters the terrestrial nitrogen cycle each year since the 1850s. Humans also add about 80 million metric tons of nitrogen per year to the biosphere by fixing atmospheric N for nitrogen fertilizer (Box 5C), which is a large source of the N found in runoff and groundwater (Figure 5.22). This additional nitrogen can have severe effects on ecosystems at local to regional scales (Vitousek et al. 1997). High rates of nitrogen deposition have disruptive effects on forests, streams, and coastal waters and estuaries (Aber 1992; Aber et al. 1998) and can lead to general forest decline and dramatic changes in stream ecology (Schlesinger 1997).

Anthropogenic nitrogen comes from a wide variety of sources: automobile exhaust, industry, power plants, and agricultural fertilizers and animal wastes. This broad range of sources is one of the main reasons why it is much more difficult to control nitrogen pollution than sulfur pollution. Sulfur is produced by far fewer kinds of sources, primarily the combustion of coal and other fossil fuels (particularly by power plants) and by particular industries. The largest contribution of anthropogenic N to the terrestrial nitrogen cycle comes from the industrial fixation of atmospheric N for agricultural fertilizer (currently $80 \times 10^{12}$ g/year). Eventually, substantial quantities of this N are volatilized or transferred as particulate matter (dust and other small particles) into the atmosphere as $NO_x$. The use of commercial N fertilizers is increasing rapidly: The amount of N in chemical fertilizers applied from 1980 to 1990, for instance, was at least as great as the amount applied before that in all of human history. Another environmental consequence of the enormous scale of the industrial fixation of nitrogen is the emission of large amounts of $CO_2$ from the fossil fuel combustion (primarily natural gas) necessary for the fixation process.

Volatilization of ammonium by domesticated cattle (from flatulence and decomposing manure) is a substantial source of anthropogenic N, particularly in some ecosystems but also globally. Nitrogen deposition by humans is spatially highly variable. Nitrogen from industrialized urban centers is carried by prevailing winds

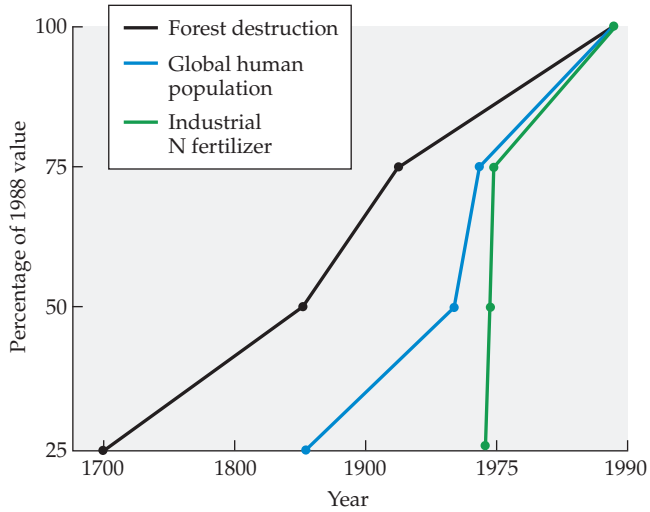



**Figure 5.22** Global changes in forest destruction, human population size, and industrial fertilizer production from 1700 to the late 1980s. The points show the years in which each variable reached 25%, 50%, and 75% of its value in 1990. Note that the time scale is nonlinear; the first two intervals represent 100 years, the next two intervals represent 25 years. Industrial production of nitrogen fertilizer has grown exponentially since the late 1940s. (After R. W. Kates et al. 1990. In *The Earth as Transformed by Human Action*, B. L. Turner et al. [Eds.], pp. 1–17. Cambridge University Press: Cambridge, UK.)

to nearby regions, where it is deposited. For example, much of the nitrogen deposited in the northeastern U.S. originates from coal-burning power plants in the upper midwestern states in the U.S.

Nitrogen deposition in the second decade of the 2000s was approximately 3 to 10 times natural levels in the areas with highest deposition: the northeastern United States, central Europe, China, and India. These numbers are lower than the highs of the 1990s, due in large part to laws limiting air pollution. Nitrogen is limiting to plant growth in many communities (see Chapter 4), so the extra nitrogen initially acts as a fertilizer, leading to increased growth. But anthropogenic nitrogen deposition can exceed the ability of the system to use the nitrogen, leading eventually to **nitrogen saturation** of the system. Its effects are nonlinear: Negative effects may not be manifested as nitrogen continues to accumulate until a critical threshold is reached. At that point, the effects of nitrogen saturation rapidly become detrimental to many plants, soil bacteria, and fungi. These effects include leaching of nutrients (such as calcium) from the soil as a result of increased acidity, loss of fine roots, steep declines in mycorrhizal fungi and consequent reductions in plant uptake of phosphorus, and increased leaching of N into ground and surface waters due to increased bacterial nitrification and reduced plant uptake.

## BOX 5C The Haber Process, the Green Revolution, and Nitrogen in Ecosystems

Nitrogen limits plant growth, including crop productivity. Farmers traditionally used manure from farm animals to supply additional nitrogen to soils; bat or bird feces (guano) were used to provide phosphorus and potassium in addition to nitrogen. The problem is that nitrogen in manure is present in low concentrations, and guano was available in limited supply and had to be imported to Europe and North America from remote islands or caves where it was more plentiful at great expense. In 1907, a German chemist, Fritz Haber, invented a relatively inexpensive and efficient (although energetically costly) method to fix atmospheric $N_2$ into $NH_3$, now called the Haber process. This process was scaled up to industrial production levels, winning Haber a Nobel Prize. The $NH_3$ could be readily converted to $NH_4NO_3$ (ammonium nitrate) and used for fertilizer, or made into gunpowder or explosives. During World War I Germany could not import guano to make explosives, because of military blockades. The ability of Germany to produce gunpowder and explosives using the Haber process almost certainly greatly extended the duration of the war. Haber also invented and oversaw the deployment of poison gas, ultimately killing hundreds of thousands in gruesome deaths.

After World War I and extending into the 1960s, the use of chemical fertilizers containing ammonium fixed by the Haber process led to a period of vastly increased agricultural productivity, in conjunction with the breeding of crops capable of using the increased nitrogen. Together with the introduction of industrial farming methods, herbicides, and pesticides, this Green Revolution resulted in enormous increases in yields, particularly of grains. This increase dramatically decreased the incidence of famines and without question saved billions of lives, or perhaps we should say, made billions of lives possible as the human population increased from under two billion at the end of World War I to almost eight billion people in 2020. Half of the nitrogen in your own body is likely to have been captured by the Haber process. The discovery of the Haber process has also had many other profound effects, some positive, some negative. Production shifted from family farmers using traditional agricultural practices to large corporate farms, particularly in North America. There was a great increase in toxic and polluting herbicides and pesticides in food and in the environment, with negative effects on people and on other organisms. Although the human population increased, starvation dramatically decreased worldwide. The large-scale use of chemical fertilizers has been a great contributor to nitrogen deposition and nitrogen runoff into streams and the ocean, with many negative consequences for the natural world. Fritz Haber's legacy is indeed a very mixed one.

Ecosystem responses to nitrogen deposition are uneven. Different systems vary greatly in their responses, depending on soils, plant species composition, climate, and other factors. Areas in the United States that show evidence of nitrogen saturation and consequent declines in forest ecosystems include the Catskill Mountains north of New York City, the Great Smoky Mountains in Tennessee, high-elevation forests in the eastern Colorado Rockies, and the San Gabriel and San Bernardino Mountains outside of Los Angeles. Many other regions globally show similar evidence of forest declines due to nitrogen saturation from nitrogen deposition.

Nitrogen deposition can also reduce plant tolerance of other stresses, including frost, periods of drought, excessive heat, and herbivory. Changes in insect herbivory patterns in forests and other plant community types can be sensitive indicators of nitrogen deposition effects. Changes in community composition, with the disappearance of sensitive species, have been attributed to the effects of nitrogen deposition in various parts of Europe. Mosses and lichens appear to be especially vulnerable to nitrogen deposition, often showing signs of stress first. In the United Kingdom, sensitive *Sphagnum* moss species have declined in the Pennine Hills due to the toxic effects of $NH_4^+$ and $NO_3^-$ deposition, and the losses of many species of mosses and lichens in Cumbria since the 1960s also appear to be a consequence of nitrogen deposition.

Changes in community composition and the losses of some species may also be an indirect result of changes in competitive interactions due to higher nitrogen input. Experiments in bogs at four sites in western Europe showed that increases in nitrogen deposition reduced the growth of the dominant *Sphagnum* moss species. This reduction was a result of shading by increased cover of taller vascular plants and led to changes in community composition (Berendse et al. 2001). Similarly, the widespread replacement of *Calluna vulgaris* (heather, Ericaceae) by dense stands of grass in heathlands in East Anglia in the United Kingdom and in other locations seems to be due to the indirect effects of nitrogen deposition on competitive relationships. Interestingly, grazing may reduce nitrogen-induced competitive replacements of other species by grasses in species-rich vegetation in the Netherlands (Kooijman and Smit 2001).

It is difficult to separate the effects of nitrogen deposition on species richness from the effects of other

environmental factors that have also changed over the same time period. Simply observing nitrogen deposition and species losses occurring together does not mean that nitrogen deposition necessarily caused the declines. However, experiments have provided evidence of a direct link between atmospheric nitrogen deposition and changes in species composition in heathlands and calcareous grasslands in the Netherlands. Rein Aerts and Frank Berendse (1988) used field experiments to demonstrate that increased nitrogen input in amounts similar to those of anthropogenic deposition led to declining cover of *Erica tetralix* (Ericaceae; a dwarf shrub, formerly dominant) and increasing cover of *Molinia caerulea* (Poaceae; a perennial grass replacing *Erica*) in wet heathlands. These changes in species composition have been attributed to the effects of increased nitrogen on competitive interactions in these naturally low-fertility habitats (Aerts et al. 1990; Aerts 1999). Together with other evidence, studies such as these have implicated atmospheric N deposition in changes in community composition and losses of plant biodiversity in the Netherlands, the United Kingdom, and elsewhere in western Europe.

## 5.6 Cycles of Phosphorus and Other Elements Play Important Roles in Ecosystems

We have examined the major ecosystem processes of two primary nutrients important to plants: carbon and nitrogen. Other nutrients, of course, are also necessary for plant growth (see Table 4.2). Here we briefly discuss the cycling of the other primary nutrient—phosphorus—and two other important elements—sulfur and calcium—because these macronutrients are ecologically important and because of the influence of human activities on their cycles. Then, we turn to perhaps the most crucial material of all, water.

### Microorganisms make phosphorus available for plants

Although phosphorus is needed for plant growth and function in smaller amounts than other macronutrients (see Table 4.2), it is fairly common for plant growth to be limited by P availability. We considered P availability in the soil and its uptake by plants in Chapter 4. Most of the phosphorus that cycles through ecosystems is found in living plants, microorganisms, decomposing plant material, and soil organic matter—that is, it is all found in organic material (Figure 5.23). The original source of most of the phosphorus that enters the biosphere is weathering of rock containing apatite minerals. The minerals in the apatite group are the only minerals common in the lithosphere that contain a substantial amount of phosphorus. As rock weathers and soils form, and as the soils develop and age, much of the phosphorus becomes bound in the interior of iron and aluminum oxide crystals, from which living organisms cannot remove it. In old soils that have become dominated by iron and aluminum oxides (such as those that are common in Africa and Australia, as well as in tropical regions elsewhere), the lack of available phosphorus is one of the most important environmental factors affecting plants. Inorganic P may also be held in more available forms on the surface of other soil minerals by anion adsorption and other reactions. The complicated soil chemistry involved in these reactions depends on soil pH and other factors. It is therefore generally difficult

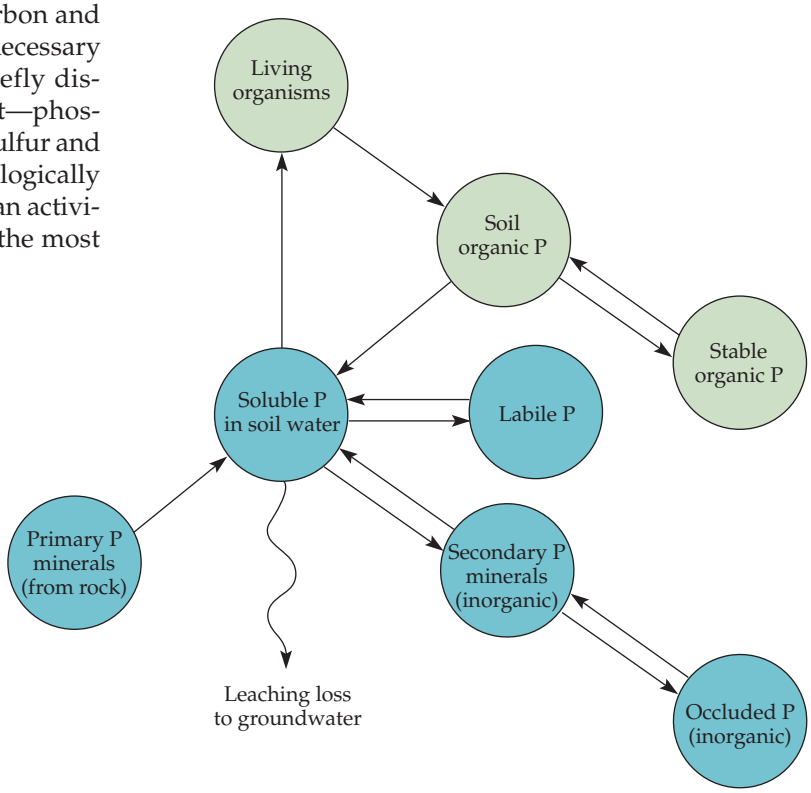



**Figure 5.23** Phosphorus from minerals in rock (primary P) dissolves in the soil solution, from which it can be absorbed by plants and other living organisms. Some of this soluble soil phosphorus is transformed into various inorganic forms or is leached. Much of the soluble soil phosphorus is recycled many times through living organisms; therefore, the vast majority of the phosphorus taken up by plants does not come directly from rock, but has been recycled through other living organisms. Biological components are shown in green, geochemical components in blue. (After N. E. Smeck. 1985. *Geoderma* 36: 185–199.)

for scientists to accurately measure the amount of phosphorus available to plants.

Unlike the global cycles of carbon, nitrogen, and water, the global cycle of phosphorus does not have a major atmospheric component (Figure 5.24), although some phosphorus is transported through the atmosphere in dust. Most of the phosphorus that is used by organisms has been recycled in organic form through other living organisms, although its ultimate source is rock weathering. Organically bound phosphorus is released in the form of $PO_4^{3-}$ (orthophosphate) by plant roots, fungi (including those in mycorrhizae as well as free-living fungi), bacteria, and algae. All of these organisms mineralize phosphorus through the action of extracellular phosphatases (enzymes that cleave ester bonds). This process is expensive, both energetically and in terms of the nitrogen invested in the phosphatase enzymes. Microorganisms as well as plants can take up the $PO_4^{3-}$ that is released, but if it is not immediately taken up, the orthophosphate does not remain available for long in the soil. Rather, it readily binds to organic particles or to soil minerals, becoming partially or completely unavailable to plants.

## Sulfur is critical for certain plant compounds

Sulfur is necessary primarily for building the amino acids cysteine and methionine, important constituents of various proteins. Like nitrogen, phosphorus, and carbon, sulfur is cycled biogeochemically in ecosystems. The weathering of rock with minerals containing sulfur (such as pyrite and gypsum) is one major source of the sulfur that ultimately becomes available to plants. Atmospheric deposition is also an important source of sulfur. A large proportion of the sulfur pool in soils exists as a part of soil organic matter. Inorganic sulfate ions ($SO_4^{2-}$) are adsorbed onto soil mineral particles, in equilibrium with $SO_4^{2-}$ dissolved in the soil solution. Bacterial mineralization and immobilization of sulfur are analogous to the same processes for nitrogen and phosphorus. Sulfur is lost from ecosystems in a variety of forms. In addition to the loss of sulfate ions from ecosystems by leaching, many plants release volatile organic sulfur (particularly hydrogen sulfide, $H_2S$), and waterlogged, anaerobic soils can also produce large quantities of sulfur gases by bacterial reduction and release it to the atmosphere.

On a global scale, there is a large pool of sulfur dissolved in seawater. In the atmosphere, sulfur-containing gases are short-lived and are found at low concentrations. However, dust storms, sea spray, and volcanic eruptions can contribute large amounts of particulate sulfur (as well as some sulfur in the form of gases and aerosols) to the atmosphere. Human activities are the largest current source of atmospheric sulfur dioxide gas ($SO_2$). The $SO_2$ that enters the atmosphere reacts with water to form sulfuric acid, which, along with nitric acid (from anthropogenic NO and $NO_2$), is the chief cause of acid precipitation. Today, stricter regulation of power plants and the replacement of coal by natural gas in power plants has greatly reduced $SO_2$ deposition in the U.S.

## Calcium is necessary for many plant processes and structures

Calcium is abundant in the Earth's crust. It is an important component of limestone and other common rocks such as gypsum ($CaSO_4 \bullet H_2O$) and calcium carbonate ($CaCO_3$). Weathering of these rocks releases calcium ions ($Ca^{2+}$), which are carried by rivers to the oceans; large pools of calcium ions are found in seawater. Marine invertebrates use this calcium to build calcium carbonate

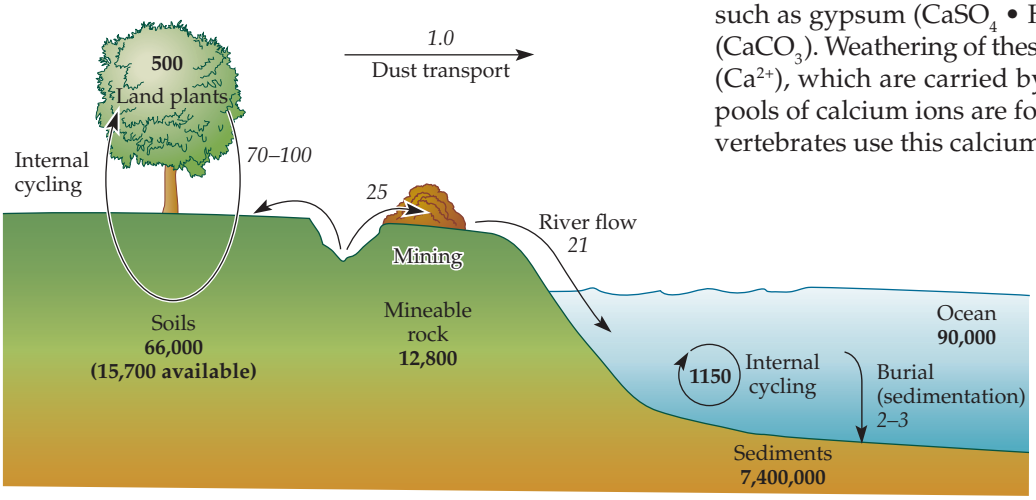



**Figure 5.24** The global phosphorus cycle, showing major fluxes (in $10^{12}$ g P/year, shown in italic type) and pools (in $10^{12}$ g P, shown in bold type). Marine sediments are the largest (although not active) pool, soils are the largest terrestrial and the largest active pool, internal ocean cycling is the largest global flux, and cycling by land plants is the largest terrestrial flux. (After W. H. Schlesinger. 2013. *Biogeochemistry: An Analysis of Global Change*, 3rd ed. Academic Press: New York.)

shells; when these organisms die, their remains form thick sediments. Through burial, metamorphosis, and crustal movement, these marine sediments may ultimately be transformed into calcium-bearing terrestrial rock. Calcium ions readily dissolve in the soil solution and are taken up by plants passively in the transpiration stream. Calcium concentrations can vary dramatically among soils, from being apparently limiting to plant growth to being present in excessive amounts. In humid regions with low-pH soils, calcium ions can be heavily leached from the soil, while in contrast, in arid regions calcium can accumulate in the soil, particularly in caliche horizons (see Chapter 4).

Within an ecosystem, the calcium contained in plant biomass is returned to the soil through litterfall and subsequent litter decomposition, as well as by **throughflow**—rainwater that passes through the canopy and travels down the surface of tree trunks, carrying dissolved ions leached from the plant surfaces. Some calcium is carried in particulate form in the atmosphere and deposited in rainfall or by dry deposition (Figure 5.25). Unlike nitrogen, which is resorbed before leaf senescence (see Chapter 4), calcium is largely lost when leaves fall. Calcium cannot easily be resorbed, because much of it is a structural component of cell walls. Therefore, woody plants need to replace most of the calcium lost during leaf senescence each year with calcium that has been recycled in the soil by decomposition.

The requirement of plants for calcium is generally very high, second only to that for nitrogen (see Table 4.2). Calcium is unusual in that it plays both chemical and structural roles in plant tissues, including regulation of growth and response to stress, stomatal function, cell division and cell wall synthesis, and structural support in both leaves and wood (McLaughlin and Wimmer 1999). Decreasing calcium availability may lead to decreased growth and higher mortality in forest trees—for example, by making them more susceptible to pathogens.

Evidence suggests that calcium depletion has occurred in many forests, with negative consequences for forest vitality (Likens et al. 1998), although other researchers have disputed these findings (Yanai et al. 1999). The potential for calcium depletion is a result of several factors, including decreasing amounts of calcium in rainfall and removal of calcium from the ecosystem in biomass (by logging and other human activities). Most important are the direct negative effects of anthropogenic nitrogen deposition (particularly in the form of $HNO_3$), which results in displacement of calcium ions in the soil, causing them to be leached into stream water and lost. Thomas Huntington (2000) studied fluxes of calcium in southeastern U.S. forests and found that losses of calcium caused by logging and soil leaching often were greater than inputs from atmospheric deposition and rock weathering, leading to severe calcium depletion in the forest soils. His results suggest that calcium depletion could result in widespread problems for southeastern forest trees currently and into the future.

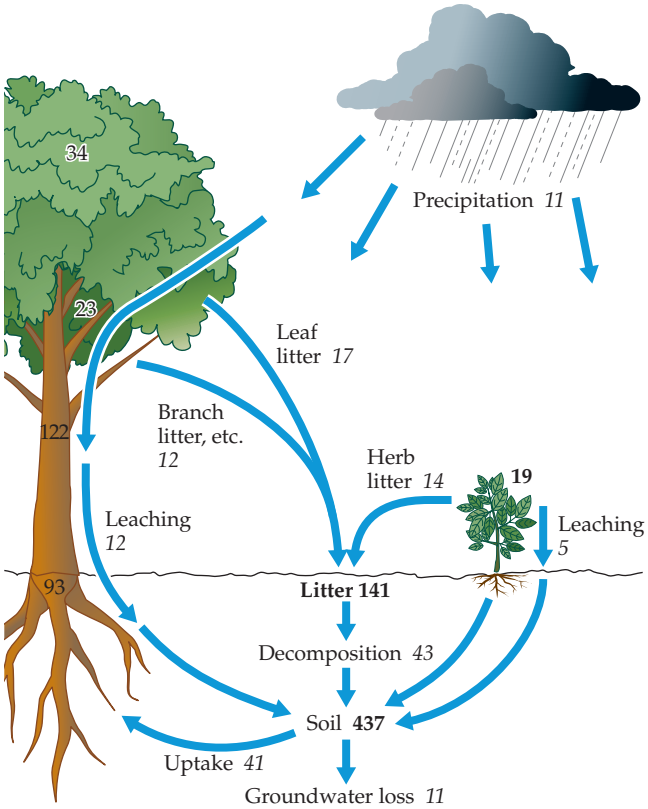



**Figure 5.25** The local calcium cycle within a forest ecosystem in the United Kingdom. Pool (bold type) amounts are given in kg/ha, and annual fluxes (arrows, in italic type) are shown in kg/ha/year. (After R. H. Whittaker. 1970. *Communities and Ecosystems*. Macmillan: New York.)

## 5.7 The Water Cycle Is Central to Life and Climate

### Water cycles at local and regional scales

As we have seen earlier in this book, at the individual plant level, water moves from the soil to the roots, up through the xylem and out through the leaves via the soil-plant-atmosphere continuum, regulated by the gradient in water potential (see Chapter 3). Soil nutrients also enter the plant carried in the soil water (see Chapter 4). We remind readers that water is the only substance that occurs naturally on Earth's surface in all three phases—liquid, solid, and vapor. The transformations between these states are important for the transfer of materials and energy in ecosystems—they include

evaporation and transpiration (change from a liquid to a vapor), condensation (change from a vapor to a liquid), sublimation (direct change from a solid to a vapor), and deposition (direct change from a vapor to a crystalized solid as frost; note that since this has nothing to do with anthropogenic deposition of substances like nitrogen we suggest the less ambiguous term fristilation).

The flux of water through plants is also very important regionally and globally. Plants can contribute a large share of the water in the atmosphere and in regional precipitation, but the effects of plants on local precipitation patterns vary greatly. Precipitation in California, for instance, comes largely from water evaporated over the Pacific Ocean. In contrast, between one-fourth and one-half of the rainfall over the Amazon basin comes from evapotranspiration from the Amazon forest itself, with the rest coming from outside the region (Salanti and Vose 1984; Eltahir and Bras 1994). Cutting large areas of the Amazon forest can therefore result in decreased evapotranspiration, decreased rainfall, and increased temperatures at the ground surface (Lean and Warrilow 1989; Shukla et al. 1990). More recent work has corroborated these earlier findings: Deforestation is likely to reduce rainfall in the Amazon and affect the water cycle regionally. This is predicted to have negative consequences for people locally and in the broader region, highlighting the connections between the forests, rainfall, and people in the Amazon (Marengo et al. 2018).

Plants can also indirectly affect water fluxes by intercepting precipitation, reducing the impact of rain hitting the ground, and decreasing runoff from the soil. If vegetation is removed from an area—for example, by clear-cutting a forest—runoff and soil erosion can be greatly increased, particularly in areas where the terrain is hilly or mountainous. This can sometimes result in flooding and mudslides, further damaging vegetation and threatening human lives and homes. Even in semiarid regions, removal of vegetation can result in reduced precipitation, increased soil warming, and the onset of desertification (Schlesinger et al. 1990; Dirmeyer and Shukla 1996).

**Potential evapotranspiration** (**PET**) is the maximum amount of water that would be lost to evapotranspiration

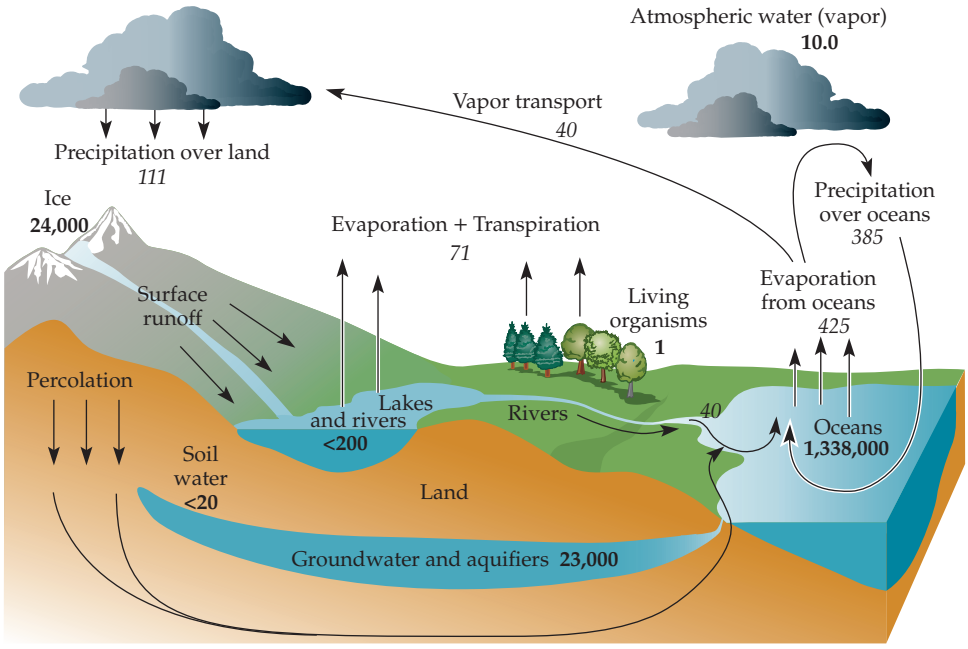



**Figure 5.26** The global water cycle. The numbers show the pools (in units of 1000 km$^3$ of water, shown in bold type) held in various major components of the global ecosystem and the fluxes among those components (in units of 1000 km$^3$/year, shown in italic type). Notice that the amount of water leaving the land to return to the oceans (40,000 km$^3$/year) is equal to the amount of water that returns to the land as water vapor evaporated from the oceans. The numbers given are approximate, as estimates from different authors vary considerably. (After W. H. Schlesinger. 1997. *Biogeochemistry: An Analysis of Global Change*, 2nd ed. Academic Press: New York; P. H. Gleick. 1996. In *Encyclopedia of Climate and Weather*, S. H. Schneider [Ed.], pp. 817–823. Oxford University Press: New York; data from M. I. Lvovitch. 1973. *Eos* 54: 28–42; M. T. Chahine. 1992. *Nature* 359: 373–380.)

in a particular place if water were freely available in the soil and plant cover were 100%. PET depends on the energy available to evaporate water, as measured by the ambient temperature. **Actual evapotranspiration** (**AET**) is equal to the amount of water that enters the system in precipitation minus the amount that is lost in runoff and percolation to groundwater (the amount of water stored in living systems is minimal relative to these amounts). You may be surprised to learn that AET can be much greater per unit area than the amount of water that would be lost by evaporation from an open body of water. This is because the surface area of the leaves from which water is evaporating is often considerably greater than the surface area of the ground—that is, the **leaf area index** (**LAI**) is often greater than 1.0. (Similarly the surface area of the roots taking up water is much greater than the surface area of the ground in which the plants are rooted.) In areas with long dry periods, PET is much greater than actual evapotranspiration, but in tropical rainforests PET and AET are approximately equal. Actual evapotranspiration is closely linked to other major ecosystem processes, particularly productivity, decomposition, and other soil processes (Schlesinger 1997). This linkage is not surprising, given the dependence of plant and microbial activity on soil water content and the close functional linkage between carbon assimilation and water loss in plants (see Chapter 2 and Chapter 3).

**TABLE 5.2** The relative importance of pathways leading to the loss of water in a variety of terrestrial ecosystems

| Biome | Evaporation (%) | Transpiration (%) | Runoff and groundwater recharge (%) |
|---|---|---|---|
| Tropical rainforest | 25.6 | 48.5 | 25.9 |
| Tropical rainforest | 10 | 40 | 50 |
| Tropical rainforest | 11 | 56 | 32 |
| Temperate forest | 13 | 32 | 53 |
| Temperate grassland | 35 | 65 | 0 |
| Temperate grassland | 33 | 67 | 0 |
| Temperate grassland | 55 | 45 | 0 |
| Temperate grassland | 56 | 34 | 10 |
| Desert | 28 | 72 | 0 |
| Desert | 20 | 80 | — |
| Desert | 73 | 27 | — |
| Desert | 65 | 35 | — |

*Source*: W. H. Schlesinger. 1997. *Biogeochemistry: An Analysis of Global Change.* Academic Press: New York, based on miscellaneous data sources.

## Local water cycles can affect global cycles

Terrestrial plants are the only living things (except humans) to have a substantial effect on the global water cycle (Figure 5.26). Most (96.5%) of the world's available water (that is, excluding water that most organisms cannot reach, such as that in deep aquifers) is stored, not surprisingly, in the oceans, with all of the fresh water in the world's rivers, lakes, ice, atmosphere, and groundwater constituting the remaining few percent. Living things store and metabolize relatively minuscule amounts of water when considered in the context of the global water cycle, although those minuscule amounts of water are essential to life. However, the amount of water that moves through plants by transpiration is large and globally important. The energy that drives **evapotranspiration** (evaporation plus transpiration) comes from solar radiation. The water cycle is very different from the cycles of mineral nutrients in that most of the water remains chemically unaltered as it passes through the biotic and abiotic components of the biosphere, with only its physical phase (liquid, gas, or solid) changing.

In intact vegetation, the balance between precipitation, evapotranspiration, and runoff varies greatly among biomes (Table 5.2). In deserts and grasslands, almost all of the water that enters the system as precipitation is lost by evaporation or transpiration, leaving little for runoff and recharge of groundwater. Both temperate and tropical forests generally lose less water by evapotranspiration than they receive as precipitation, so considerable amounts often run off into rivers and streams and percolate into the groundwater. Over the Earth as a whole, rivers carry about a third of the precipitation that falls over land back to the oceans. While the amount of rainfall over the oceans is greater than that over land (see Figure 5.26), the evaporation of water from the oceans is also much greater than the evapotranspiration of water over land. The net movement of water vapor from the oceans to the land through the atmosphere is balanced by the runoff of water from the land to the oceans in the form of groundwater and river flow. While only a minuscule amount of water is stored as water vapor in the atmosphere, the flux of water vapor continuously through the atmosphere is enormous.

## Summary

- Ecosystems are composed of the living organisms and the nonliving components of their environment with which they interact, together forming a single system.
- The fluxes of energy and materials in ecosystems and the magnitude and nature of the pools in which materials are stored form the basis of ecosystem ecology. These fluxes and pools make up ecosystem cycles.
- Different elements cycle through the ecosystem at very different rates, and they move through different parts of the ecosystem, including plants, the atmosphere, and the soil and soil microorganisms. Living organisms play a large role in the biogeochemical cycling of some elements, such as nitrogen and carbon, and have little or no role in the cycling of other elements, such as sodium and aluminum.
- Gross primary production (GPP) is the total energy (or carbon) fixed by producers in an ecosystem, while net primary production (NPP) is GPP minus the losses to respiration by primary producers. NPP varies enormously over the Earth's surface.
- NPP in terrestrial ecosystems is determined by climate but is also affected by many other biotic and abiotic factors.
- Decomposition transforms dead organisms, shed parts, root exudates, and excreted materials into soil organic matter and, ultimately, into inorganic nutrients, $CO_2$, water, and energy.
- The form in which nitrogen can be taken up by plants is an important aspect of the nitrogen cycle and depends strongly on microbial activity.
- Organisms have relatively fixed stoichiometries; differences in stoichiometry among plants, microorganisms, herbivores, and the physical environment interact with nutrient turnover times to regulate the dynamics of ecosystem and global nutrient cycles.
- Terrestrial plants and humans are the only living organisms that have large effects on the global water cycle. Living things contain and metabolize very small amounts of water in their cells, but large amounts of water move through plants in transpiration.
- Improving our understanding of biogeochemical cycles is critical both for understanding how plants and plant communities change over time and for understanding the environmental challenges presented by anthropogenic change.

# PART II

# From Individuals to Populations

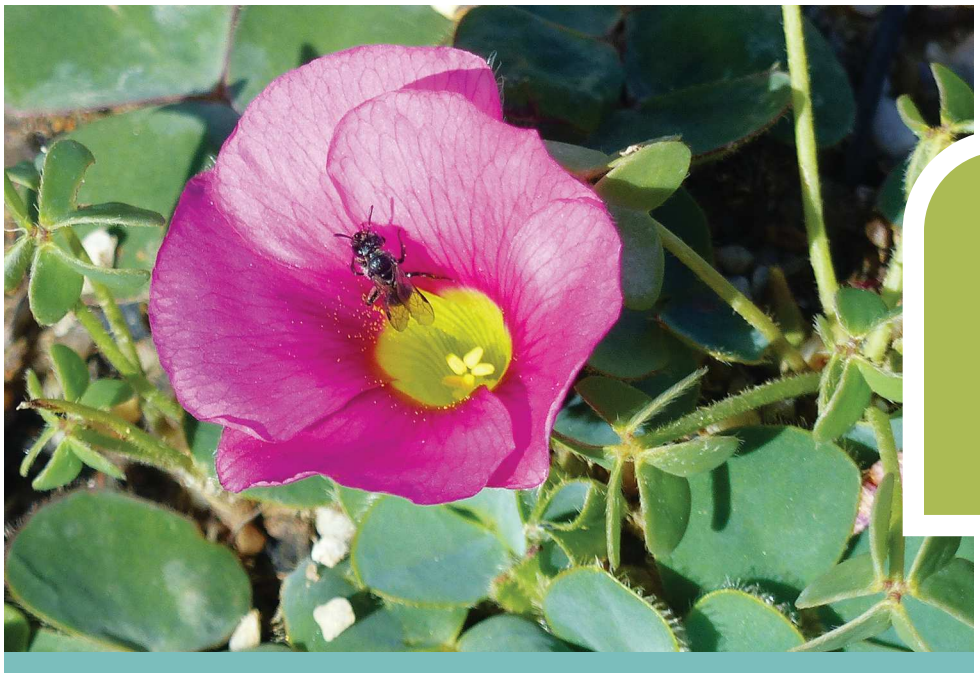
Courtesy of G. Fox


**6** Individual Growth and Reproduction 143

**7** Plant Life Histories 177

**8** Population Structure, Growth, and Decline 199

**9** Evolution: Processes and Change 231



Courtesy of G. Fox

6

# Individual Growth and Reproduction

The growth and reproduction of individuals are critical components of the growth (or decline) of populations. Growth and reproduction are also key ways in which plants interact with other plants, animals, and the physical environment around them through uptake of nutrients, use of space, pollination relationships, and the dispersal of fruits and seeds. The movements of pollen, seeds, and fruits are the main factors responsible for initiating spatial patterns in plant populations. The evolution of specific reproductive traits marks key episodes in the history of terrestrial plants. How do these traits evolve under the influence of natural selection? What are the ecological consequences of variation in these traits in terms of population spread, speciation, and the ability of plant species to invade new environments? Do genetically engineered crops present a danger to their wild relatives? We begin this chapter at the within-plant level of anatomy and physiology, but then we make the transition to considering growth and reproduction at the level of the whole plant and of plant populations.

Above: *Oxalis purpurea* (Oxalidaceae), near Stellenbosch, Western Cape, South Africa.

## 6.1 Growth Begins with Seed Germination

Seeds are dry, with slow metabolic rates. They may be dormant, or prepared to germinate as soon as they are produced. Dormancy may have lasted for weeks, or years, but unless the seed dies, eventually dormancy is ended and the seed becomes primed to germinate. As environmental and internal factors trigger germination (see Chapter 7), the seed imbibes water, it swells, and respiration increases as it becomes more metabolically active. The seed coat softens, and the **radicle**—the structure that will develop into the root structure—emerges first and begins to expand into the soil and take up more water. In many seed plants, the cotyledons emerge next. They are initially white, but as they are exposed to light, chloroplasts mature and the small plant becomes photosynthetic. The apical meristem begins to differentiate tissues behind it as it elongates the stem and starts to produce the first true leaves.The meristem behind the root tip continues to produce cells that differentiate into the expanding root system.

## 6.2 Plants Grow by Adding Repeated Units to Their Bodies

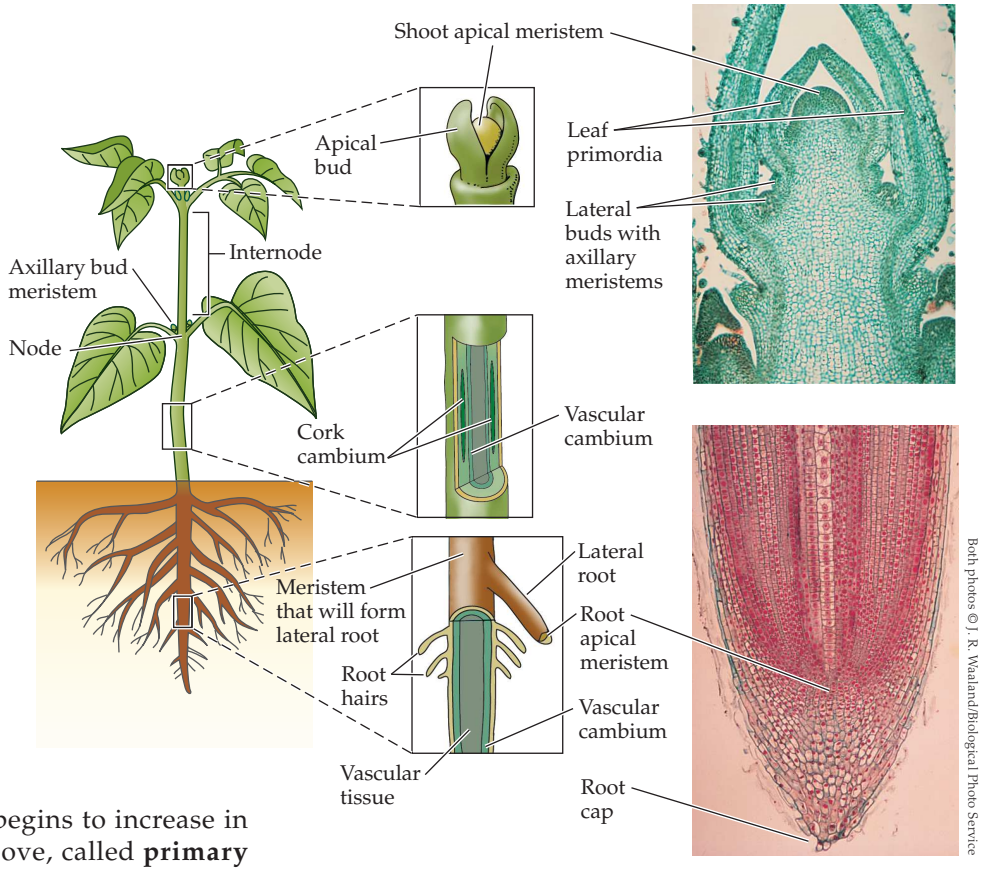
The growth of plants is **modular**; that is, plants grow by adding repeated units, or modules, to their bodies. We can compare growing plants to children's construction toys in which complex structures are built up by adding simple, repeated units to the original structure. As with such toys, there are only a few types of repeated units in modular plant growth; all plant structures are iterations of these units.

The basic rules of plant growth have important consequences for the ecology of plants. In particular, the ability to add (or lose) individual modules means that plants have a great deal of plasticity in their size and shape. This plasticity affects their ability to respond to damage, to capture resources, and to interact with one another. Many plants can survive the loss of large portions of their bodies to herbivores. Commonplace as this may seem, it is remarkable: Few animals could survive a similar loss of body parts! Similarly, many plants regularly lose parts of the **shoot**—the aboveground part of a plant, consisting of stems and leaves—to abiotic factors such as freezing temperatures, drought, and fire—and nevertheless survive.

The addition of new modules depends on the activity of **meristems**, which are collections of undifferentiated cells (cells that do not yet have specialized functions) (Figure 6.1). At the tip of a growing stem or branch is an **apical meristem**. The growing stem is created by the production of new cells behind the apical meristem. Eventually, these new cells differentiate into the many different kinds of cells making up the shoot, such as epidermal, xylem, phloem, and parenchyma cells. As the stem elongates, some cells originating in the apical meristem periodically generate a **node**—a point of cell proliferation where leaves or flowers will develop. Nodes, therefore, are the places where leaves and flowers are attached to stems. In many plants, the stem is visibly swollen at each node, and a primordium (bud) forms there for each potential new leaf or flower. The stem segments between nodes are called **internodes**. When the apical meristem generates a node, it may create groups of cells called **intercalary meristems**. The internodes may continue to expand, lengthening the shoot, if the intercalary meristems are activated. Some plants have short or even unnoticeable internodes. These plants, called **rosette plants**, include cabbage, dandelions, many alpine plants, and many low-growing weedy herbaceous plants (see Figure 7.7). In many of these species, the onset of flowering activates the intercalary meristems, and the plant "bolts," or lengthens, due to the expansion of the internodes. In other rosette species, bolting occurs when the apical meristem produces a lengthened stem above the rosette and then differentiates into a terminal inflorescence. The intercalary meristems also have other important functions. One of these occurs in grasses, where intercalary meristems at the plant base make it possible for the plant to regrow after fire, or after the leaves have been eaten by a grazer.

At each node there is also an **axillary meristem** in the **leaf axil**, the place where the leaf and stem join. The axillary meristem produces an axillary bud. The axillary meristems are often inactive while the apical meristem is intact, because the apical meristem produces hormones, especially auxins, that inhibit the proliferation of cells in the axillary meristems; this is called **apical dominance**. If the apical meristem is removed—for instance, when an herbivore chomps off the tip of the plant—the axillary buds are freed from apical dominance and they can begin to create new stems, leaves, or flowers. This basic modular growth response to damage is a major way in which plants recover from herbivory and continue to grow, although now in a lower, bushier form.

Flowering is initiated by a complex set of hormonal signals—which are not yet fully understood—in response to both the developmental stage of the plant and environmental signals, including changes in daylength and in resources. When this happens, the apical meristem ceases stem elongation and intiates the development of an **inflorescence**, the floral structures of the plant. At that point the apical meristem becomes dedicated to this process and ceases to exist as a meristem. Likewise, the axillary buds may develop into




Both photos © J. R. Waaland/Biological Photo Service

**Figure 6.1** The basic structures involved in plant growth. In the shoot, the apical meristem produces a bud—a miniature new shoot in the leaf axil that contains an axillary meristem and the primordial cells for either a new branch or lateral organs (such as leaves and flowers). The points at which these lateral organs, or buds, are attached are called nodes. The leaf primordia expand to become leaves. Stems increase in girth by division of the vascular cambium; this is called secondary growth. (After W. K. Purves et al. 1995. *Life, The Science of Biology, 4th ed.* Oxford University Press/Sinauer: Sunderland, MA.)

inflorescences and then will cease being meristematic and producing other structures.

A young plant or new shoot begins to increase in size by the process described above, called **primary growth**. All perennial gymnosperms and most perennial angiosperm dicots (and many annuals in these groups) continue their growth by another process, called **secondary growth**. Secondary growth is largely a process of increasing girth (thickness) by producing woody tissues (secondary xylem). In a new stem, the vascular bundles are separated from one another, but as the stem matures, the vascular bundles grow together and coalesce to form a sheath surrounding the inner part of the stem. In the center of this vascular tissue, between the xylem and the phloem, a new meristem, the **vascular cambium**, is formed just inside the stem, under the phloem and epidermis. The vascular cambium forms a thin sheath that extends for much of the length of the stem. A similar process occurs in roots. Secondary xylem, or **wood**, is typically much stronger and better protected than primary xylem and is a major source of the structural strength that enables perennial plants to survive for more than a short time.

In monocots, increase in girth is more complicated. Palms that grow as trees, for example, have no vascular cambium, and the growth in a palm stem is concentrated in a single very large bud (the primary thickening meristem) that widens early in life and then begins elongating, so the palm stem grows as a cylinder. Early in life they also add numerous **adventitious roots** (roots that originate from the stem), strengthening the main stem and further increasing in girth. Other tree-like monocots have other mechanisms to increase in girth. *Pandanus* (screwpines, Pandanaceae), which are especially common on tropical and subtropical islands, rely on extensive systems of adventitious roots (Figure 6.2A). Some of the very large succulents like *Dracena* (Asparagaceae) (Figure 6.2B) and *Aloe* (Asphodelaceae) (Figure 6.2C) are able to add entirely new vascular bundles including both xylem and phloem around the edges of their stems.

Root apical meristems control root growth. The apical meristem of a growing root is located behind, and protected by, a cap of loosely held cells that sloughs off as the growing tip of the root pushes through the abrasive soil. Roots do not have systems of nodes, internodes, or most of the different meristems characteristic of shoots. Roots branch when special differentiated cells resume active division, and each branch forms its own root apical meristem. One important feature of roots is their very fine root hairs. Most uptake of water and minerals occurs through these hairs (or through mycorrhizal structures; see Chapter 4), but each hair usually functions for only a few days. Thus, continued growth of roots is often important for the uptake of water and nutrients.

(A)

Asit K. Ghosh Thaumaturgist/CC BY-SA 3.0

(B)

Esculapio/CC BY-SA 3.0

(C)

Courtesy of G. Fox

**Figure 6.2** Increase in girth in treelike monocots can occur in several ways. (A) *Pandanus* species (screwpines, Pandanaceae), which are very common in tropical and subtropical islands, have extensive systems of adventitious (prop) roots. (B) *Dracena draco* (dragon tree, Asparagaceae) from the Canary Islands and (C) aloes like *Aloidendrum dichotomum* (Asphodelaceae) from South Africa add new vascular bundles including both xylem and phloem around the edges of their stems.

## 6.3 Plant Growth Affects Resource Acquisition

### Shoot architecture determines light interception

Plants have a great variety of **architectures** (arrangements of regenerating parts) and shapes. The growth forms of plants can reflect their environments in at least three ways. First, plant growth form is in part a result of evolutionary adaptations to the timing of unfavorable seasons (generally cold or drought). A widely used classification of perennial growth forms is shown in Figure 6.3. Each growth form has different consequences for survival and resource retention. An excellent example of the evolution of a diversity of growth forms is the morning glory genus, *Ipomoea* (Convolvulaceae). Some Mexican and Central American species in this genus are drought-deciduous trees (that is, phanerophytes), while others are vines with succulent root crowns in which the leaves and stems die back during the dry season (that is, both cryptophytes and hemicryptophytes). Farther north, in the United States, *Ipomoea* species are annual vines and other herbaceous forms.

Variation in plant shape can also be a result of phenotypic plasticity (see Chapter 9) in response to environmental conditions. Plants growing under poor resource conditions often have a different shape than plants of the same species growing in richer microhabitats. In some species, this variation reflects evolutionary adaptation to variation in the environment; in other species, however, the fact that plants have different shapes under poor conditions is a simple reflection of the stress induced by those conditions. Not all plasticity is adaptive.

Some forms of plasticity are adaptations that enable individuals to acquire more resources, disperse more effectively, or improve their competitive ability. One of the best-studied examples of such plasticity is tree architecture. Trees are plants that have invested most of their biomass in wood, which allows vertical growth, rather

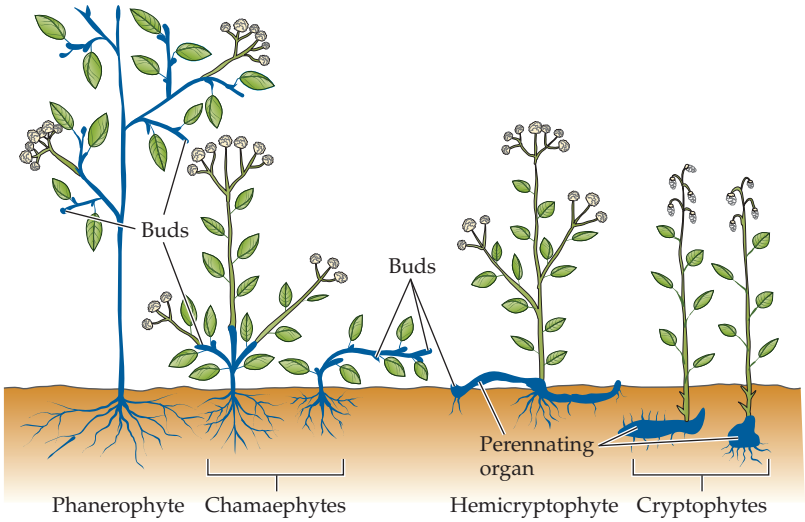



**Figure 6.3** The Raunkiaer system of perennial growth form classification is based on the position of buds or regenerating parts of a plant. Perennating tissues are shown in blue; deciduous tissues are green. The four categories are phanerophytes, trees or tall shrubs with buds more than 25 cm above the ground; chamaephytes, shrubs with buds less than 25 cm above the ground; hemicryptophytes, perennial herbs with buds at the ground surface; and cryptophytes, perennial herbs with perennating organs below the ground surface.

than in leaves or reproduction. The principal selective advantage of producing a tall trunk seems to be competitive superiority for light. In many forest trees, shaded individuals have a characteristic shape that is different from that of the same species when it reaches the canopy, or when it grows in a gap or along an edge. Many species have patterns of branching and growth that vary in response to light availability, and light interception is often a key factor affecting plant shape. Forest trees typically shed lower branches that are so shaded that their leaves fall below the light compensation point for long periods of time, so the tree as a whole experiences higher photosynthetic carbon gain than respiratory loss over the year (see Chapter 2). In some cases, the same species retain their lower branches when grown in the open, because those branches have positive net carbon balances under those circumstances. Most species have a characteristic pattern of **self-pruning**—shedding of branches—that tends to keep the net carbon uptake of a branch positive.

Self-pruning, however, is not necessarily just a passive response to shading; it can also be an important part of plant strategies for competition. When plants grow in dense stands (as in forests), selection often favors allocation of resources to upper branches rather than to lower branches because the upper branches not only gain more light, but also can overtop neighbors and deprive them of light. Such competition for light appears to have driven the evolution of trees. In some trees in fire-prone habitat, such as longleaf pine (*Pinus palustris*, Pinaceae) in the southeastern United States, self-pruning of lower brances may also serve to reduce the chances that flames from ground fires will spread into the tree crowns (the upper parts of the canopy).

The shortest-lived plants are nonwoody, and the longest-lived plants are woody. Woody plants are associated with greater longevity than herbs, for well-understood reasons—the xylem cells that make up wood, and other woody tissues such as fibers impregnated with very tough materials such as lignin, resist attack by herbivores and pathogens and are much more resistant to decay by fungi. However, many other factors also affect longevity. As a result, there are some very long-lived herbaceous plants: *Silene acaulis* in the Alaskan tundra (which grows as a single "cushion") appears to reach ages of at least 300 years (Morris and Doak 1998), and the distinctive gymnosperm *Welwitschia mirabilis* (Welwitschiaceae), which grows mainly as two leaves atop a short woody stem, appears to reach much greater ages (Henschel and Seely 2000). Some genets of clonal sedges have been carbon-dated to several thousand years of age, though no single ramet is nearly that old (de Witte and Stöcklin 2010). There are also short-lived plants—including many annuals—that have secondary growth and therefore must be considered woody.

## The growth of clonal plants affects their ability to take up patchy resources

In growth form **clonal** plants—those with numerous ramets—can vary in more complex ways (see Figure 9.10 for definitions of *genet* and *ramet*). The spatial distribution of ramets can vary, as can the size and shape of each ramet. The two most important ecological factors affecting the spatial distribution of ramets are competition among genets (including individuals of different species) and spatial variation in the distribution of resources. These factors are not necessarily independent of each other. If resources (say, available soil nitrogen) are concentrated in distinct patches, individual genets are subject to selection for their abilities both to acquire these resources and to prevent their neighbors from getting them.

Clonal plants have a variety of patterns of spatial dispersion. Several kinds of explanations for this variability,

none of which are mutually exclusive, have been investigated. These explanations involve four different factors: the mechanics of clonal spread, the degree to which plants actively explore the environment seeking resource patches ("foraging"), the degree to which plants grow by simply following the distribution of resources, and the degree to which the clone is integrated.

Some of the variability in spatial dispersion may be caused simply by differences in the biological mechanisms of clonal spreading. Each method of spreading, however, can potentially result in a wide range of patterns of ramet dispersion. Spread by vegetative (nonsexual) connections, for example, can lead to the slow growth of a compact patch of one genet, or it can lead to a more diffuse structure, with a number of smaller patches of ramets. The spread of genets by the dispersal of **plantlets** (small plants created vegetatively) or asexual seeds (described below)—units that obviously are not physiologically integrated—may also create a variety of spatial patterns.

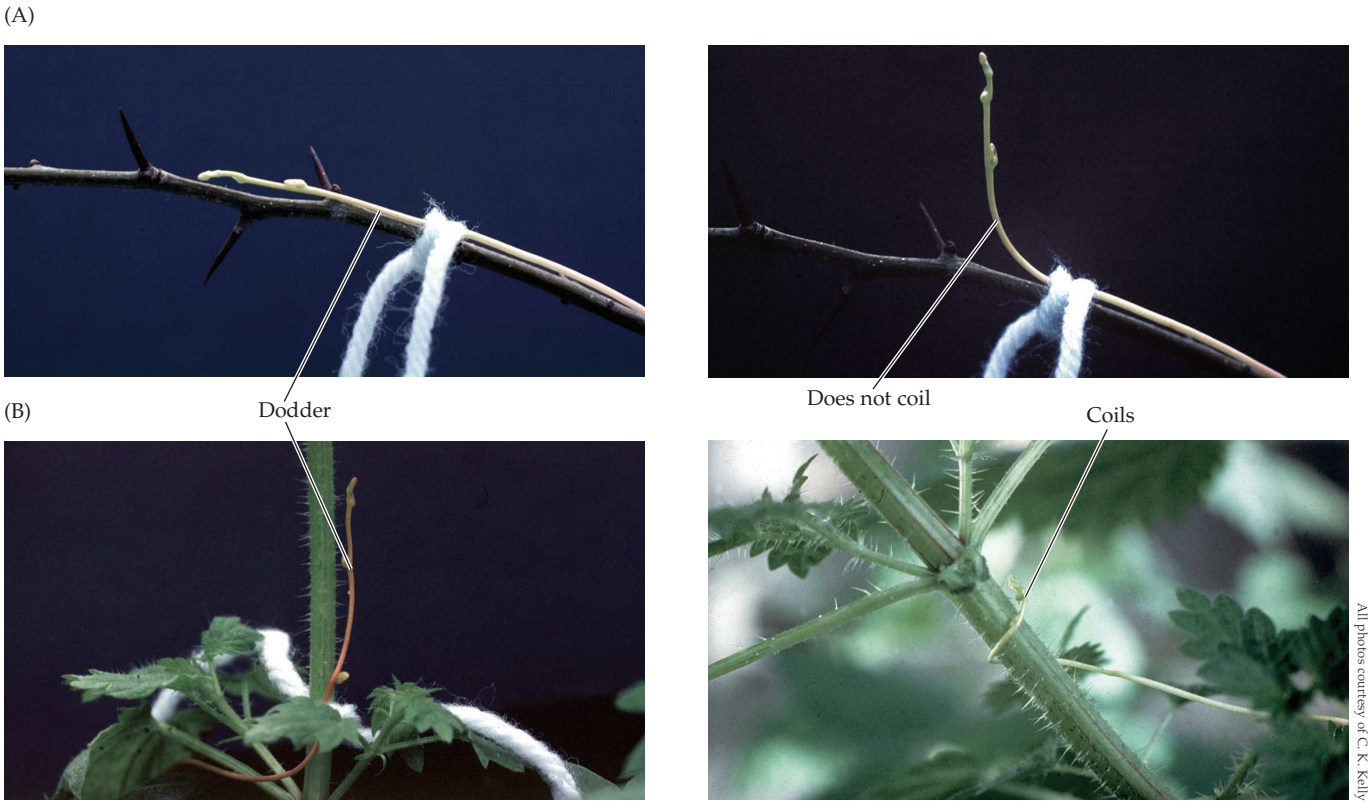
It is also clear that some variation in clonal form may be a more or less passive response to spatial variation in resource availability. Once a resource-rich (or toxin-rich) patch is encountered, roots and new ramets tend to follow the concentration gradient of increasing resources (or avoid the gradient of increasing toxins). Amy Salzman (1985) showed that a clonal ragweed, *Ambrosia psilostachya* (Asteraceae), was able to avoid concentrations of saline soil in this manner.

This kind of clonal growth has been called "foraging" as an analogy with the way in which animals search for food. Colleen Kelly (1992) made a useful distinction, however, between plants passively following a resource gradient and those actively "seeking" new patches of resources. She suggested that one can distinguish the two responses in cases in which the plant growth occurs before any resource is actually taken up. Her studies of *Cuscuta europaea* (dodder, Convolvulaceae)—a parasitic plant that coils around and sinks absorptive tissues into its host before any resource uptake can occur—showed that this plant does forage in her strict sense. Dodders attacked different host species at different rates. Remarkably, the rate at which each host was attacked was predicted well by optimal foraging models derived from animal behavior studies (Figure 6.4). Kelly suggested that this kind of strict-sense foraging might occur in many other plants. Other researchers use the term *foraging* to refer to any phenotypically plastic pattern of clonal growth that may have an effect on the acquisition of spatially variable resources (de Kroons and Hutchings 1995).

A large body of studies has shown that clonal plants readily vary their patterns of spread under different abiotic and biotic environmental conditions. Andrew Slade and Michael Hutchings (1987), for example, showed that a single genet of the British clonal herb *Glechoma hederacea* (ground ivy, Lamiaceae) varied its growth form (e.g., number of ramets, number of stolon branches, stolon internode lengths) as the availability and spatial arrangement of soil nutrients were varied. Important though this kind of result is, it does not provide information about whether these responses are adaptations—that is, whether they have been shaped by natural selection (Gould and Lewontin 1979). One step toward this kind of inference was a study by Mark van Kleunen and Murkus Fischer (2001) on the responses of clones of the circumboreal herb *Ranunculus reptans* (creeping spearwort, Ranunculaceae) to spatially variable competition from *Agrostis stolonifera* (creeping bentgrass, Poaceae). They found not only that *R. reptans* varied its growth form (vertical angle of the first stolon internode and specific stolon length) in response to competition, but that genotypes varied in their responses. This among-genotype variation indicates the potential for the population to respond to selection on these traits; however, it does not provide any information about whether these traits have been shaped by selection in the past (see Chapter 9). More work is needed before we can know the extent to which growth form has been shaped by selection.

Clonal plants vary considerably in the extent to which they remain physiologically integrated. Ramets connected by a root or rhizome can share water, carbohydrates, and other nutrients. Such sharing may benefit the plant in several ways. It may allow a genet to spread into microhabitats where a seedling could not survive. The genet may grow faster by increasing the number of ramets. If there are very small-scale differences in the environment (e.g., greater availability of nitrogen in one spot and phosphorus in another), the differences can be averaged by the plant, sometimes leading to improved overall vigor for the genet. Thus, by itself, the fact that some clonal plants remain integrated does not tell us whether this integration provides a fitness advantage, nor whether it was shaped by selection. The hypothesis that a trait such as clonal integration is adaptive must be tested explicitly. Similarly, while it is widely believed that clonal integration may play an important role in invasions of some exotic species (Liu et al. 2016), evidence for this claim is only observational and not demonstrated experimentally. Indeed, this argument begs the question: Why are many invaders not clonal, and why are many clonal plants not more highly integrated?

In the early years of study of clonal architecture, Lesley Lovett-Doust (1981) characterized plants as having either "phalanx" or "guerrilla" growth, spreading either grouped tightly like a classic advancing army or, like guerrilla forces, through ramets penetrating their competitors' turf. One problem with this is scale: the categories are subjectively determined based on what individual ecologists see as being tightly grouped. Another complication is that plants use multiple resources and have multiple competitors, so a species might have phalanx growth for competition with one species for one resource, and guerrilla




All photos courtesy of C. K. Kelly

**Figure 6.4** *Cuscuta europaea* (dodder, Convolvulaceae) is an obligate parasite. Like the other 400 or so *Cuscuta* species, *C. europaea* can use a number of host species, which vary in quality as resources. To attack a host plant, it coils around the host and then sinks absorptive tissues into the host's phloem. The "decision" to attack a particular host is made before any resource uptake, and it is made within minutes of contact with the host, as shown in this choice experiment. (A) A strand of dodder tied to a branch (left) of *Crateagus monogyne* (hawthorn, Rosaceae) fails to coil around this host (right); *Crateagus* is sometimes accepted as a host, but dodder grows poorly on it. (B) A strand of dodder tied to *Urtica dioica* (stinging nettle, Urticaceae) coils readily; dodder grows well on *Urtica*. Host choices are predicted well by models of optimal foraging from animal ecology.

growth in competing with another species for another resource. The phalanx/guerrilla categorization has faded in importance, but it still appears in some literature on invasive plants. We suggest that progress will be made if we study specifically how plants explore space at different spatial scales and how that depends on the environment.

## 6.4 Plants Reproduce both Sexually and Asexually

What is reproduction in a plant? It may seem surprising to begin a discussion of reproduction by asking what it is—after all, we have a pretty good idea of what our own reproduction involves. But the biology of plants means that their reproduction is not so clearly defined. Reproduction must, by definition, involve the formation of a new individual. But the concept of an individual is not as well defined in plants as it is in animals (especially vertebrates). A genet may be made up of many ramets. While everyone would include sexual reproduction in the category of reproduction, production of new ramets is sometimes best thought of as asexual reproduction and sometimes as growth, depending on the ecological context. For example, if we are concerned with natural selection on *Populus tremuloides* (quaking aspen, Salicaceae), which forms gigantic clones consisting of many individual trees (see Figure 9.11), we look at the entire clone as the individual and consider reproduction to occur only when new genets are formed. But if we are interested in the spread of aspen on a mountain, it may be useful to think of aspen as reproducing in two ways, through new ramets (asexual or vegetative reproduction of new stems) and through new genets (sexual reproduction).

### Many plants reproduce vegetatively

**Vegetative reproduction** (reproduction by vegetative growth of a new ramet) is extremely widespread among plants. It is an example of **apomixis**, a general term for asexual reproduction. It occurs in most herbaceous

perennials as well as in many shrubs and a few trees, and it can take a number of forms. Often the result is a collection of ramets that are physiologically integrated with other ramets, at least initially. For example, some plants generate new ramets by sending out modified shoots. *Fragaria* (strawberries, Rosaceae) spreads by **stolons** (runners)—branches or stems that spread at or just above the surface of the soil and generate ramets at nodes touching the ground. *Eichhornia crassipes* (water hyacinth, Pontederiaceae), a South American native, has become a serious freshwater invasive pest in many tropical and subtropical regions after human introduction because of its ability to spread rapidly through stolons. It is currently a serious problem in many waterways in Africa, the southern United States, and northern Australia. Many grasses (bamboos, for instance), gingers, and irises spread by **rhizomes**, which are underground horizontal stems growing near the soil surface. The dense rhizomes of *Hedychium gardnerianum* (Kahili ginger, Zingiberaceae), a Himalayan plant introduced into Hawaii, make it a problematic invasive plant that crowds out many native species. Other species, including familiar plants in the lily family (Liliaceae) such as tulips (*Tulipa*), onions (*Allium*), and daffodils (*Narcissus*), spread by division of **bulbs**, underground rosette stems that store nutrients. Many woody plants spread by the formation of **suckers**, meristematic buds on some of their near-surface roots capable of generating new stems and leaves.

Some plants (such as some seagrasses and cholla cacti; see Figure 8.1 and Box 8A) spread by **clonal fragmentation**, in which pieces of the plant break off that are capable of rooting to form new independent plants; the new ramets are thus never physiologically integrated with other ramets. *Elodea canadensis* (Canadian pondweed, Hydrocharitaceae), a common aquarium plant, was introduced into Great Britain in the 1800s. In almost every part of Britain, only male plants were introduced, so reproduction is necessarily asexual. *Elodea* spread so extensively by clonal fragmentation that British canals became severely clogged, and it is speculated that this plant may have contributed to the predominance of railways rather than barges in British commerce (Simpson 1984). Some plants, such as *Kalanchoe daigremontianum* (mother of thousands, Crassulaceae), produce little plantlets on the edges of leaves that can become dislodged and form new independent ramets. Physiologically independent ramets are also created in plants that reproduce with **bulbils** (tiny bulb-like organs vegetatively produced in inflorescences or leaf axils), which are formed above ground in monocots such as *Agave* (century plants, Agavaceae), *Allium* (onions and garlics, Liliaceae), *Dioscorea* (true yams, Dioscoreaceae), and *Lilium* (lilies, Liliaceae), in dicots such as *Polygonum viviparum* (Polygonaceae) and *Saxifraga cernua* (Saxifragaceae), and in ferns such as *Cystopteris bulbifera* (Dryopteridaceae).

## Some plants produce seeds asexually

**Agamospermy**, the production of seeds asexually, also occurs in a number of plant taxa (many people use the term *apomixis* to refer to this process). There are several mechanisms by which asexual seed production can occur, described by a complex terminology (Gustafsson 1946, 1947a, b). Most of these mechanisms involve partial meiosis without a reduction division, so the chromosome number remains unchanged, but in some cases there is no meiosis at all. In most cases, however, recombination is prevented, so the new embryos are **clones**—genetic duplicates—of the plants on which they are formed.

Agamospermy is fairly widespread among plant families. It occurs in some familiar plants, including members of the Asteraceae such as dandelions (*Taraxacum*) and hawkweeds (*Hieracium*); citrus plants (Rutaceae); a number of plants in the Rosaceae, such as raspberries and their relatives (*Rubus*) and cinquefoils (*Potentilla*); some buttercups (*Ranunculus*, Ranunculaceae); many grasses (Poaceae); and nettles (*Urtica*, Urticaceae). While sexual reproduction can provide populations with a possible long-term advantage because it can promote genetic diversity, asexual reproduction like agamospermy can also be advantageous. In particular, it preserves favorable combinations of genes, makes it possible to reproduce at extremely low population densities, and eliminates the need for traits like showy flowers, which can be quite costly in resources. Moreover, individuals reproducing asexually can reap a considerable advantage in Darwinian fitness over their sexual relatives: they pass on twice as many genes to their offspring. Much research has focused on the evolution of sexual reproduction (Orive 2019).

## The sexual life cycles of plants involve alternation of generations

Plant life cycles are fundamentally different from those of animals because all plants have **alternation of generations**: They have a haploid generation that produces haploid gametes, alternating with a diploid generation that produces haploid spores. Alternation of generations set the stage for the diversification and success of plants on land. Most of the plants that we see around us are the diploid generation, because that is the generation that predominates in the seed plants (gymnosperms and angiosperms), the most abundant plants in the modern world. The individuals that actually mate are haploid. In some plant phyla, such as the Bryophyta (mosses), however, the haploid generation is actually the larger, longer-lived, dominant generation; the familiar moss plants you see growing in forests are haploid.

The haploid generation is called the **gametophyte** because it produces gametes. The diploid generation is called the **sporophyte** because it produces spores (haploid somatic cells). Only the gametophytes technically have a

gender (are male or female), which is reflected in the gender of the gametes they produce. However, in flowering plants, we sometimes speak loosely of a male or a female plant (sporophyte) if the individual produces spores that result in only one type of gamete, either male (pollen) or female (embryo sacs). Figure 6.5 gives a general picture of plant life cycles; consult an introductory biology or botany text or a web source for a more complete discussion. Many conifers produce small, short-lived cones in which the male gametophyte is produced and from which pollen is released, as well as the more familiar woody, long-lasting cones in which the female gametophyte is produced and where seeds are formed.

Features of each of the two plant generations constrain the range of possibilities for the other generation (Niklas 1997). If gametophytes are the dominant generation, living independently of the sporophytes (as they do in mosses and ferns), they can live only in moist places to allow reproduction. Free-living gametophytes must be large enough to nourish the new sporophyte, at least initially, but must also be low enough to the ground that they can be reached by swimming sperm. Moreover, they must be near one another if the sperm are to get from plant to plant. If sporophytes grow to a large size, they cause the demise of the parental gametophyte on which they grow, limiting its size (as in ferns). By contrast, gametophytes maturing inside the sporophyte (as in the seed plants) get their nutrition from the sporophyte. This arrangement enables the gametophytes to be much smaller. Most importantly, it also allows reproduction to occur in dry environments and frees the gametophytes from the constraint of having to be near one another for fertilization to occur.

Most familiar plants are the sporophytes of gymnosperms (such as the conifers) or angiosperms (flowering plants). The reproductive structures of gymnosperms typically are cones or cone-like structures. The reproductive structures of angiosperms are flowers. Within either cone or flower, the gametophyte is produced and fertilization occurs. Both gymnosperms and angiosperms produce seeds.

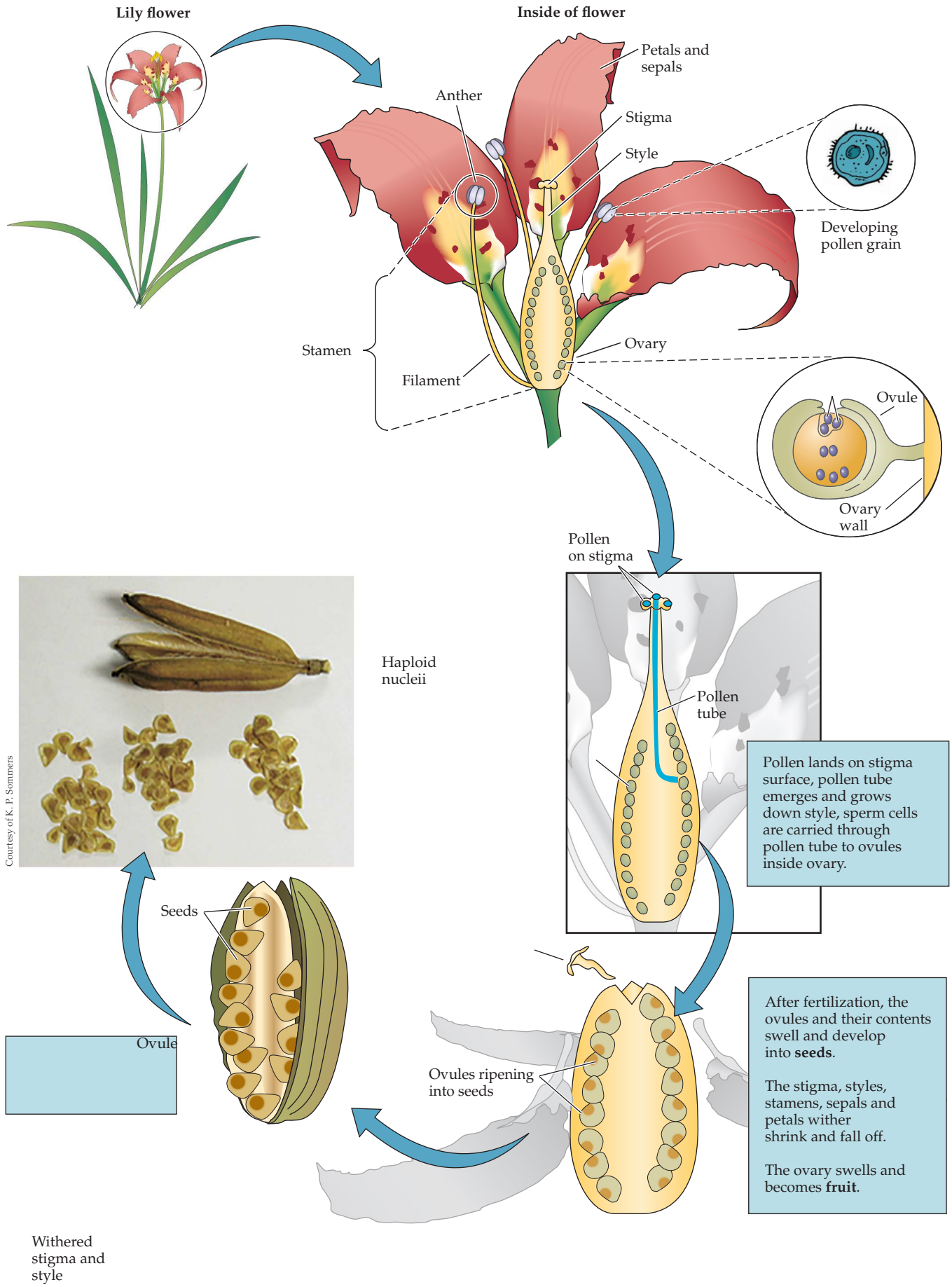
In angiosperms, fertilization results in the flower's ovary maturing into a fruit with seeds inside. Although not all angiosperms produce showy flowers or fruit that is edible by humans, all angiosperms produce flowers, seeds, and fruits. Flowers can be large and conspicuous or small and inconspicuous, depending largely on the mechanisms by which pollination occurs in a particular plant group. Flowers sometimes act to attract certain animals (and discourage others) as pollinators.

When particular cells in flowers undergo meiosis, the haploid daughter cells that result are called spores. These spores grow mitotically into either **pollen grains** (also called microgametophytes) or **embryo sacs** (megagametophytes). An embryo sac usually consists of seven cells, one of which ultimately becomes the egg cell. It develops within a sporophytic structure called an **ovule**, which is located inside the **ovary**. The spores that become pollen grains undergo two further mitotic divisions. One division creates a vegetative cell and a generative cell. The latter divides again, forming two sperm cells. These three cells are haploid, and they make up the entire male gametophyte in angiosperms. The pollen grains develop in a sporophytic structure called an **anther**, located at the end of a stalk called the **stamen**.

Pollen germinates after landing on a stigma, the receptive surface of the ovule (Figure 6.5). Upon germination, the vegetative cell grows into a pollen tube, which penetrates the stigma and style and, ultimately, the ovary and an ovule inside the ovary. The sperm cells are transported down the pollen tube, where one of them fertilizes the egg cell, forming a new diploid zygote. This zygote grows and divides by mitosis, forming the embryonic tissues of a new sporophyte. In almost all angiosperms, the second sperm cell fuses with a specialized cell in the embryo sac that has two nuclei. This triploid cell usually grows to become **endosperm**, which serves as the primary nutritive source for the developing embryo. In some angiosperms (especially orchids) this **double fertilization** occurs but the triploid tissue does not develop and the seeds are remarkably small.

Once fertilized, an ovule becomes a **seed**. In angiosperms, as the seeds develop, the ovary starts to swell; petals, sepals, and related structures shrivel and fall off; and a fruit develops. The seeds grows inside the developing ovary, which becomes a **fruit**, the mature ovary of a flower produced by the maternal sporophyte. The tissue in the seed coat comes from the ovule. The tissues in a fruit, and some of the seed tissues, come from the maternal sporophyte; only the embryo and endosperm inside the seed are new genetic individuals. The seed can be dispersed by wind, water, or animals and can sometimes remain dormant for a long time. The fruits of angiosperms can further enhance the protection and dispersal of the seeds enclosed within them. Fruits can be large or small, colorful, sweet, and fleshy or dry and hard, toxic or nutritious, depending on the evolution and ecology (especially the seed dispersal mechanisms) of the plant producing them.

All angiosperms—from bamboos and rice, wheat, and corn (grasses, Poaceae) to maple, oak, and orange trees (species in the genera *Acer*, *Quercus*, and *Citrus*, respectively) to roses (*Rosa* spp.), *Acacia* spp. (Fabaceae), *Eucalyptus* spp. (Myrtaceae), the many dipterocarp rainforest trees of Southeast Asia (Dipterocarpaceae), and goldenrods and ragweeds (*Solidago* spp., *Ambrosia* spp., Asteraceae)—have flowers, fruits, and seeds. Their flowers, fruits, or seeds may or may not be colorful and obvious;


Lily flower
Inside of flower
Petals and sepals
Anther
Stigma
Style
Developing pollen grain
Stamen
Filament
Ovary
Ovule
Ovary wall
Pollen on stigma
Pollen tube
Pollen lands on stigma surface, pollen tube emerges and grows down style, sperm cells are carried through pollen tube to ovules inside ovary.
Haploid nucleii
Courtesy of K. P. Sommers
Seeds
Ovule
Ovules ripening into seeds
After fertilization, the ovules and their contents swell and develop into **seeds**.
The stigma, styles, stamens, sepals and petals wither shrink and fall off.
The ovary swells and becomes **fruit**.
Withered stigma and style

◀ **Figure 6.5** Reproduction in a typical angiosperm, as illustrated by *Lilium catesbaei* (pine lily, Liliaceae). Flowers are the reproductive structures of the sporophyte. Meiosis inside the ovule (which is sporophytic tissue inside the ovary) leads to a haploid unicellular megaspore. The megaspore develops by mitosis into the female gametophyte, which has eight haploid nucleii in seven cells, and is enclosed in the ovule. The cell with two nuclei will become the endosperm after fertilization. Meiosis inside the anther (a part of the sporophyte) leads to a haploid unicellular microspore. The microspore develops into a binucleate pollen grain—the male gametophyte. When a pollen grain lands on a stigma, it grows into a pollen tube, which grows down the style and into the ovules. One of the nuclei divides and becomes two sperm cells, which are transported down the pollen tube. One sperm cell fuses with the egg cell to form the diploid embryo, and the other sperm cell fuses with the binucleate cell to become the triploid endosperm. Both the embryo and endosperm divide mitotically to become multicellular. The ovule and its contents swell and become seeds. A seed thus contains three different genetic individuals: the diploid seed coat (formed from the cells in the ovule), the diploid embryo, and the triploid endosperm. Some angiosperm families have variations of this general pattern.

the details may differ, but the general pattern is the same as the one illustrated for *Lilium catesbaei* (Figure 6.5).

In the other seed plants, the gymnosperms, the basic patterns of gametophyte development and seed development are similar, with the key differences being that the female gametophyte consists of many more cells, double fertilization does not occur (so the nutritive tissue is haploid and endosperm is absent), and development usually takes substantially longer.

The tremendous diversity of seed plants and their dominance over terrestrial habitats are due largely to these reproductive structures, which permit them a great deal of independence from water for mating and dispersal. Finally, the interaction of several different kinds of tissues—from both gametophyte and sporophyte—has allowed the evolution of mechanisms for limiting matings. These characteristics of plant life cycles are basic to plant reproductive ecology and to the evolutionary success of the seed plants, especially of the angiosperms.

## 6.5 The Movement of Pollen Is an Important Aspect of a Plant's Life Cycle

### The pollen of many plants is moved by the wind

When most people think of flowers and pollination, they think of showy flowers that are visited by birds or insects. But for a large number of plants (at least 10% of plant species; Friedman and Barrett 2009), pollen is carried from flower to flower mainly by wind. There are two quick ways to convince yourself that wind pollination is important. First, almost all pollen transfer in grasses, rushes, and most temperate-zone trees (including all conifers, oaks, beeches, maples, elms, willows, and many others) is by wind. Second, one of the most common human medical complaints (at least in developed countries) is allergy to pollen ("hay fever" is one type)—the overreaction of our immune systems to the foreign proteins in the huge amounts of pollen in the air. Animal-pollinated plants with conspicuous, colorful flowers typically have much larger and fewer pollen grains that are not carried in the wind and are unlikely to be encountered by human respiratory systems and mostly do not cause these allergies.

While many animal-pollinated species produce attractive flowers, wind-pollinated flowers are rarely showy. There is generally no selective advantage to producing metabolically costly large petals, scents, nectar, and floral pigments if they are not needed to attract pollinators. In fact, there is often a disadvantage to making large petals and other organs that interfere with pollen transfer by wind. The grass family, for example, is almost entirely wind-pollinated, and grass flowers lack petals and sepals (Figure 6.6). Many wind-pollinated species have large, feathery stigmas to enhance pollen capture. Pollen in wind-pollinated plants is light and rarely sticky, whereas insect-pollinated plants often have heavy, sticky pollen.

Wind-pollinated plants produce massive quantities of pollen because they cannot exert much influence over where their pollen blows. The quantity of pollen produced can be expressed as the ratio of pollen to ovules produced by an individual or population. Wind-pollinated plants typically have much higher pollen:ovule ratios than their insect-pollinated relatives. For example, Janet Steven and Donald Waller (2004) compared several species of *Thalictrum* (meadow rue, Ranunculaceae). They found that the insect-pollinated *T. sparsiflorum* had a pollen:ovule ratio of 4240, while the wind-pollinated species *T. alpinum*, *T. fendleri*, and *T. dioicum* had pollen:ovule ratios of 29,980, 16,880, and 54,380, respectively. Two important caveats must be given here, however. First, although it is true that within plant lineages, wind-pollinated species typically have much higher pollen:ovule ratios than related animal-pollinated species, a particular value for the pollen:ovule ratio does not necessarily indicate the pollination mechanism. One cannot tell, for example, from the pollen:ovule ratio for *T. sparsiflorum* that it is animal-pollinated. In some plant lineages this is a high pollen:ovule ratio, but in *Thalictrum* it is a low ratio. Second, in many plants, pollen is carried both by wind and by animals. The presence of petals does not prohibit wind pollination, nor does their absence mean that animal pollination is necessarily unimportant.

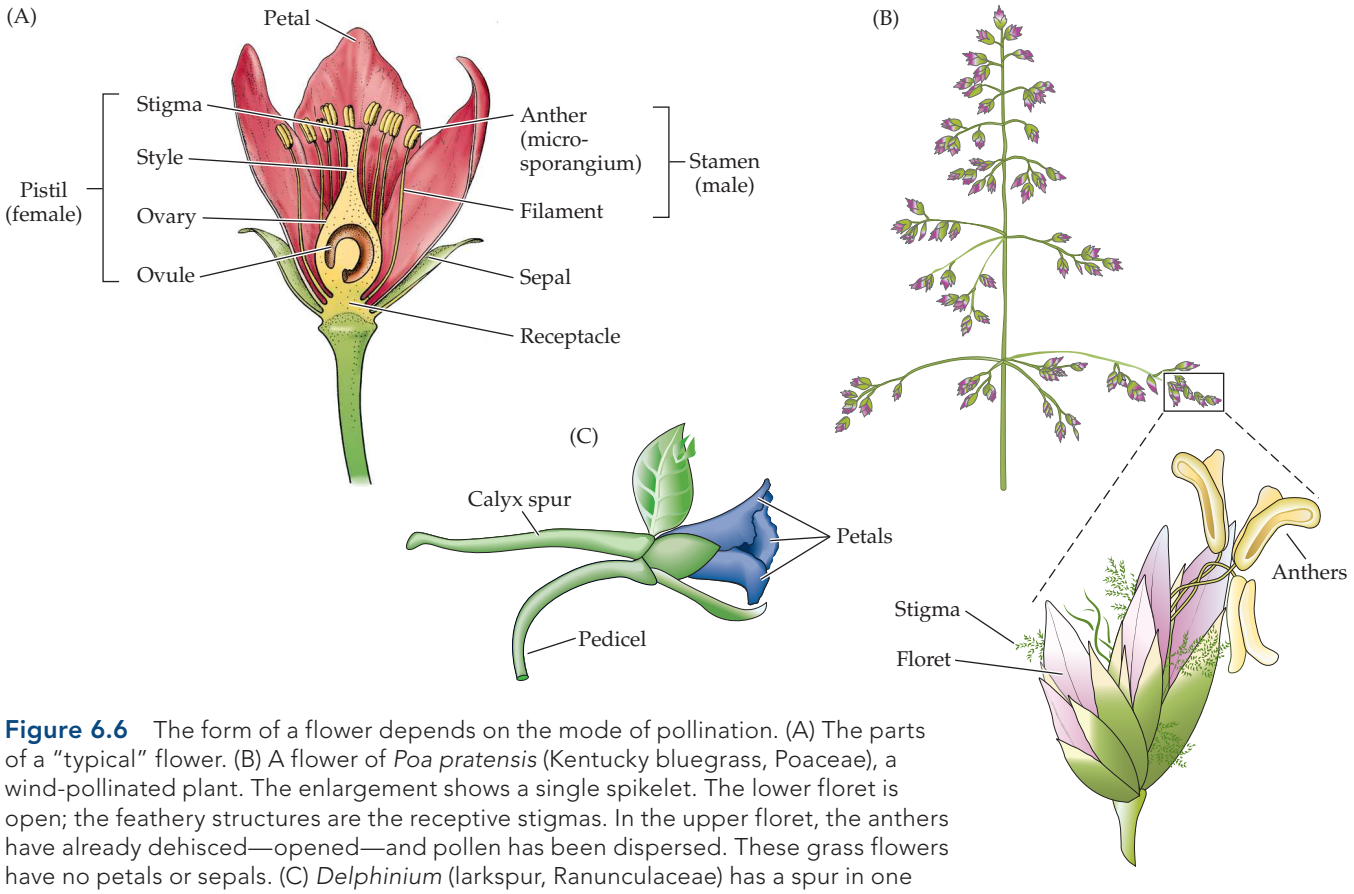



**Figure 6.6** The form of a flower depends on the mode of pollination. (A) The parts of a "typical" flower. (B) A flower of *Poa pratensis* (Kentucky bluegrass, Poaceae), a wind-pollinated plant. The enlargement shows a single spikelet. The lower floret is open; the feathery structures are the receptive stigmas. In the upper floret, the anthers have already dehisced—opened—and pollen has been dispersed. These grass flowers have no petals or sepals. (C) *Delphinium* (larkspur, Ranunculaceae) has a spur in one petal that contains nectar, restricting its animal pollinators to those with long tongues.

Pollen transfer by wind is much reduced in sheltered habitats where air movement is reduced. Thus, wind pollination is most common in plants of open habitats and favorable seasons. Wind-pollinated temperate forest trees usually flower well before their leaves emerge in spring. Pollen movement among wind-pollinated trees is generally greatest in early spring when trees are leafless. Different wind-pollinated species may have concentrated and largely nonoverlapping periods of pollen movement, which might increase pollination effectiveness. Spring also often happens to be the windiest time of the year, and wind can carry pollen very long distances. On June 5, 1998, local hunters at Repulse Bay in extreme northern Canada along the Arctic Ocean noticed large amounts of pollen at the edges of ponds (Campbell et al. 1999). It proved to be pollen from *Pinus banksiana* (jack pine, Pinaceae) and *Picea glauca* (white spruce, Pinaceae) that had been blown on June 1 from Quebec, over 3000 km away!

It is tempting to think that wind pollination is evolutionarily primitive. It seems so simple and inelegant when compared with some of the complex adaptations some plants have for animal pollination. And it is clear that adaptation to animal pollinators has been a driving force in plant evolution. However, both ancient lineages (such as gymnosperms) and more recently evolved taxa (such as grasses) have wind pollination. Some wind-pollinated taxa have given rise to animal-pollinated taxa, while others have been derived from animal-pollinated ancestors.

## Visual displays are important for attracting animal visitors

Flowers exist for one reason only: sex. Everything about them has to do with sex. This is obvious in the case of the floral parts that are directly involved in producing or receiving pollen or in the growth of seeds and fruits. But it is also true of the other floral parts, such as petals, that function mainly to influence the transfer of pollen.

Animals visit flowers because they generally contain rewards—mainly nectar, which is rich in sugars and sometimes amino acids, and pollen, which is rich in protein. Many people think of floral traits such as shape, color, and scent as being important in attracting animal pollinators—as advertisements of the rewards the animal can expect to acquire at the flower. But the interaction between plants and their floral visitors often involves a great deal of mutual cheating, false advertising, and theft. As we will see below, some plants appear to advertise rewards, but do not deliver them. Some animals visit flowers and remove rewards, but do not pollinate the flowers. Moreover, one species' attractant may be another species' repellent.

Displays of bright red flowers catch our eye (Figure 6.7). Although most people also enjoy looking at white

flowers, we often do not notice them as easily as we notice red ones. Animals vary in their ability to see and discriminate among colors (Kevan 1978). For example, humans see colors between the wavelengths of about 380 nm (violet) and 780 nm (deep red), and our eyes are most sensitive to blue (436 nm), green (546 nm), and red (700 nm). Insects, on the other hand, can usually see colors at shorter wavelengths than humans. Bees see wavelengths mainly from 300 nm to 700 nm. Thus, bees are less sensitive than we are to longwave light and have difficulty discriminating among shades that we perceive as red, but there are colors in the ultraviolet part of the spectrum that they can see—those between 300 and 380 nm—that are invisible to us. Birds' eyes are most sensitive to colors in the middle and red parts of the spectrum. A word of caution: the fact that certain colors are invisible to bees, birds, or humans does not mean that the flowers themselves are invisible to those taxa! Rather, flowers that are visually attractive to bees may not be as noticeable, attractive, or easily distinguishable to birds, and vice versa. Many bee-pollinated flowers have yellow hues, while bird-pollinated flowers are often red or orange. Flowers pollinated by moths are usually white to pale yellow, while bat-pollinated flowers are usually white to brown. Butterflies are often attracted to yellow or blue flowers. A note of caution is again in order, however: Many flowers reflect light in several different wavelengths. Flowers that we perceive as red may be easily distinguished by bees because they also reflect in the blue wavelengths.

Flower color plays a more complicated role in pollination than these observations might suggest. Many flowers have contrasting colors in the **perianth** (the petals and sepals together). For example, the central portion of the flower may contrast sharply with the rest of the flower, as in many irises (Iridaceae). In most species that have been studied, these contrasts seem to have one of two functions. In some cases, the contrast itself is the attractant for pollinators. In other cases, one of the colors functions as a "nectar guide"—pollinators are first attracted to the flower as a whole and are then attracted by the color contrast to the areas containing nectaries, anthers, and stigma (Figure 6.8).

Courtesy of S. Scheiner

**Figure 6.7** *Lobelia cardinalis* (cardinal flower, Campanulaceae) has bright red, tubular flowers. This widespread herb from moist areas of North America is frequently pollinated by hummingbirds.

(A)

(B)

Both photos courtesy of T. Eisner

**Figure 6.8** *Rudbeckia hirta* (black-eyed Susan, Ranunculaceae) under white (A) and ultraviolet (B) light. The color contrast that we see only under ultraviolet light is visible to bees in ordinary sunlight and serves as a nectar guide.

Flower colors can vary over time, and numerous studies have shown that these variations make a difference in attracting pollinators. In most cases, older flowers fade, although there are some, such as those of the prickly pear and cholla cacti, that grow darker over time. These changes in flower color seem to function much like nectar guides. The old flowers add to the attractiveness of the floral display as a whole, but once an animal has been drawn to the plant, it tends to be attracted to young, fresh flowers; these are the ones containing nectar or pollen.

Color depends on the incident light on the flower. This means that a flower can have a different apparent color in the morning than in the afternoon, in a shadier or a brighter spot, or at different times during the season, even if the pigments present and the structures of the petals are unchanged. How important is this variation to pollinators? Kellen Paine and his coworkers examined intraspecific flower color variation in 14 mainly bee-pollinated species in New Mexico (Paine et al. 2019). In a majority of populations, this variation in color made almost no difference in visitations by bees.

Other parts of plant bodies have also been modified to act as attractants. **Floral bracts**—specialized leaves found below many flowers—are often colored and appear to act as part of the plants' display. A well-known example is *Euphorbia pulcherrima* (poinsettia, Euphorbiaceae; see Figure 2.16A), whose red bracts and upper leaves are far more colorful than its tiny yellow flowers. Many vase-forming bromeliads also have leaves that are more brightly colored than their flowers (see Figure 2.19A).

Flowers are often arranged in aggregations called inflorescences. There are several types of inflorescences, ranging from the capitulum (head) of composites (Asteraceae) to the catkins of beeches and oaks (Fagaceae) to some rather loosely arranged racemes in plants such as buttercups (Ranunculaceae) (Figure 6.9). Inflorescence structure is important in pollination ecology. Having flowers gathered into an inflorescence may act to increase the size of the attractive display without altering the flowers themselves. It is likely that the form of an inflorescence also affects the behavior of pollinators.

## Animal visitors are attracted to plants with floral odors or acoustic guides

Many flowers have noticeable odors, which clearly act as attractants. Why have scent as well as visual attractants? It is generally thought that scent acts over much longer distances than visual attractants, since many animals are able to detect extremely low concentrations of molecules in the air. On the other hand, scents are hard to locate precisely. Once an animal is in the general area of a patch of flowers, it is likely to find them much more quickly by vision than by scent.

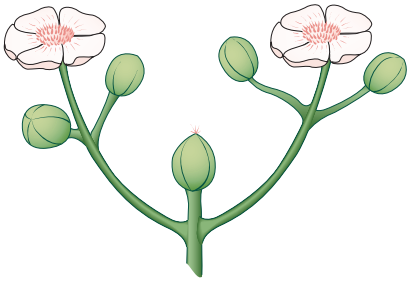
Much less is known about floral odors than about colors, partly because the science of measuring odors is not nearly so well developed as the science of optics. However, this has been changing in the last decade with the development of portable devices that can measure the very low concentrations of odorant molecules. It is now understood that odors can vary independently of colors. Genes that produce yellow or red flowers usually have no effect on flower odor. All of us have been surprised to find no odor, or an unpleasant one, coming from a visually attractive flower. Similarly, pollinators experience floral odors quite separately from the way they experience colors.

One point should, upon reflection, be obvious: odors that attract one group of animals may repel others. Casual observation of the responses of humans, dogs, and flies to various odors should convince you of this! Bees are often attracted to what we perceive as sweet odors, and bats to musty odors. A number of plants, especially the milkweeds in the family Apocynaceae and the aroids (Araceae), have odors similar to animal dung or rotting tissue, which are strongly attractive to the flies and gnats that pollinate these plants. Some plants have odors that mimic insect mating pheromones, which cause male insects to visit the flowers and sometimes attempt to mate with them. A study in Greece (Kantsa et al. 2019) revealed that there are strong associations between pollinating insect families and different classes of volatile organic compounds disseminated by plants.

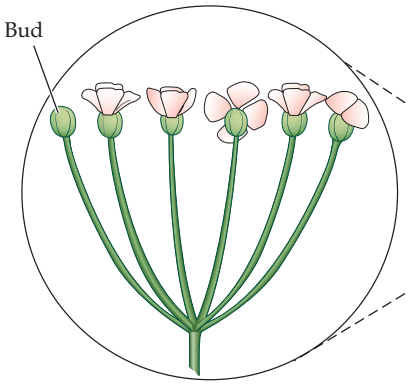
Olfactory and visual features of flowers are not necessarily independent of one another. In *Dracula lafleurii* (Orchidaceae), an orchid from the Andes mountains of Ecuador, a number of floral features are visually similar to those of mushrooms. Also, the flowers produce the same volatile compounds as mushrooms, which along with the visual cues, attract flies. By using 3D-printed flowers, Tobias Policha and associates (2016) were able to independently manipulate floral shapes, colors, and odors and ask how each plays roles in pollination. The visually similar features play an important role in dispersing the aromatic compounds. In this species, floral pattern is more important than floral colors.

**Figure 6.9** Some forms of inflorescences (arrangements of flowers on a branch). A principal division is between (A) determinate inflorescences (which have a flower at the end of each flowering branch) and (B) indeterminate inflorescences (which do not). A determinate inflorescence has a fixed maximum number of flowers; the terminal flower normally blooms first. By contrast, in an indeterminate inflorescence, flowering usually begins at the bottom, and additional flowers can be added by the apical meristem if resources are available. The distinction is often important ecologically because plants with indeterminate inflorescences can more finely adjust their production of flowers to available resources. ▶

(A) Determinate inflorescences

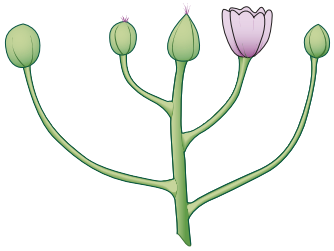


Cyme, opposite branching

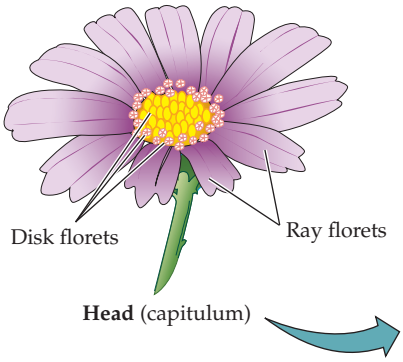



**Umbel**

Bishop's goutweed (*Aegopodium podagraria*)



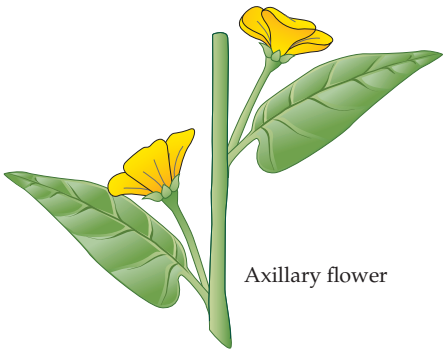
Cyme, alternate branching

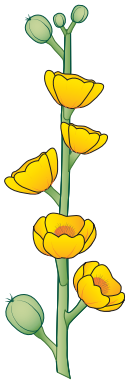



**Head** (capitulum)

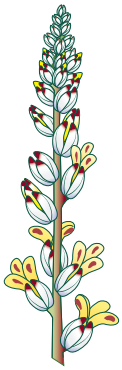
*Cosmos bipinnatus*

(B) Indeterminate inflorescences

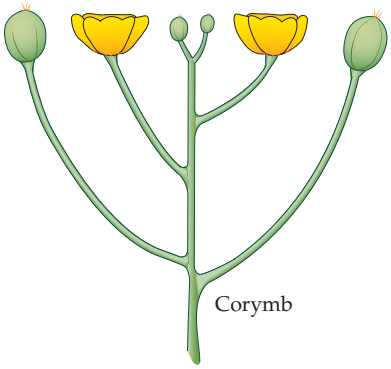


Raceme



**Spike**



All photos courtesy of D. McIntyre

*Lachenalia mutabilis*

Floral odors may attract less desirable visitors as well as pollinators. Carlos Nunes and his collaborators (2016) found that in the Atlantic Forest of southeastern Brazil, *Dichaea pendula* (Orchidaceae) produced the volatile compound 2-methoxy-4-vinylphenol that attracted both pollinating euglossine bees and flower-eating weevils. Trade-offs (see Chapter 8) like this play a large role in pollination systems.

Courtesy of Tia-Lynn Ashman

Tia-Lynn Ashman

In dioecious species, male and female plants may use odors differently. A review by Tia-Lynn Ashman (2009) found that in most of the 33 species studied, males produced more volatile molecules than females, and in a majority of species there was at least one component of the floral odor that males produced and females did not.

Possibly the most unusual flower attraction system known is that of the tropical vine *Mucuna holtonii* (Fabaceae), a bat-pollinated species (von Helversen and von Helversen 1999). The upper petal on the flower is raised when the flower opens. This petal is concave and acts as a sound mirror—an acoustic guide—reflecting back bats' sonar signals and letting them know that the flower is open. To reach the nectar, a bat has to press its snout deep into the flower. When it does so, another petal, which is holding the stamens, bursts open, throwing the pollen onto the bat's rump. Other bat-pollinated species may also have acoustic guides.

## Plants often need to limit unwanted visits

Many plants have adaptations not just for attracting certain visitors, but for repelling others. Why would a plant want to limit visitors? Once an animal visitor has removed pollen from a flower, that pollen has to reach another flower of the same species for it to contribute to the reproductive success of the plant producing it. If the next flower the animal visits is of another species, then the pollen is likely to be deposited in the wrong place. From the perspective of the plant, the best pollinator visits only one species, ensuring that the pollen gets to the right place. Limiting visits to only a small group of pollinators would increase the chance that the next plant visited would be of the same species. Numerous animals that are not effective pollinators are nevertheless attracted to flowers. From the animal's viewpoint, pollination is almost always an accidental consequence of this visitation. Thus, many plants have undergone selection to deter certain floral visitors—but, as we all know from personal experience, it can be hard to get rid of unwanted guests!

Changes in flower shape are the main way in which access to flowers is restricted. The earliest flowers are thought to have been pollinated by insects unspecialized for this role, such as beetles. Not surprisingly, flowers that are typically beetle-pollinated today are often open and bowl shaped, a generalized shape that is accessible to all visitors. On the other hand, many flowers are tubular, and some have long spurs or smaller conical structures behind the main flower tube (Figure 6.6C), in which nectar collects. Only animals with long tongues (or other long mouthparts), such as birds, butterflies, or moths, can easily reach the nectar and pollinate these flowers.

The type of reward a plant offers also restricts visitations. Beetles, for example, are generally unable to use nectar, whereas birds, bats, and bees use it easily. Many species eat as well as disperse pollen, which is a high-protein food source; nectar provides sugars, water, and sometimes amino acids or other nutrients. Some orchid bees (family Euglossidae) collect floral scents (by scraping the floral cells that secrete odor-producing oils), which they use in mating displays. Others use the oils in scents for nest building. The timing of reward availability is important as well: bat- and moth-pollinated flowers usually produce nectar only at night.

Mechanisms such as these do not, of course, provide any guarantees to the plant against wasted nectar and pollen. Many flowers are visited by large numbers of animals that transfer little pollen. In a study of *Calathea ovadensis* (Marantaceae), a tropical Mexican herb, Douglas Schemske and Carol Horvitz (1984, 1988) showed that moths and butterflies were the most common visitors, but bees accounted for the most pollination. Among the bees, the most efficient pollinators turned out to be relatively unimportant because they were uncommon. An extreme case of wasted resources is that of "nectar robbers," animals that insert their mouthparts between petals, or bore or chew through the side of the flower, and drink the nectar without ever coming into contact with the pollen. Very small insects such as ants may also act as nectar robbers because they can visit the nectaries without contacting stamens or stigma.

One remarkable mechanism of restricting pollinators is **buzz pollination**, in which pollen can be effectively extracted from flowers only by vibrating the anthers at an appropriate frequency (Vallejo-Marín 2019). This is done mainly by certain bee taxa. Some 6% of plant species—including many in the Solanaceae (nightshades) and some familiar plants like many *Vaccinium* (blueberries and cranberries, Ericaceae)—use buzz pollination.

## How strongly are floral characteristics associated with particular pollinators?

Color, odor, reward type, timing of rewards, and shape all affect the numbers and types of floral visitors. It is not surprising, therefore, that certain combinations of these attributes often seem to be associated with particular types of pollinators. From what we have said, one might expect that bee-pollinated flowers would be yellowish,

smell sweet, be broad enough for the bee to contact the anthers and stigma, and produce nectar during the daytime. Bird-pollinated flowers might be expected to be red or orange, produce copious nectar during the daytime (but little scent), and have long tubes or spurs—and so forth. Such associations of floral attributes with particular types of pollinators (usually taxonomic classes or orders) are called **pollination syndromes** (Faegri and van der Pijl 1979). These syndromes are usually thought of as being matched by adaptive suites of pollinator traits—plant-pollinator interactions are thought of as tightly coevolved mutualisms (Howe and Westley 1997). Some ecologists have stood the idea of pollination syndromes on its head, treating floral traits as diagnostic of particular pollinators: if flowers have traits X, Y, and Z, the reasoning goes, then the plant's important pollinators are in animal taxon A.

There is little doubt that many plants and pollinators have coevolved—that the plants have strongly influenced the evolution of their pollinators, and vice versa. Some well-known cases offer evidence for tight coevolution of particular pairs of plants and pollinators (Box 6A). But there also are reasons to think that the strength and generality of pollination syndromes have been overstated. In an influential critique, Nickolas Waser and his collaborators (1996) pointed to evidence suggesting that animals outside these syndromes accomplish much pollination—birds at "bee flowers," bees at "butterfly flowers," and the like. A principal criticism of the use of the pollination syndrome idea is that the syndromes are tendencies—and

## BOX 6A Specialized Plants and Pollinators

Plants' relations with their pollinators range from being highly specialized—with the plant depending on visits from only one or a few species of animals—to quite generalized. There are many interesting cases of specialization and adaptation by plants or their pollinators. In most cases, the animals are not as specialized as the plants, as they are visiting and pollinating plants for longer periods than any single plant species is in flower.

Flowers of *Yucca* (Agavaceae) are pollinated exclusively by *Tegeticula* moths. The moths pollinate the yucca flowers while laying their eggs in an ovary, and the larvae eat a fraction of the seeds in that ovary. If too many eggs are laid in a particular ovary, the plant will abort it. Some yuccas abort a sizable number of pollinated flowers.

In the orchid family (Orchidaceae), pollen grains are clumped together in structures called pollinia. The size of the pollinia makes it relatively difficult to move them from flower to flower. In some tropical epiphytic orchids, the flowers are shaped so that an insect—usually a bee—must crawl through a particular opening to get to the nectar or oil rewards, and in doing so, it cannot avoid hooking the pollinia in a way that facilitates their transport to the stigma of the next flower visited. Perhaps the most famous case is that of *Ophrys* (first studied by Charles Darwin [1877]), which has an odor similar to the mating pheromones of female bees. Males visit these flowers and "pseudocopulate" with them, thus transferring pollinia.

Many plants in the family Araceae have foul-smelling flowers that attract gnats and other small flies. Two common North American examples are *Symplocarpus foetidus* (skunk cabbage) and *Arisaema triphyllum* (jack-in-the-pulpit). The most extreme case is that of *Amorphophallus titanum* (voodoo lily), a plant native to Sumatra. This plant not only has the largest inflorescence in the plant kingdom, but may also have the most powerful odor, which is reputed to resemble a combination of rotting fish and burnt sugar (none of the authors has yet had the pleasure). There are reports of people fainting from the odor. Many African succulents in the dogbane family (Apocynaceae, another family that often has pollen grains clumped together in pollinia) are fly-pollinated. These plants attract flies by having flowers that smell much like decaying meat or, in some cases, animal dung.

Wild figs (*Ficus*, Moraceae) are mainly pollinated by specialized wasps. The flowers in this family are borne in a unique way: they are on the inside of the inflorescence, so a wasp can reach them only by boring a hole in it. The wasp then lays eggs in the inflorescence. When the brood hatches, they mate (usually, this is a brother-sister mating), and the wingless males then die (so every wild fig contains dead wasps). As the young females leave the inflorescence, they are dusted with pollen, and when they visit another inflorescence to lay their eggs, they pollinate it.

Courtesy of Allard Pierson, UoA, archive of Hugo de Vries. Photograph by C.G.G.J. van Steenis

*Amorphophallus titanum* (voodoo lily, Araceae), a plant native to Sumatra, has the largest single inflorescence of any plant. The man in the photograph is Hugo de Vries, an important early geneticist and one of those responsible for bringing Gregor Mendel's work to the attention of biologists.

Fascinating as these examples are, it is easy to get the wrong impression: as noted in the text, most plants have much less specialized mechanisms of pollinator attraction.

often rather weak ones—rather than laws. By concentrating on these tendencies, ecologists may be missing important variation. Under some fairly broad conditions, both plants and pollinators can be expected to evolve to be pollination generalists, rather than specialists. Thus, focusing on syndromes has caused us to neglect these generalist plant-pollinator interactions. Recent research by Javier Valverde and collaborators (Gómez et al. 2015; Valverde et al. 2019) on generalist plants in the genus *Erysimum* (wallflower, Brassicaceae) has begun to uncover new patterns of association between plants and pollinators. Research on the importance and generality of pollination syndromes is ongoing; reviews (Ollerton et al. 2009; Rosas-Guerrero et al. 2014) concluded that somewhere between about 30% and 75% of angiosperm species might be usefully described with pollination syndromes. Methodological differences beween these studies may account for their apparently different conclusions (Ollerton et al. 2015).

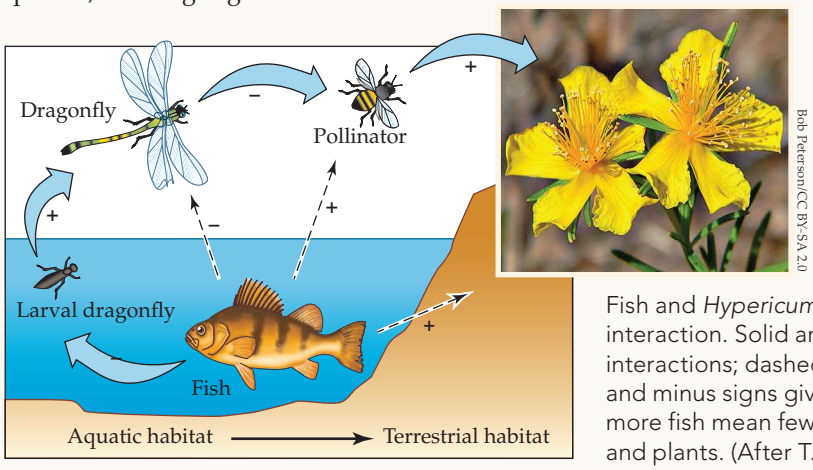
A recent review by Gordon Smith and associates (2019) found that males and females within pollinator species of insects frequently differ considerably in their pollination behavior. They can, for example, differ in visitation rates to flowers and in outcomes such as the rate at which nectar or other rewards are removed or the chance that a visit results in outcrossing. This should not be a surprise: insects vary within their populations, just as plants do. Future research in pollination may need to incorporate this additional complexity.

Pollination can involve some fairly complex interactions; some examples are given in Box 6B. While

## BOX 6B Some Complex Plant-Pollinator Interactions

While plants and their pollinators can each exert strong selective pressure on the other, they are also involved in other ecological interactions. Traits that affect pollination can reflect natural selection that results from these other interactions as well. Sharon Strauss and her coworkers (2004) have studied several such cases of complex selection. In a study of *Raphanus sativus* (wild radish, Brassicaceae) in California, they found that plants with differing flower colors tended to differ in their ability to induce the production of secondary compounds that act as defenses against herbivory (see Chapter 11). Plants with yellow or white flowers were preferred by pollinating bees over those with pink or bronze flowers. However, those same plants were less able to induce defensive responses, resulting in greater herbivore damage. The two traits are genetically linked, so changes in flower color affect defenses, and vice versa. In a study of *Brassica rapa* (field mustard, Brassicaceae), Strauss and colleagues (1999) used artificial selection to create one population with high concentrations of the defensive compound myrosinase and another population with low concentrations. Those with high concentrations were much more resistant to attack by flea beetles and experienced much less damage to flowers. However, pollinators tended to spend more time at flowers of low-myrosinase plants; other studies have shown that pollination success in *B. rapa* increases with time spent by the pollinators. Thus, herbivores and pollinators can exert selection in opposing directions.

Plant-pollinator interactions can also be involved in another kind of complex interaction called a trophic cascade, in which predators reduce the population size of their prey, setting off still more changes at the next trophic level. A study by Tiffany Knight and her associates (2005) revealed such an indirect interaction between a plant and a fish. They studied bee pollination of *Hypericum fasciculatum* (St. Johnswort, Hypericaceae) on the shores of a set of Florida ponds. Some ponds contained fish, while others did not. Many of the fish species were predators of dragonfly larvae, and dragonflies are important predators of bees. They found that *H. fasciculatum* growing on shores of ponds with fish had significantly more visits by pollinators. A pollen-addition experiment (see Box 6D) revealed that plants growing by fish-free ponds were significantly more pollen limited than plants growing by fish-containing ponds. While fish-plant interactions are unusual, it is likely that many other plant-pollinator interactions can be affected by trophic cascades.





Bob Peterson/CC BY-SA 2.0

Fish and *Hypericum fasciculatum* are involved in a complex interaction. Solid arrows connecting organisms show direct interactions; dashed arrows show indirect interactions. The plus and minus signs give the direction of the effect—for example, more fish mean fewer dragonfly larvae, but more pollinators and plants. (After T. M. Knight et al. 2005. *Nature* 437: 880–883.)

examples like the fish-plant connection described there are certainly unusual, it is likely that there are many other complex interactions involving pollinators. There has been much recent attention in the popular press, and to some extent in the scientific literature, about declines in pollinator species and some of their potential consequences. These are discussed in Box 6C.

### Aquatic plants have special adaptations for pollination

Some flowering plants, such as water lilies (*Nymphaea* or *Nuphar*, Nymphaeaceae), live in water. In most of these plants, the flowers are borne above the water's surface, and wind or insects pollinate the plants. But in plants of some 31 genera in 11 families, the flowers are borne underwater. As with wind pollination, it is inherently unlikely that pollen grains in water will find receptive stigmas. Furthermore, tides and currents are likely to make pollen dispersal unpredictable.

Aquatic plants have generally reduced this unpredictability by making their pollen dispersal units larger (and thus more likely to contact a stigma) and, in many cases, by dispersing them mainly at the water's surface (so that they travel in a two-dimensional plane, rather than a three-dimensional volume). A remarkable example is the freshwater plant genus *Vallisneria* (Hydrocharitaceae; Figure 6.10). In this genus, staminate (male) flowers containing pollen are released underwater and float to the surface. The petals curve downward to keep the flower floating on the water, and two sterile anthers act as tiny sails. The flowers are thus blown around until they encounter pistillate (female) flowers containing ovules, which are held at the surface by long pedicels (stalks). The pistillate

## BOX 6C Is There a Pollinator Crisis?

There has been much attention recently to the potential consequences of declines in populations of pollinating insects. In general, there are few data on the geographic extent or magnitude of the decline of insect populations, and the scientific evidence on pollinator declines is sparse. In reviewing data on insect declines, David Wagner (2020) reported evidence that several bumblebee species are near extinction or have become extinct in Europe, North America, and Asia, and many others are experiencing declines. There have been few studies of the 20,000 or so solitary bee species in the world.

In a study of 187 bee species in the northeastern U.S., Bartomeus and associates (2013) did not find a general pattern of species loss, though they did find evidence of population collapse in three bumblebee species and six species of solitary bees. A more recent study in New Hampshire reported evidence for decline of eight species of solitary bees (Mathiasson and Rehan 2019). A study of 139 bee species in the United Kingdom found declines for 25% and increases for about 10% of the species studied (Powney et al. 2019).

Many of these declines appear to be associated with land use patterns. There have been no substantial changes in numbers or diversity of euglossine bees at three forested locations in Panama (Roubik 2001), but there have been some declines at other Neotropical sites associated with deforestation, agricultural expansion, and urbanization (Frankie et al. 2009; Nemésio 2013). Wagner (2020) reports similar results from the temperate zone: declines associated with intensive agriculture, but not in relatively undisturbed natural sites. There were no reports of widespread pollinator limitation in intact natural communities.

Many reported declines have been in domestic honeybees. Commercial hives are subject to several diseases, which sometimes spread rapidly and kill many bees. There is ongoing research on whether recent declines in honeybees are unusual in this regard or whether they are caused (or magnified) by anthropogenic factors like insecticides or climate change.

A related concern involves differences in the ways that plant and pollinator species may respond to climate change. A "phenological mismatch" could occur if a pollinator responds to climate change such that its peak abundance is earlier or later than the flowers it normally pollinates. While it seems clear that this can happen, it is not clear that there is a general cause for alarm. Most insects visit multiple plant species, and most plants are pollinated by multiple insects. Particular relationships may well become less important or even disappear, and new ones may emerge. A recent review by Heather Kharouba and associates (2018) found little evidence of any trend since the 1980s in either increasing or decreasing synchrony between plants and pollinators.

Many well-intended people and organizations promote the idea of using more native plants in gardens, with a view to assisting pollinator species. A recent analysis of this idea by Anna Johnson and associates (2017) suggests caution. Greatly increasing the populations of urban wildflowers may have negative effects on extant populations of native plants in cities due to increased competition, reduced quality of pollinator services, and other effects. An analogue is the well-studied increase in urban populations of noisy miners in southeastern Australia. These aggressive birds have increased in urban areas largely because of native plantings; the miners, however, are able to drive out many other bird species. The lesson is that well-intended ecological manipulations can have unexpected negative effects.

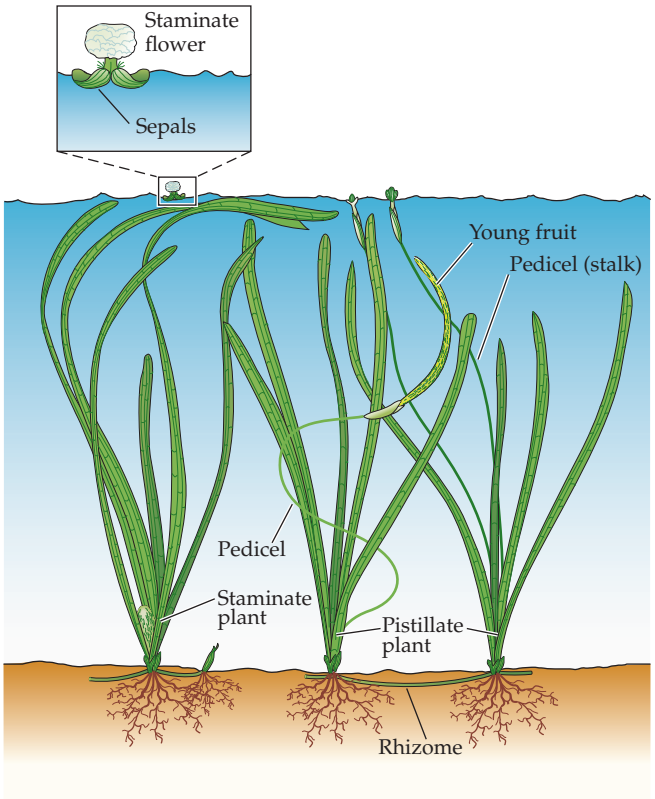



**Figure 6.10** The aquatic plant *Vallisneria* (Hydrocharitaceae) has special adaptations for pollination at the water's surface.

flowers create small depressions in the water surface, and the staminate flowers fall in, pollinating the pistillate flowers.

In most aquatic species, pollen, rather than a whole flower, is the dispersal unit. In many cases, the pollen is sticky and tends to form rafts that float on the surface until they contact pistillate flowers. Pollen release in some subtidal marine species occurs only near very low (spring) tides, which also increases the chances of contact between pollen and stigmas. In species with entirely underwater pollination, such as the turtle grass *Thalassia* (Hydrocharitaceae), pollen tends to be elongated or embedded in strands of mucilage—features that also increase the chances of contacting stigmas.

A remarkable case of pollen elongation occurs in the seagrass genus *Zostera* (Zosteraceae). The pollen germinates before release, forming a pollen tube in the process. The pollen tube greatly increases the length of the pollen and enhances its chance of becoming caught in a pistillate flower. Pistillate flowers of aquatic plants also show a number of features, such as unusual stigma shapes, that act to increase the probability of contact with pollen.

## 6.6 Plants Have Complicated Mating Systems

Plants have remarkably complex sex lives. In this section, we look at plant **mating systems**—biological factors that govern who mates with whom. Discussion of mating systems is sometimes confusing for students until they realize that it involves two different issues. The first involves **inbreeding**, which refers to whether individuals pollinate themselves or close relatives. Biologists also use *inbreeding* to refer to the consequences of mating between close relatives. The second issue concerns **sexual selection**, variation among compatible mates in their mating success.

### Inbreeding is mating between close relatives

Inbreeding can be important mainly because mating with close relatives (including self-fertilization) increases the chance of offspring having two copies of deleterious recessive alleles. The result is **inbreeding depression**, reduced fitness among the offspring. But plants (and other organisms) that inbreed extensively suffer from little or no inbreeding depression. The reason is straightforward: if there is much inbreeding, deleterious recessives tend to occur as homozygotes, and they are quickly selected against. As a result, over a sufficient period of time these populations tend to eliminate most of the deleterious alleles and most inbreeding depression.

There is almost always some inbreeding among plants because individuals in a population are likely to share common ancestors. As in a small human village, everyone is a cousin of some sort. But strong inbreeding occurs when plants pollinate themselves or mate mainly with closely related neighbors. Reproductive traits can promote or limit such inbreeding.

If inbreeding results in inbreeding depression, it is advantageous to limit matings with relatives. Self-incompatibility (SI) systems can accomplish this. There are two general types of SI systems. Typically, the genetic loci involved in self-incompatibility are referred to as *S* loci, and individual alleles are numbered $S_1$, $S_2$, and so forth. In **gametophytic SI**, the pollen grain's haploid genotype at the *S* locus prohibits it from growing down the stigma and style of a sporophyte sharing the same allele. By contrast, in **sporophytic SI**, the diploid genotype of the pollen parent prohibits it from mating with any sporophyte sharing either of its *S* alleles. The two systems act differently because of the difference in the time the *S* alleles are expressed. In gametophytic SI, many matings are "semicompatible": if pollen from an $S_1S_2$ plant is deposited on the stigma of an $S_1S_3$ plant, the $S_2$ pollen will grow and can fertilize both $S_1$ and $S_3$ ovules. In sporophytic SI, pollen from an $S_1S_2$ plant

cannot fertilize any ovules on an $S_1S_3$ plant or an $S_2S_3$ plant. Thus, while self-fertilization cannot occur under either system, sporophytic SI limits mating with more types of relatives than does gametophytic SI.

Is inbreeding bad for populations? Not necessarily! If there is little inbreeding depression, inbreeding can allow for local populations to maintain favorable combinations of genes. In fact, where localized adaptation is important, **outbreeding depression**—reduced fitness due to mating with unrelated plants—can occur. On the other hand, outcrossing can be advantageous when populations benefit from new combinations of genes or when inbreeding depression is severe. Accordingly, some species are obligately **self-fertilizing** (individuals can only pollinate themselves), while other species are **outcrossing** (they have mechanisms preventing self-fertilization and sometimes matings with other close relatives), while others fertilize their own ovules some of the time and mate with other plants some of the time. This multiplicity of mating patterns is accomplished by a variety of mechanisms. The conditions under which one mating system or another should be favored by natural selection constitute an area of active research, both theoretical (e.g., Barrett and Harder 1996; Charlesworth 2006; Busch and Delph 2011; Karron et al. 2011) and experimental (e.g., Spigler and Ashman 2011; Gervasi and Schiestl 2017; Barrett 2019; Delph 2019).

## Plants may vary in gender

Gender expression in plants is complex and varied. Some individual angiosperm plants are **cosexual**; that is, they function as both males and females simultaneously. The most common form of cosexuality—and the most common type of gender expression among plants—is **hermaphroditism**. For example, 75% of plants in Restinga de Jurubatiba National Park, Brazil, are hermaphrodites, and surveys in many other regions show similar proportions (Matallana et al. 2005). Hermaphroditic individuals have **perfect flowers** containing both functional stamens and functional stigmas. Under **monoecy**, individual plants have some **staminate flowers** (flowers with functional stamens only) and some **pistillate flowers** (with functional stigmas only); sometimes perfect flowers also occur on the same individual. By contrast, under **dioecy**, at least some plants in a population have only pistillate or only staminate flowers. Dioecious individuals cannot self-fertilize but may still be able to mate with many close relatives.

Gymnosperms also exhibit variation in gender expression. Pines (*Pinus* spp., Pinaceae) and many other conifers have male cones and female cones on the same tree. Other species, like ginkgos (*Ginkgo biloba*, the only living species in the Ginkgophyta), a broad-leafed, deciduous gymnosperm, have separate male and female trees. The evolutionary ecology of gender expression in gymnosperms has received almost no research attention, however, in contrast to the large literature on gender expression in angiosperms.

These three systems of gender expression can exist in combination with one another. Under **gynomonoecy**, both pistillate and perfect flowers occur on the same individual. By contrast, under **andromonoecy**, both staminate and perfect flowers occur on the same individual. Andromonoecy is much more common than gynomonoecy. Under **gynodioecy**, some plants in a population have only pistillate flowers, while others have either perfect flowers or a mixture of staminate and pistillate flowers. By contrast, under **androdioecy**, some plants in a population have only staminate flowers, while others have either perfect flowers or a mixture of both staminate and pistillate flowers. All of these combinations are unusual, but gynodioecy is much more common than androdioecy.

Still another form of gender expression is **sequential hermaphroditism**: beginning life as one gender (almost always male) and then switching to the other. In some sequential hermaphrodites the switch is not instantaneous, and individuals may function as hermaphrodites for a time.

Some cosexuals are capable of self-fertilization (sometimes within the same flower and sometimes only between flowers), while others are not. A review by Donna Vogler and Susan Kalisz (2001) of several hundred studies showed that about 25% of wind-pollinated cosexual plant species are mainly inbreeding; almost all other wind-pollinated cosexual species are strongly outcrossing. Among animal-pollinated species there is a different pattern: numbers are nearly evenly divided between strongly inbreeding species, strongly outcrossing species, and species with mixtures of inbreeding and outcrossing. Previously, plant ecologists had to rely on indirect estimates of inbreeding and outcrossing using the kinds of pollination experiments described in Box 6D. Today, using genetic markers, we can directly estimate outcrossing rates and even determine the parents of a particular seed or seedling in many cases (Smouse and Sork 2004; see Box 9B).

What is even more remarkable is that some of these different forms of gender expression can occur simultaneously within a single plant population. Within a population of hermaphrodites, not all individuals function equally as both males and females. Some individuals may mainly contribute pollen and set few seeds, acting mainly as males. The pollen on other individuals may be inviable or unable to germinate; those individuals are functionally strictly female. When individuals that appear to us to be

## BOX 6D Pollination Experiments

Many questions about plant reproductive ecology can be answered with very simple experiments. Often, we have two questions about a particular species: Are the plants self-incompatible? Is pollen a limiting resource? To answer these questions, we can do various combinations of three types of manipulations: cover flowers or inflorescences with mesh bags to exclude pollinators, remove stamens from (emasculate) plants, and hand-pollinate plants. The results of these treatments are then compared with results from unmanipulated plants, which allow us to estimate the typical rate of seed set. Bagged flowers indicate how many seeds in the unmanipulated plants resulted from outcrossed pollination. Emasculated plants show how many seeds in the unmanipulated plants resulted from self-pollination. If we get no seeds from bagged flowers, hand pollination with pollen from the same or different individuals tells us whether the species is self-incompatible or simply requires animals for pollination. Doing both bagging and emasculating tells us whether the species is capable of producing seeds by agamospermy.

A related question is whether pollen is a limiting resource. At first glance it might seem that this could be answered with the same sort of experiment: if hand-pollinated plants produce more seeds than naturally pollinated plants, the plants are pollen limited. However, this is too simple. One reason is physiological: A number of studies have found that hand-pollinated plants may produce more seeds in the current season, but fewer at subsequent times. Evidently some plants given much more pollen than usual may use stored resources to increase seed production. A study by James Thomson (2019) on *Erythronium grandiflorum* (glacier lily, Liliaceae) showed that pollen limitation that seemed apparent on some time scales vanished on a longer time scale, and flowers that open at different times during the season faced different selective regimes. Pollen availability and the consequences of adding pollen also can vary among microhabitats in a population. An influential paper by Tia-Lynn Ashman and associates (2004) concluded that apparent pollen limitation can be a result of random fluctuations around a pollen-resource equilibrium, an adaptation to stochastic pollination environments, or a chronic syndrome caused by an environmental perturbation. Assessing pollen limitation, then, requires a demographic analysis that asks how increased availability of pollen changes the population growth rate of a population (see Chapter 8). Since pollen availability can fluctuate naturally among years, answering this question may be challenging.

hermaphrodites actually are functionally unisexual, we speak of **cryptic dioecy**. Thus, the functional gender of a plant may be very different from its apparent physical gender; an individual's male reproductive output may be quite different from its female reproductive output.

Since variation in sexual expression can greatly affect the genetic composition of populations, it should be no surprise that angiosperm plant gender has been the subject of much study among evolutionary ecologists, conservation biologists, and population geneticists. Eric Charnov (1982) developed the basic theory that has motivated many studies of the evolution of gender. The central idea is that of fitness maximization: In different settings, fitness may be maximized by increasing the allocation of resources to both male and female functions, or by increasing allocation to one at the expense of the other and resulting in the evolution of the mating system from individuals acting as both males and females. The genetic basis of variation in gender is understood in some plant species; there are well-known cases in which the controls are quite simple. Because monoecy and gynodioecy are thought to provide two separate evolutionary routes from hermaphroditism to dioecy, these mating systems have been the focus of much more study than one might expect, based on their relative rarity. A recent study of evolutionary patterns among multiple plant lineages by Emma Goldberg and collaborators (2017) found no general trend across lineages for evolution of more differentiation (e.g., more dioecy) or less (more hermaphroditism); selection acted differently in each plant taxon.

Mating systems are not always fixed across a species. In a survey of 741 populations of 105 species, Michael Whitehead and associates (2018) found that nearly two-thirds of species had at least one outcrossing population. Some species were nearly all outcrossing, some nearly all inbreeding, and some had populations with almost every possible rate of outcrossing between the two. *Mammillaria dioica* (pincushion cactus, Cactaceae), from southern California, is gynodioecious in some populations and trioecious (having males, females, and hermaphrodites, a very rare condition) in others.

Discussion of plant gender clearly involves a lot of terminology. To make matters worse, different biological disciplines use these terms to mean somewhat different things. Zoologists usually use *hermaphroditic* to mean that both male and female sex organs are produced by the same individual (since there is no animal analogue of a flower). Geneticists use *monoecious* to mean the same thing. This terminology is not imprecise. Scientists all mean something quite specific when using these terms, but each subdiscipline means something a bit different. In practice, this is not usually a problem: one just needs

to be aware of these differences and use the context to determine the meaning an author intends.

## Competition occurs among plants and among pollen grains

As Oscar Wilde remarked, "Nothing succeeds like excess." Many plants appear to have followed this advice. Almost all plants start out with many ovules that never turn into viable seeds. These excess ovules may exist for three reasons: because pollen is a limiting resource, because the flowers have another function, or because the plant is hedging its bets (producing more ovules than can usually be matured, because under unusually favorable conditions these ovules can be matured). We consider the first two reasons here; bet hedging is discussed more fully in Chapter 8.

For animal-pollinated plants, pollen may be limiting because of a lack of pollinators (see Box 6D). If there are more flowers than pollinators to visit them, or if the pollinators are inefficient at transferring pollen, the plants may be competing for pollinators. This competition can occur both among plants of the same species and among species in a community. These two types of competition can have different consequences for flowering patterns (see Chapter 8).

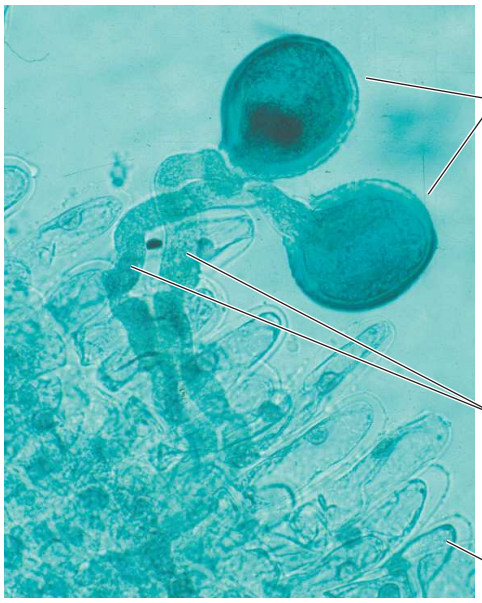
Competition for pollinators among individuals of the same species is analogous to competition among males (or females) for mates in animals. When this concept was first developed in the 1980s—notably by Mary Willson (Burley and Willson 1983; Willson 1983)—it was met with some skepticism because the analogy between pollen grains on a stigma and, say, male birds competing for the attention of females seemed strained to some botanists. However, there is now a large body of evidence that pollen donors and pollen grains do indeed compete. Some individuals in a population may end up with many matings (lots of pollen), while others may get very few. Most evolutionary ecologists consider sexual selection to be a subset of natural selection, although common usage sometimes treats this area of selection theory separately.

Courtesy of Mary F. Willson

Mary Willson

There are two broad categories of sexual selection: male-male competition and female choice. In plants, male-male competition is thought to result in extremes of floral displays, such as large inflorescences in which most of the flowers function only to attract pollinators. These flowers are thus functionally male; they reproduce only by having their pollen fertilize the ovules of flowers on other plants, and not by being pollinated themselves. Such cryptic dioecy is particularly common in legumes (Fabaceae) and in the Bignoniaceae. Competition among pollen grains after they land on a stigma can also result in sexual selection. There is a race for the ovules as pollen tubes grow down through the style (Figure 6.11). Experimental studies have shown that in some species, the maternal sporophyte chooses among potential pollinators through biochemical interactions that affect pollen germination and the rate of pollen tube growth. When this results in more fit offspring, it is adaptive (Marshall and Folsom 1991). Some maternal sporophytes, such as *Cucurbita pepo* (zucchini, Cucurbitaceae), have been found to selectively abort certain seeds or fruits before they mature (Stephenson and Winsor 1986).

(A)



Courtesy of Allison A. Snow

(B)

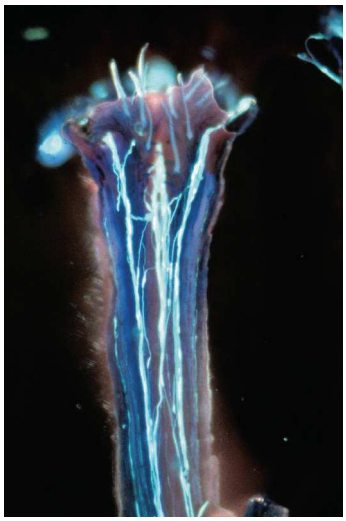

Courtesy of J. Ghazoul

**Figure 6.11** When pollen grains germinate, their pollen tubes grow, penetrate the stigmatic surface, and grow down the style toward the ovules. (A) Two germinated pollen grains of *Erythronium grandiflorum* (glacier lily, Liliaceae). (B) Pollen tubes growing in the styles of *Shorea siamensis* (Dipterocarpaceae), a rainforest tree from Thailand. To visualize these tubes, the sample is stained with aniline blue and viewed under ultraviolet light with a fluorescence microscope.

## Most mating is between neighboring individuals

Some degree of inbreeding is common in plant populations. The reason is simple: the largest proportion of matings occurs among neighboring plants, which tend to be relatives because most seeds do not travel far.

Most pollen ends up near the sporophyte on which it originated. In animal-pollinated species, there are several reasons for this pattern. First, most plant species have clumped spatial distributions. Moreover, some patches of plants are typically more rewarding to pollinators than others, and the pollinators usually concentrate on the more rewarding patches, reducing the average distance traveled between plants. Finally, most pollen is transferred to the next flower visited (Richards 1986). In *Lupinus texensis* (bluebonnet, Fabaceae), for example, most pollinator movements between plants are less than 0.6 m, and over 90% of movements are less than 1.2 m (Figure 6.12). Gene flow does not precisely match pollinator movement (again suggesting that pollinator visitation is not equivalent to pollination), but it also declines rapidly with distance from the source.

Pollen movement by wind also usually results in most pollen being distributed near its source, but for different reasons. Wind usually blows in short gusts, most of which have low velocity, so most pollen movements are over short distances. In trees, pollen is released at varying heights, and the pollen released at the greatest height usually travels the farthest (Richards 1986). Tree shape and proximity to conspecifics will affect the distance pollen travels. In pines (*Pinus*, Pinaceae) female cones are disproportionately clustered at the top of the tree and male cones at the bottom, which may tend to reduce self-fertilization. Since most wind-pollinated plants (such as grasses) occur in large stands, most pollen tends to land on neighboring plants. In some populations, plants have variable characteristics that can affect the distance that pollen travels. In *Plantago lanceolata* (ribwort plantain, Plantaginaceae), Stephen Tonsor (1985) found that individuals differed in the extent to which they produced pollen as single grains or as clumps of 2, 3, or 50 or more grains. A wind tunnel experiment showed that single grains travel much farther than larger clumps (Figure 6.13). Thus, some individual plants are more likely to mate with distant individuals, and some are more likely to mate with close neighbors.

## Plants may mate preferentially with individuals with similar phenotypes

Plants with similar phenotypes—for example, those with similar flower colors or flowering times—may mate with one another more often than they would by chance alone; this is called **assortative mating**. In the case of flowering time, this is especially obvious: early flowerers are more likely to mate with other early flowerers than with late flowerers. Assortative mating for flower color

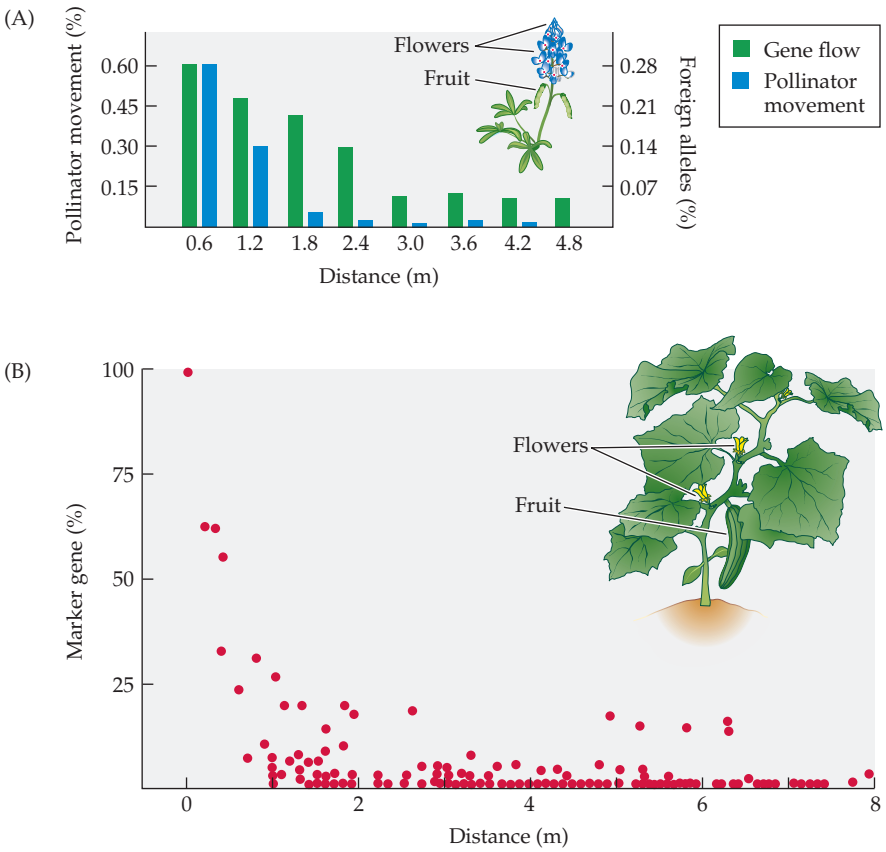



**Figure 6.12** (A) The distances traveled by pollinators and gene flow in seedlings of *Lupinus texensis* (bluebonnet, Fabaceae). While the general distribution patterns are the same, using pollinator movement alone to estimate gene flow would underestimate it, because pollen from one flower may be deposited on the second, or third (and so forth), flower that is visited. (B) The appearance of marker genes in seedlings of *Cucumis sativus* (Cucurbitaceae). The tail of the dispersal distances is wide; in this case as well, pollen may be carried over from flower to flower. (A after B. A. Schaal. 1980. *Nature* 284: 450–451; B after S. N. Handel. 1983. *Evolution* 37: 760–771.)

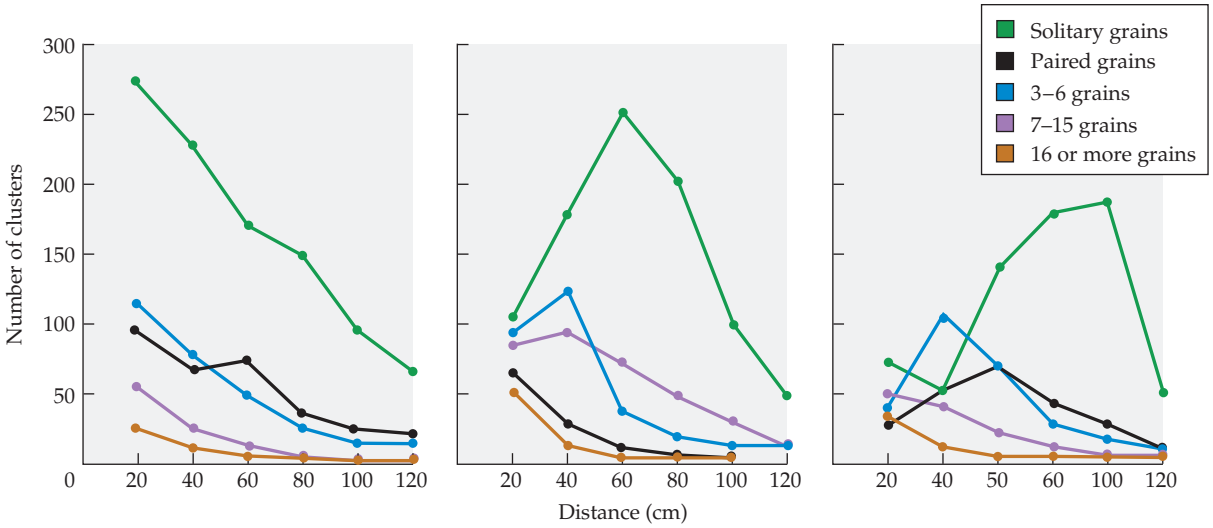



**Figure 6.13** Variation in the distance traveled by *Plantago lanceolata* (Plantaginaceae) pollen clumps in a wind tunnel as a function of size in three trials. In each trial, a plant was placed at one end of the wind tunnel, and rows of Vaseline-coated microscope slides were placed at 20 cm intervals along the floor of the tunnel. Each line shows the number of solitary grains or clumps that traveled each distance. (After S. J. Tonsor. 1985. *Evolution* 39: 775–782.)

often occurs because many pollinators have preferences (sometimes on an individual basis) for colors. A bee, for example, may travel mainly from yellow flower to yellow flower or red flower to red flower.

When individuals tend to mate with others with dissimilar phenotypes, we refer to **negative assortative mating** or **disassortative mating**. Two common forms of this phenomenon in plants are **self-incompatibility** (meaning that matings can occur only between individuals with certain genetic differences) and **heterostyly** (flowers have styles and stigmas of different lengths, and matings within morphs occur only rarely; Figure 6.14). These types of negative assortative mating cause obligate outcrossing.

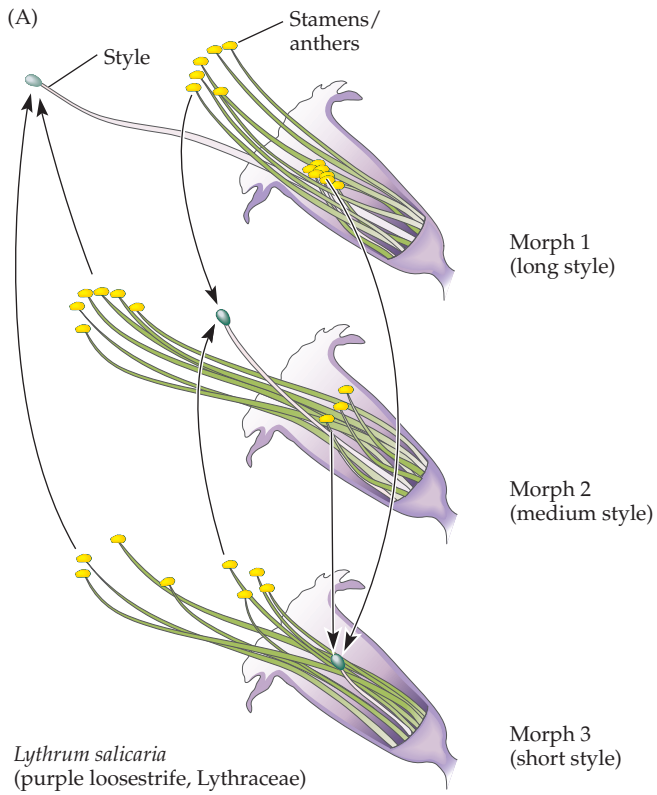
## Fitness can depend on a population's composition

**Frequency-dependent selection** occurs when the fitness of a genotype depends on whether it is common or rare in a population. An example is seen in the evolution of mating systems in *Lythrum salicaria* (purple loosestrife, Lythraceae). This species is an invasive weed in wetlands of northern North America (see Chapter 13). It is still in the process of invading the continent, which makes it an ideal species for studying evolution in action. In the study described here, scientists were able to directly observe populations evolving over a 5-year period.

*Lythrum salicaria* has an unusual mating system, called tristyly. The style is the female reproductive part that supports the stigma, on which pollen lands to fertilize the ovules (see Figure 6.6). This mating system gets its name because each of the plants in a population has one of three different forms of the style: short, medium, or long (see Figure 6.14A). The form of the style is genetically determined. The styles are matched by three stamen lengths (the stamen supports the anthers, which hold the pollen). A single individual has one style length and the other two stamen lengths. This system is a way of enforcing outcrossing: animal visitors do not transfer pollen between individuals with the same style length.

Frequency-dependent selection comes about because of this inability of individuals to mate with other individuals of the same form. Imagine a purple loosestrife population with 99 long-styled individuals and 1 short-styled individual. That short-styled individual would be able to mate with all the other plants in the population, as it could fertilize all of them with its pollen. The long-styled plants would be limited to pollinating just the one short-styled individual. So, in the next generation, the genes of the short-styled individual would end up in the offspring of all 100 plants (in its own offspring through its female gametes and in those of the other 99 plants through its pollen). In contrast, each of the long-styled individuals would have offspring through at most 2 plants (themselves and the short-styled plant). As a result, the genes of the short-styled plant would increase in frequency. In general, the rarest mating type (here, the one with the least common style length) will always enjoy this sort of mating advantage. Our purple loosestrife population would therefore continue to evolve until all three morphs were at equal frequency, with each morph constituting one-third of the population.

Just such a pattern of evolutionary change was observed by Christopher Eckert and associates (1996). They surveyed 24 newly established populations of *L. salicaria*



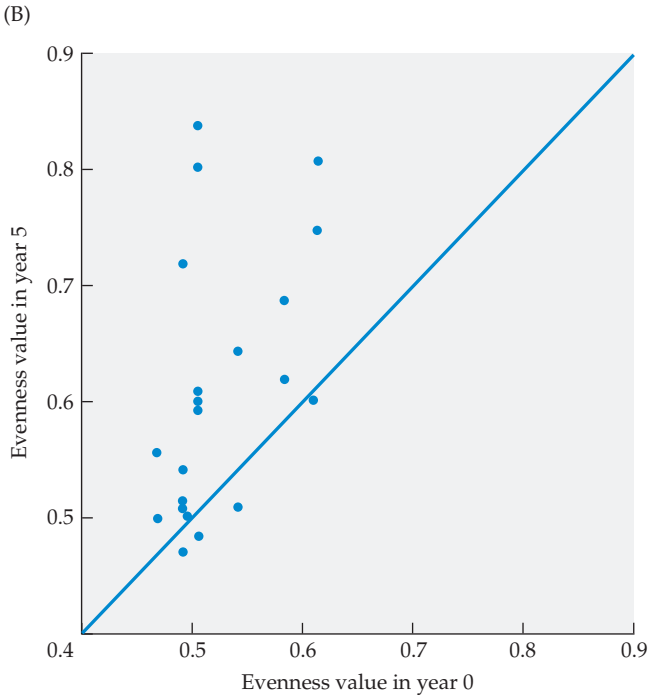




**Figure 6.14** (A) Tristyly in *Lythrum salicaria* (purple loosestrife, Lythraceae). The plant has three different forms, or morphs, of stamen and style, a system that promotes outcrossing. An individual with a given style length can mate only with one of the other two morphs. (B) Change in style length evenness over a 5-year period in 24 populations of *L. salicaria*. An evenness value of 1 would indicate that the three morphs were equally frequent. In all but 4 of the populations, the frequencies were more even in year 5 than in year 0, indicating that the rare morph had increased in frequency. (B after C. G. Eckert et al. 1996. *Heredity* 77: 581–588.)

across the Canadian province of Ontario. During the establishment of each population, by chance alone, some forms were at high frequency, while others were at low frequency. In some populations it was the short-styled form that was rare, in others it was the medium-styled form, and in still others it was the long-styled form. The researchers predicted that over time, the populations should evolve to an equal frequency of the three forms. They tested this prediction by returning to the populations 5 years later. The prediction was met: in all but four cases, the rarest form had increased in frequency (see Figure 6.14B). Most of the populations were still not yet at equilibrium, with each form making up a third of the population; that will require several more decades. However, the populations had consistently evolved in the predicted direction, something that would be highly unlikely to occur by chance.

This example is an instance in which natural selection was strong enough that we could observe it in action in a short time. Until recently, selection of this strength was thought by most ecologists and evolutionary biologists to be rare, but there is now much evidence that it occurs fairly commonly. Part of the strength of this study is that this change was observed across a number of populations. Thus, it is highly unlikely that the change was due to anything other than natural selection.

## Mating systems have other important consequences

Self-pollination and nondispersing seeds often co-occur in plants. If a particular genotype is successful in a particular site, it may be advantageous to produce some nondispersing offspring of the same genotype. This may be part of the reason why some plant species have some **cleistogamous** flowers (which never open and always self-pollinate), but the evidence is equivocal. In some species, cleistogamous flowers are located near ground level; in annual grasses of the genus *Amphicarpum* (Poaceae), they are actually underground. More-familiar open flowers are **chasmogamous**. Cleistogamous and chasmogamous flowers may be produced simultaneously, as in many species of the

temperate zone herbs *Viola* (Violaceae), *Oxalis* (Oxalidaceae), and *Impatiens* (Balsaminaceae). In some species, cleistogamy appears to be environmentally determined. Flowers of the cactus genus *Frailea* (Cactaceae) are cleistogamous except in extreme heat and so are considered to be facultative selfers. In several British plants, including *Lamium amplexicaule* (Lamiaceae) and *Hesperolinon* (Linaceae), flowers produced in the dry season are cleistogamous (Richards 1986).

Plant mating systems often strongly affect the number of pollen grains that are produced by an individual or a population relative to the number of ovules (the pollen:ovule ratio). As Table 6.1 shows, self-fertilizing species tend to have much lower pollen:ovule ratios than outcrossing species, and the greater the assurance of selfing—as in cleistogamy—the lower the ratio.

In many cases, ecologists are motivated to study pollination and plant mating systems because they are curious about the underlying biology or because they are interested in natural selection on traits that are tightly linked to fitness. Recently, however, ecologists and evolutionary biologists have expressed increasing interest in other issues of broad importance where pollination and mating systems are key elements. These include studies of speciation and of conservation-related issues concerning landscape ecology, population fragmentation, invasive species, and the possibility of spread of genes from genetically modified organisms (GMOs).

Many traits related to plant mating systems, such as flower color and shape and flowering time (see Chapter 8), can strongly affect the chance of mating between any two individuals and thus can affect gene flow between populations that differ in these traits. A striking example of how such assortative mating can play a role in speciation was discovered by H. D. "Toby" Bradshaw and Douglas Schemske and colleagues (Bradshaw et al. 1998; Schemske and Bradshaw 1999; Bradshaw and Schemske 2003; Ramsey et al. 2003). In *Mimulus* (monkeyflower, Scrophulariaceae), a simple genetic change probably played an important role in the speciation event (see Chapter 9) separating *M. lewisii* from *M. cardinalis*. There are normally very few matings between these two species, which are usually pollinated by bumblebees and hummingbirds, respectively. Having shown that flower color is strongly influenced by a locus called *YELLOW UPPER* (*YUP*; Bradshaw et al. 1995, 1998), the researchers used a careful breeding program to insert the *M. lewisii* form of the gene into *M. cardinalis*, and vice versa. *M. cardinalis* with the *M. lewisii* allele had dark pink flowers (instead of their normal red flowers) and received 74 times as many bee visits as unmanipulated *M. cardinalis*, while *M. lewisii* with the *M. cardinalis* allele had yellow-orange flowers (instead of their normal pink flowers) and received 68 times as many hummingbird visits as unmanipulated *M. lewisii*.

**TABLE 6.1** The ratio of the number of pollen grains produced per flower to the number of ovules produced per flower in species with different levels of outcrossing

| Mating system | Number of species examined | Mean pollen: ovule ratio (standard error) |
|---|---|---|
| Cleistogamous | 6 | 4.7 (0.7) |
| Obligate selfer | 7 | 27.7 (3.1) |
| Facultative selfer | 20 | 165.5 (22.1) |
| Facultative outcrosser | 38 | 796.6 (87.7) |
| Obligate outcrosser | 25 | 5859.2 (936.5) |

*Source*: R. W. Cruden. 1977. *Evolution* 31: 32–46.

Populations of *Quercus lobata* (valley oak, Fagaceae), a distinctive and much-loved tree that grows in the Central Valley of California, have suffered serious declines in density, mainly due to severe reductions in recruitment (Brown and Davis 1991). A major concern is the possibility that remaining stands of *Q. lobata* will become reproductively isolated due to habitat fragmentation (see Chapter 15) and begin suffering from inbreeding depression. Using a set of molecular markers, Victoria Sork and colleagues (Sork et al. 2002; Austerlitz et al. 2004; Dutech et al. 2005) estimated that pollination was taking place mainly among trees within 50 to 88 m of one another. By studying the number of genes shared by adults at varying distances, they estimated that historically, pollination took place mainly within 353 m. Because evidence suggests that few acorns disperse very far, these results suggest that the flow of genes by pollen dispersal has declined greatly during the twentieth century.

Another consequence of habitat fragmentation can be a reduction in the availability of pollinators. Taylor Ricketts and colleagues (2004) studied bee pollination of coffee (*Coffea arabica*, Rubiaceae) crops in Costa Rica. Coffee plants will self-pollinate, but pollination by bees increases the amount of fruit set. Study sites within 100 m of forest remnants received significantly more visits from native bees than those farther away. The exotic honeybee *Apis mellifera* was less affected by distance from forest fragments, but honeybee visits varied dramatically from year to year.

People have been concerned about the consequences of gene flow between crops—especially crops modified with recombinant DNA (often called genetically modified, or GM, crops)—and their wild relatives. The major concern is the possibility that "superweeds" might result from such hybridization: for example, a crop into

which an herbicide tolerance gene has been inserted might mate with one of its wild relatives and transfer the gene for herbicide tolerance. Evidence shows (Ellstrand et al. 1999; Ellstrand 2003) that crops frequently hybridize with their wild relatives, providing some basis for this concern. Interestingly, though, there are no documented cases of a transgene (a gene transferred by human technology from one species to another) spreading from a crop to its wild relatives, despite the extremely widespread use of GM crops in North America. Well over a hundred herbicide-tolerant populations of weeds have been reported, but all of these were the result of the strong selection imposed on the populations by the application of herbicides. A heated debate continues—in both the scientific and the public policy realms—over the potential for gene transfer between GM crops and wild plants. The principal argument against GM crops is one of caution: because there is a potential for gene transfer and therefore for creating serious environmental problems, critics argue, it is wise to err on the side of caution and limit or eliminate GM crops. The strongest counterargument is that such hybrids are likely to suffer severe fitness reductions, either because of the nature of the differences between crops and wild plants, or because traits modified with recombinant DNA are particularly likely to lead to reduced fitness and therefore will usually be selected against in wild populations (Crawley et al. 2001; Stewart et al. 2003; Kelly et al. 2005). Scientists arguing for use of GM crops have also advocated various measures to reduce the possibility of gene transfer (Trewavas and Leaver 2001; Stewart et al. 2003).

Hybridization between invasive species and their native relatives is also an issue of much concern. Several *Spartina* species (cordgrass, Poaceae) are problematic invaders in salt marshes around San Francisco Bay, California. In addition to competing with many native plants and animals (including several threatened species), they dramatically alter habitats, covering open mud flats and filling channels. A native species, *S. foliosa*, is threatened with local extinction because of hybridization with the invasive *S. alterniflora* (see Chapter 14 for more discussion).

## 6.7 Fruit and Seed Characteristics Affect Dispersal across Space and Time

Since most plants are rooted in place, they disperse as fruits, seeds, or related structures. These dispersal units are sometimes referred to collectively as **diaspores**. All flowering plants have fruits, and fruits have a tremendous range in size, shape, and number of seeds, characteristics of other tissues, color, odor, and chemical properties. A fruit is just a mature ovary, so even plants such as grasses have fruits, called **grains**, although these single-seeded fruits contain little maternal tissue beyond the seed coat of the seed itself and some associated tissues (such as awns and barbs) that serve in dispersal or protection. Fruits have two major functions: they protect the developing seeds, and they can affect the dispersal of the seeds. These functions often conflict with each other, and variation in their relative importance leads to much of the variation in fruit characteristics.

### The structures of seeds and fruits affect their dispersal

Many fruit characteristics are best understood as traits that affect the behavior of animals feeding on fruits and seeds. On the one hand, most seeds are quite nutritious, and seeds are therefore important animal foods. Most human caloric intake, for example, comes from the grains of grasses (e.g., rice, wheat, and maize). In fact, the rise of agriculture and the evolution of human civilizations depended on grain—the fruits of grasses. There is, therefore, a selective advantage to plants in protecting seeds from being eaten before they are mature—otherwise those offspring are simply lost. On the other hand, after seed maturity, there is often a selective advantage in attracting animal consumers to disperse the seeds. Plants have various ways of balancing these two functions. Some seeds can pass through animals without being destroyed; for example, while humans consume vast quantities of wheat and rice seeds, tomato seeds pass through our digestive tracts without any loss of viability.

Many fruits attract **frugivores** (fruit eaters) after seed maturity. Unripe fruits typically have little in the way of attractive odors, flavors, or colors, and sometimes these may even be repellent. Moreover, in many taxa, unripe fruits contain compounds that act as defenses; animals eating them tend to become sick, especially when they are eaten in large quantities (Herrera 1982a, b). You may have experienced this phenomenon yourself if you have ever eaten unripe fruit. As they ripen, however, most fruits change color, begin to give off attractive odors, and become palatable through the conversion of starches to sugars and other chemical and physical transformations.

If seed dispersal is to occur, animals must not spend too much time feeding at any one plant. If the plant and its fruits are too attractive to the animals, they will not move away, and the seeds will not be dispersed. How do plants encourage frugivores to move along? First, the fruits on an individual plant may mature asynchronously. For example, flocks of birds or troops of monkeys often travel from tree to tree, eating only the ripe fruits. Second, if ripe fruit remains very mildly toxic, an animal will eat only a few fruits of a given species at once.

Seeds can be dispersed by animals, by wind, or less commonly, by water. Animals can disperse seeds internally or externally. Seeds can be dispersed singly or in groups. Furthermore, seeds can be dispersed in time as well as in space. All of these factors have different effects on patterns of seed dispersal. These patterns are important because they determine the demographic and genetic structure of plant populations, the amount of competition within and among species, and species migration patterns.

There are two modes of dispersal by wind: some diaspores float on the breeze, and some flutter to the ground. Dandelion seeds (*Taraxacum*, Asteraceae), which children enjoy blowing around, are the archetype of the former sort (Figure 6.15A). This form of dispersal is very common among plants that colonize disturbed habitats. Many of them are annuals that survive by being very good dispersers, but poor competitors. We discuss such trade-offs in more detail in Chapter 7.

Maples (*Acer*, Sapindaceae) are a good example of plants that produce fruits that flutter to the ground (Figure 6.15B). Clearly, most maple fruits are unlikely to travel nearly as far as most dandelion seeds. Most maple seeds end up falling very close to the parent tree, although a few may be blown a substantial distance. Carol Augspurger and Susan Franson (1987) showed that the shapes of diaspores can have a large effect on how far they travel (Figure 6.16). Even after initial dispersal, significant diaspore movement by wind can occur. For example, seeds of *Daucus carota* (wild carrot or Queen Anne's lace, Apiaceae) disperse very late in autumn and can then be blown across the snow in winter.

Courtesy of Carol K. Augsperger

Carol Augspurger

Dispersal by animals involves more varied mechanisms and leads to more varied dispersal patterns. In species that rely on animal dispersal through frugivory, the seeds have thick, hard seed coats and are either regurgitated by the animal or passed through the digestive tract in its feces (Figure 6.15D). This mechanism not only disperses the seeds, but provides them with fertilizer. Some fruits contain a mild laxative to speed their passage through the gut, thereby preventing their digestion and speeding their dispersal. Some seeds require passage through the gut of an animal to facilitate germination. In other cases, the fruit may be eaten and the seeds spit out (as in watermelon: *Citrullus lanatus*, Cucurbitaceae, at least before the seedless hybrids became ubiquitous).

An interesting problem in frugivory was discovered in the 1970s when scientists noted that *Sideroxylon grandiflorum* (tambalacoque, Sapotaceae), a native of Mauritius (an island in the Indian Ocean), was near extinction.

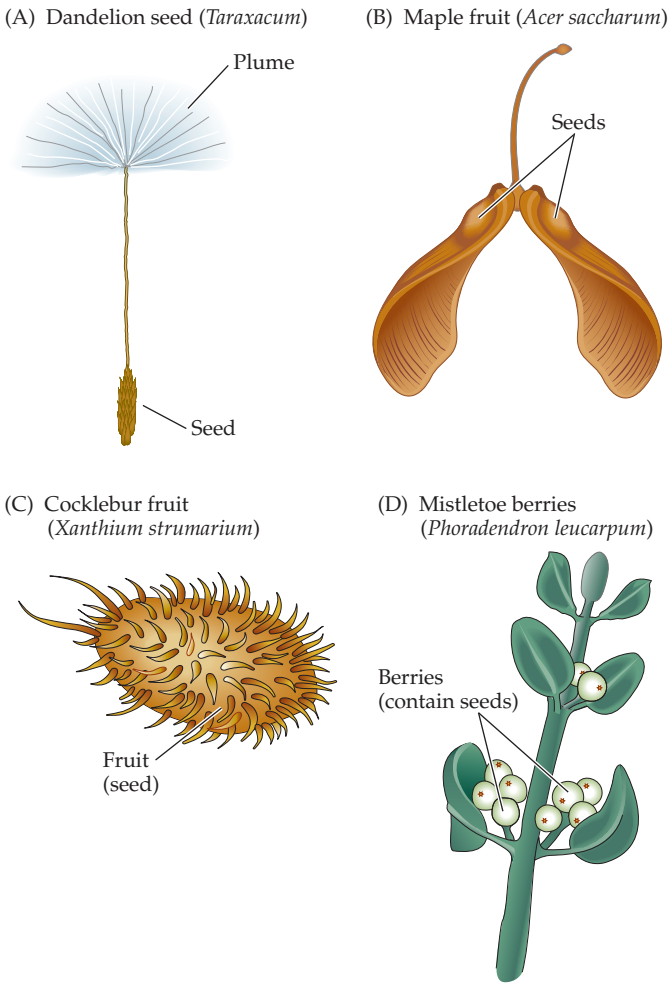



**Figure 6.15** Many diaspores have structures that facilitate dispersal by particular mechanisms. (A) Some wind-dispersed diaspores, such as the achenes of *Taraxacum* (dandelions, Asteraceae), float on the breeze. (B) Other wind-dispersed diaspores, such as those of *Acer* (maples, Sapindaceae), flutter to the ground. (C) Some animal-dispersed diaspores, such as this *Xanthium* (cocklebur, Asteraceae) fruit, are carried externally, stuck to an animal's fur or feathers. (D) Other animal-dispersed diaspores, such as the berries of the parasitic *Phoradendron* (mistletoe, Santalaceae), pass through the gut of a vertebrate before being dispersed. The *Phoradendron* seeds within the fruits are coated with a sticky mucilage that readily adheres to the branches of potential host trees; frugivorous birds often rub their cloacas on branches to remove the seeds from their bodies, thus effectively planting the seeds.

There were only a handful of individuals left, and these all appeared to be centuries old. The seeds of this plant have very thick coats, requiring **scarification** (abrasion) to germinate. An early hypothesis was that much of this scarification was provided by passage through the digestive systems of dodos (*Raphus cucullatus*, large

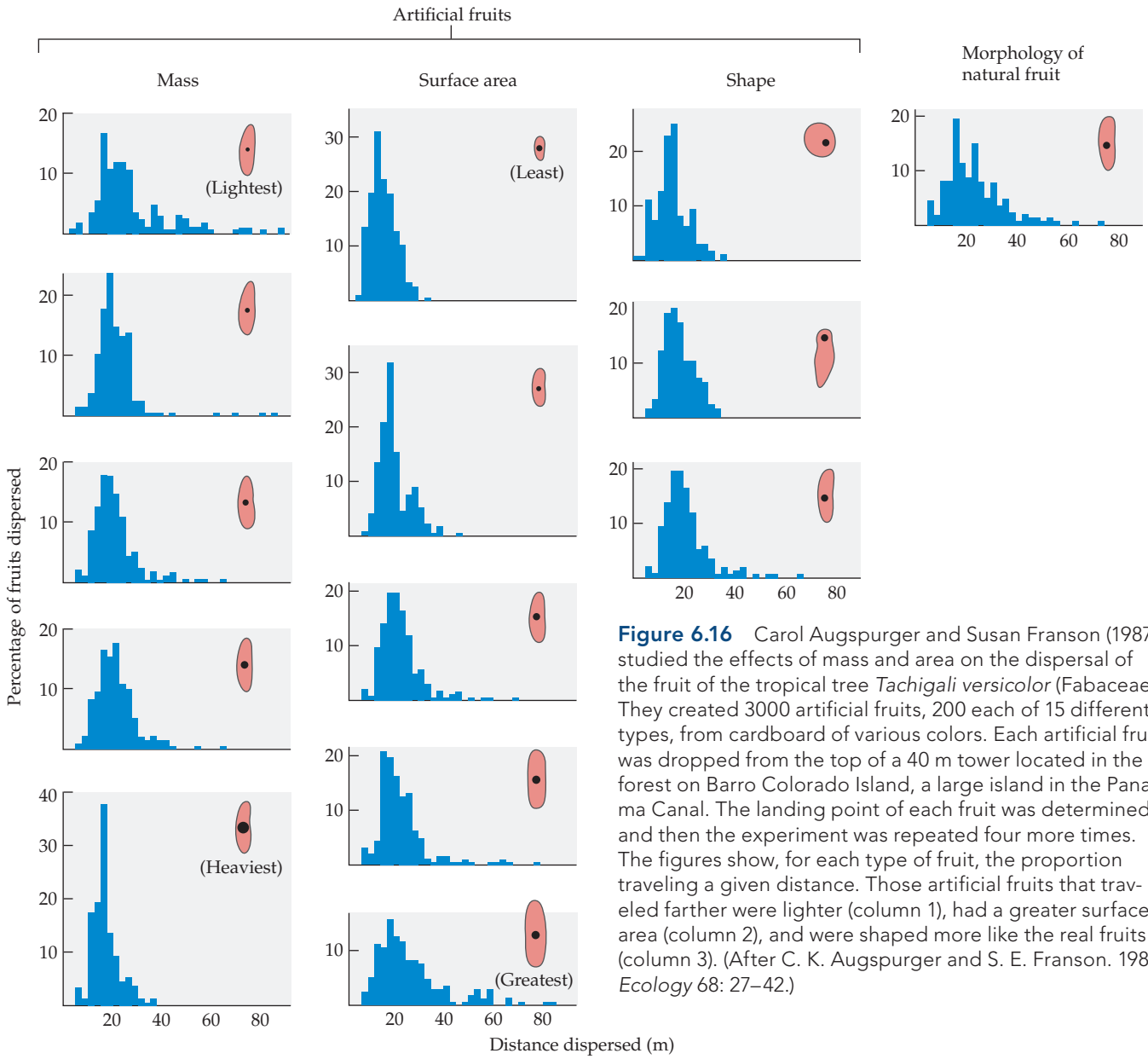



**Figure 6.16** Carol Augspurger and Susan Franson (1987) studied the effects of mass and area on the dispersal of the fruit of the tropical tree *Tachigali versicolor* (Fabaceae). They created 3000 artificial fruits, 200 each of 15 different types, from cardboard of various colors. Each artificial fruit was dropped from the top of a 40 m tower located in the forest on Barro Colorado Island, a large island in the Panama Canal. The landing point of each fruit was determined, and then the experiment was repeated four more times. The figures show, for each type of fruit, the proportion traveling a given distance. Those artificial fruits that traveled farther were lighter (column 1), had a greater surface area (column 2), and were shaped more like the real fruits (column 3). (After C. K. Augspurger and S. E. Franson. 1987. *Ecology* 68: 27–42.)

flightless pigeons), which were driven extinct in the seventeenth century by hunting by European sailors. As a result of the dodo extinction, the hypothesis went, the seeds were unable to germinate. Subsequent research has suggested alternative culprits: introduced monkeys readily eat the fruits before they are ripe, thus eliminating most recruitment; introduced pigs eat the fruits and any seedlings that might appear. Whatever the cause, the species remains near extinction, although as of the writing of this textbook, there are still surviving trees.

**Granivores** (animals that eat seeds or grains; see Chapter 14) sometimes act as important seed dispersers. Animals intending to eat seeds often drop them accidentally. Animals that bury seeds for later use also act as dispersers if they fail to return. Both birds and mammals cache seeds in this fashion. For example, *Pinus albicaulis* (whitebark pine, Pinaceae) in the mountains of the northwestern United States and southwestern Canada is mainly dispersed by granivorous birds, especially Clark's nutcracker. The seeds of this tree are not released from the cones when they are mature, and the tree depends on birds to remove and disperse them. The nutcrackers pry the seeds loose from the cones, remove them, and cache them, storing from 1 to 15 seeds together in each of many caches just under the soil surface. They later retrieve the seeds and eat them. Although Clark's nutcrackers have excellent memories, they occasionally forget where some of their caches are, or they

may be killed or driven away, leaving their uneaten seeds to germinate and grow into trees (Lanner 1996).

Sometimes the mode of animal dispersal can be subtle. Many species of ants collect seeds for food. In some plant species, however, the seeds are attached to a lipid body called an **elaiosome**. Instead of eating these seeds, ants carry the seeds to their nest, remove and eat the elaiosomes, and toss the seeds outside. Thus, seed dispersal by animals may result in a highly clumped pattern if many seeds end up in a refuse heap. Other seeds disperse by getting stuck in the hair or feathers of an animal (Figure 6.15C), which can be a very effective mode of long-distance dispersal.

While most seeds end up relatively close to the parent plant, a small percentage of seeds may be dispersed long distances, and this long-distance dispersal can have important demographic consequences. Note, however, that a long distance for one species could be a short distance for another. When we speak of "long-distance dispersal," we mean roughly that individuals are moving far enough from their parent population that they can no longer interact with individuals in that population (Jordano 2017). It helps to have a rough idea of the actual distances involved; for many plants, dispersing more than a few hundred meters can move them out of their parent population.

In Chapter 17, we discuss the migration of trees northward in North America following the retreat of the glaciers. Some tree species are estimated to have migrated as fast as 400 m per year. Since it takes several years for a tree to become large enough to produce seeds, migration at this rate must have depended on very long-distance dispersal events. However, estimating the probability of seeds establishing at very great distances from their parents presents a statistical difficulty: these events are, by definition, rare. Without too much difficulty we can use some well-established models, called **dispersal kernels**, to estimate the probability of a seed dispersing at various distances from the parent, but it is quite difficult to obtain enough data to estimate these probabilities for very long distance dispersal (Portnoy and Willson 1993). Interestingly, wind-dispersed species migrated faster than animal-dispersed species (see Table 17.1); one might expect the reverse, since animal-dispersed species often appear to have more long-distance dispersal events.

There are two major approaches to estimating the dispersal properties of seeds. The first involves mechanistic models of seed dispersal, generally by wind. Using equations from aerodynamics and data on the weight and shape of seeds, ecologists have developed models that have successfully predicted the distributions of observed distances traveled by seeds (Greene and Johnson 1989, 1996; Nathan et al. 2001). An important limitation of these models is that they must generally make restrictive assumptions about the behavior of wind—for example, that wind speeds are constant and that there is no turbulence. Since these assumptions do not generally hold, many ecologists use a different approach that involves asking what statistical description best fits the observed distribution of dispersal distances (Clark et al. 1998, 1999a). The central idea is simple: Imagine a single tree in a large open field. At a number of distances from the tree, place **seed traps**—containers designed to capture windblown seeds—and use the densities of seeds arriving at each trap to estimate the parameters for a dispersal kernel. A simple example of a dispersal kernel is a negative exponential curve, which assumes that the probability of dispersing $x$ meters declines with distance at a constant rate; kernels involving more complicated biological assumptions have also been developed. One can fit different dispersal kernels (based on different assumptions) to the data to ask which is best supported by the data (Figure 6.17). This approach can be readily extended from a single source plant to many source plants (such as trees in a forest) by using a widely used model-fitting approach that

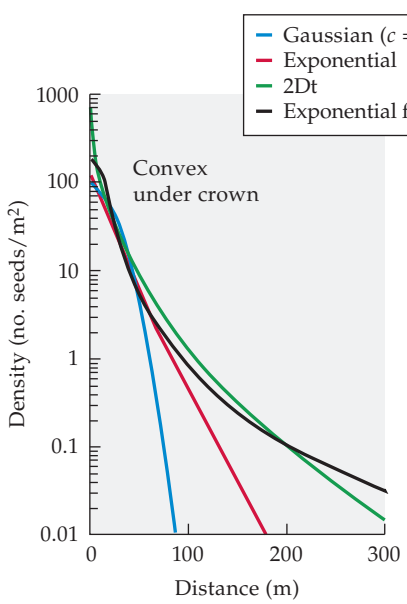



**Figure 6.17** Dispersal models for seed dispersal in *Acer rubrum* (red maple, Aceraceae) data (see Figure 6.15B). Red maples have spread into many new habitats throughout eastern North America during the twentieth and twenty-first centuries. James Clark and collaborators (1999b) fit four different dispersal kernels (probability models) to the data. The exponential (red) model assumes a constant probability of dispersing each distance. The Gaussian (blue) model assumes a bell-shaped distribution in two directions (only one is shown here). The exponential fat-tailed (black) and 2Dt (green) models both have "fat tails" (as the distance becomes large, they give probabilities greater than the exponential model). These fat tails mean that there is more long-distance dispersal, and the population can spread at a higher speed. The 2Dt model predicts more short-distance and long-distance dispersal than the exponential fat-tailed model; it is an extension of the Gaussian model that allows individual seeds to vary in dispersal properties. The 2Dt model is more recent but has rapidly become more widely used than the other models. (After J. S. Clark et al. 1999. *Ecology* 80: 1475–1494.)

asks what parameters for the model would make the observed data most likely. These parameters can now be estimated more accurately because genetic markers make it possible to track a given seed back to a particular adult (see Box 9B). These methods can have practical limitations, however, in the cost and labor necessary to genetically "fingerprint" each adult in a large population. A key problem is that most dispersal kernels are fairly similar to one another except in their tails—and their tails are difficult to estimate statistically.

Ecologists have developed methods to link dispersal kernels with demographic models (Neubert and Caswell 2000). While these approaches require knowledge of advanced mathematics, they have given us two important insights. First, many discussions of the rate of spread of invasive species have assumed that by knowing something about seed dispersal alone, one can predict which populations will spread quickly. This assumption turns out to be wrong: the rate of spread depends on both the dispersal kernel and the demographic model (see Chapter 8), and sometimes varying the demographic parameters has a larger effect on the rate of spread than varying the dispersal parameters. Second, in the long run, the rate of spread depends mainly on what occurs at the edge of a spreading population. Thus, efforts to promote the spread of a threatened species, or control an invasive species, should sometimes focus on populations where they are least dense and hardest to study.

Colonization of newly disturbed habitats or new islands necessarily depends on long-distance dispersal. Once a plant becomes established in such a habitat, short-distance dispersal is responsible for local population growth. For example, the seeds of *Cocos nucifera* (coconut, Arecaceae; Figure 6.18) disperse from island to island by floating in seawater—it is one of the few species whose seeds can survive this experience. On a given island, however, most seeds travel only short distances, as you might imagine would be the case for such large seeds (although they travel from tree to ground very fast!).

Dispersal is not always advantageous. Some plant species live in highly patchy environments, in which dispersal away from the maternal plant may be likely to cause a seed to land in completely unsuitable habitat. Many desert systems have this characteristic, and indeed many desert plants have mechanisms that limit seed dispersal. For example, seeds of the annual plant *Chorizanthe rigida* (Polygonaceae) in the Sonoran Desert remain on the dead maternal plant for long periods (as an aboveground seed bank), and when they do finally reach the ground, sharp spines on the diaspores ensure that most seeds do not travel far. Seeds of many cleistogamous flowers are borne at or even under the soil surface. Many plants on oceanic islands have lost seed dispersal mechanisms (like plumes) that are maintained in their mainland relatives, keeping the offspring on the island instead of out in the ocean.

## Plants can disperse across time via seed banks

Dispersal in time, as well as in space, can be important for some plant species. Once a seed is buried in the ground, it can remain there for years or decades. The **seed bank** is the collection of seeds in the soil (or less commonly in nondispersing fruits or cones); this term is sometimes used to refer to the seeds of a single species and sometimes to the seeds of an entire community. There is a tendency for short-lived plants (such as annuals) to have long-lived seeds, while long-lived plants (such as trees) and plants with large seeds tend to have short-lived seeds, although this is not a hard-and-fast rule. The short-lived plants can therefore have offspring waiting potentially for many years to germinate

Courtesy of S. Scheiner

**Figure 6.18** The native range of *Cocos nucifera* (coconut, Arecaceae) is tropical beaches across the Pacific and Indian Oceans. This plant was growing on a beach north of Cairns, in northeastern Australia.

until there are good conditions, while the long-lived plants can keep producing seeds each year so that finally when there is a good year for germination and growth, there will be viable seeds available. As a result, the abundances of species in the seed bank may have little relationship to those of the plants growing in the same spot. Many species growing above ground have few seeds in the soil. Other species may exist only in the seed bank at a given time. Species that specialize in colonizing disturbed habitats by wind dispersal, for example, may exist above ground for only a few years following a disturbance.

Seed banks also occur above ground in populations that have serotinous fruits or cones. **Serotinous** fruits or cone scales are bonded shut with resins and open only when temperatures are sufficiently high (see Figure 13.14). Normally the required temperatures are reached only during a fire. The seeds are thus dispersed only under conditions suitable for germination (bare mineral soil and increased nutrient availability) and are protected from granivorous animals at other times (Groom and Lamont 1997; Hanley and Lamont 2000; Lamont and Enright 2000). Serotiny is especially important in many conifers and in a number of species of Australian Proteaceae such as *Banksia* and *Hakea*.

We actually know very little about how long the seeds of many species can remain dormant, because the time scales necessary for such experiments are so long. In 1879, William J. Beal began a famous experiment at Michigan Agricultural College (now Michigan State University). He took a set of 20 glass jars and in each put a mixture of soil and 50 seeds of each of 19 common weed species and one crop species. He then buried the jars uncapped, but upside down. The idea was to allow moisture, small arthropods, microbes, and so forth access to the seeds, while not allowing new seeds to get in. Beal and his successors dug up one jar every 5 years for the first 40 years, and at 10-year intervals after that time, and tested the seeds to see whether they were still viable (Priestley 1986; Telewski and Zeevaart 2002). For 16 species, all seeds that germinated did so within 50 years, while the seeds of 4 species continued to germinate after that time. For 2 of these species—*Malva pusilla* (round-leaved mallow, Malvaceae) and *Verbascum* spp. (Scrophulariaceae; common mullein, *V. thapsus*, and moth mullein, *V. blattaria*, were not distinguished)—some germination occurred among 120-year-old seeds. There were also some interesting patterns of germination and dormancy. For example, seeds of *M. pusilla* germinated after 5, 20, 100, and 120 years of burial, but not at other time intervals. As the experiment was designed to look for maximum longevity rather than variation in germinability, we cannot estimate the latter. However, half the species showed gaps in germination like that seen in *M. pusilla*, albeit not so dramatic, suggesting that complex patterns of dormancy may exist in many species.

The existence of a seed bank can have important demographic consequences for a population. It can also help to buffer a population against year-to-year variation in demographic rates (bet hedging; see Chapter 8). Seed banks can also affect evolution in two ways. First, they slow the response to selection by maintaining genotypes produced in previous years. Second, mutation rates in seeds can be substantial, and seed banks can thus act as a source of genetic novelty (Levin 1990).

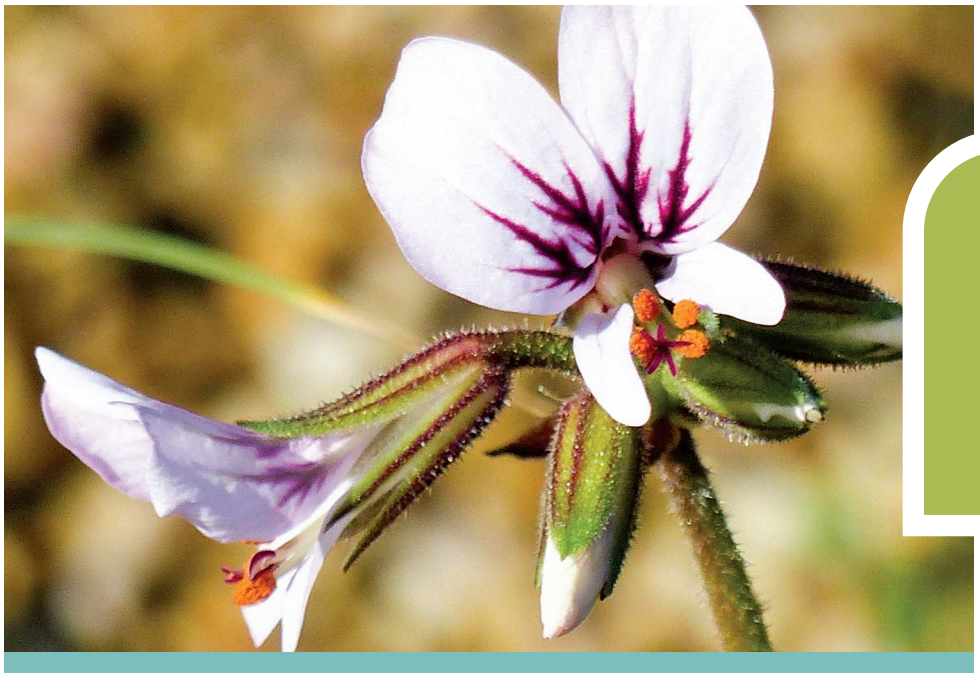
## Summary

- Growth and reproduction are key ways in which plants interact with the other organisms and the physical environment around them through their uptake of nutrients, use of space, pollination relationships, and dispersal of fruits and seeds.
- Plants grow by adding repeated modules such as leaves, flowers, and branches.
- The architecture of plants affects light interception.
- The growth of clonal plants affects their ability to take up patchy resources.
- Both sexual and asexual reproduction are common in plants.
- Some asexual reproduction is vegetative, but some is from seeds.
- The sexual life cycles of plants involve alternation of generations.
- The evolution of pollen and seeds laid the groundwork for much of the diversification of plants.
- Although there are remarkable floral adaptations and cases of coevolution with specific pollinators, most pollination in most species is probably by generalized animal taxa.
- The movement of pollen can occur by wind, animals, and sometimes water.
- Visual displays, floral odors, and acoustic guides can attract pollinators, and individuals may compete to attract pollinators.

(*Continued*)

## Summary (*continued*)

- Plants have complex mating systems that are often designed to either avoid inbreeding or increase outbreeding, and most mating occurs among neighboring individuals.
- Individuals within a species may vary in whether their flowers have stamens or stigmas.
- Competition occurs among plants and among pollen grains.
- Animals can be an important mode of seed dispersal, but animal-dispersed species must accomplish seed dispersal while avoiding seed consumption.
- Most seed dispersal is local, but rare long-distance dispersal events can be demographically important.
- The seeds of some species remain dormant in the soil for many years.



Courtesy of G. Fox

7

# Plant Life Histories

Why do plant species have such a wide variety of life cycle patterns? **Annual** species, such as most weeds and field crops, begin reproducing within a relatively short time after germination, produce many seeds in one or a few bouts of reproduction, and then die within the year in which they germinated. Other plants—the giant *Lobelia rhynchopetalum* (Campanulaceae; see Figure 18.3A) of the Bale Mountains in Ethiopia, *Argyroxiphium* (Hawaiian silverswords, Asteraceae; see Figure 8.11), and *Digitalis* (foxgloves, Scrophulariaceae; Figure 7.1)—wait for years before they reach a sufficient size for reproducing, then they pour their resources into reproduction, producing seeds in a single bout before they die. Organisms that reproduce in a single bout are called **semelparous** (the botanical term is **monocarpic** because the plant produces flowers once), while those that reproduce repeatedly are **iteroparous** (**polycarpic**). **Perennial** plants, including woody species such as most trees and shrubs, and herbaceous species such as spring-flowering garden plants, must reach a certain minimum size before reproducing (which can take months to years). Some of those species produce relatively few seeds each time they flower, but they survive for a long time and reproduce many times over their life-span. Many oak (*Quercus*, Fagaceae) species may produce some acorns every year but also a very large number every few years, a pattern called masting. The "King Clone" of *Larrea tridentata* (creosote bush, Zygophyllaceae; see Figure 3.6A) in the Mojave Desert is estimated to be over 11,000 years old. After decades reaching maturity, creosote bushes flower (if water is available) at least yearly (Vasek 1980).

A plant's schedule of birth, mortality, and growth is called its **life history**. As you can see, there is tremendous variation in life history patterns. Some of that variation is due to factors internal to a plant—how it grows and develops (see Chapter 6);

Above: *Pelargonium myrrhifolium* (Geraniaceae) in fynbos habitat near Stellenbosch, Western Cape Province, South Africa.

Courtesy of K. Whitley

**Figure 7.1** *Digitalis purpurea* (foxglove, Scrophulariaceae) is a biennial—a short-lived monocarpic perennial. These plants take 2 or more years to reach maturity, then devote an enormous fraction of their available resources to a single bout of reproduction. Relatively few plant species are strict biennials; in most so-called biennials, some plants reproduce in 2 years, and some in more.

other variation is due to external factors—patterns of seasonality, water availability, and so forth. In this chapter we will look at that variation and its causes. We discuss schedules of birth, mortality, and growth over a plant's life-span and the causes and consequences of variation in those schedules—for example, the implications of being perennial versus being annual. We also discuss phenology, the timing of growth and reproduction within a year. Variation in plant life histories has evolved over time, and this variation has important consequences for population dynamics (see Chapter 8).

## 7.1 Trade-Offs Are a Central Cause of Variation in Life History Patterns

The key to understanding the huge variation in life history patterns is the concept of trade-offs—that increasing one thing necessarily means decreasing something else. The idea of trade-offs caused by limited resources is central to most thinking about what determines life histories. To see why this is so, ask yourself what would be the most favorable plant life history if no trade-offs were necessary. Such a plant would produce unlimited quantities of seeds, each of which was large enough to maximize its chance of becoming established and growing quickly. New cohorts of seeds would be produced continually. The plant itself would grow rapidly and live indefinitely. Clearly, this combination of traits is not possible. Even if we substituted definite numbers for words such as *unlimited* and *indefinitely*, it would not be possible for a plant to maximize all of these life history components simultaneously; trade-offs are unavoidable.

Trade-offs have two causes. First, if you have a fixed amount of total resources, using those resources for one function (e.g., growth) means that you cannot use them for another function (e.g., reproduction). Second, it is not possible to maximize everything simultaneously. For example, trees that grow faster generally have xylem with thinner cell walls, with weaker wood and less sturdy trunks and branches. They do not live as long as slow-growing trees; you can either grow fast or live a long time, but not both. Of course, things are often more complex than these simple dichotomies. If you put all of your resources into growth now, you might end up with more resources for reproduction later. Besides trade-offs among traits, the same trait in different, related individuals may exhibit trade-offs. For example, seeds from the same mother might vary physiologically such that some germinate in drier years and others in wetter years. Trade-offs such as those between survival and reproduction, or between different times of reproduction, are the focus of the theory of life history evolution (Fox and Scheiner 2019).

Allocations of resources among traits often happen in a hierarchical fashion. For example, consider a plant that produces both male and female flowers (see Chapter 6). That plant may allocate some energy to growth, and some to producing offspring. The energy allocated to offspring might be further divided between the male and female flowers. Such a hierarchy does not imply that these actions necessarily happen in a set order; however, the consequences of one partition of resources (e.g., between growth and reproduction) create constraints on other partitions. So in this case, the amount of resources that remain that can then be divided between the male and female flowers will be set after energy has been allocated to growth.

### Trade-offs are difficult to measure

There are an enormous number of studies comparing species or higher taxa that show a negative correlation between traits such as longevity and fertility. Nevertheless, a substantial number of studies at the population level—where selection must be acting on these trade-offs—not only fail to show a negative correlation, but often show the opposite. This kind of result led a few scientists in the 1980s to argue that there really are

no trade-offs. The problem was clarified when Arie van Noordwijk and Gerdien de Jong (1986) pointed out that a positive correlation between, say, reproductive allocation and survival-related traits could actually be expected if the total amount of resources available varied among individuals. They suggested a useful analogy: For people operating under a fixed income, the amount of money available to spend on housing is negatively correlated with the amount available to spend on cars. Nevertheless, if you looked at the data, you would find that people who spend more on housing also spend more on cars—because they have greater total wealth. The trade-offs are real, but they operate in the context of variable access to resources.

A critical distinction must be made between traits that must be traded off because the physical rules of the universe demand it, and traits that are traded off because natural selection has molded species that way. An example of the first category is resource allocation to roots versus flowers: Limiting resources (carbohydrates, nitrogen, or water) used to produce and maintain roots cannot also be used to produce flowers. A possible example of the second category is an apparent trade-off between the speed of wood growth and wood toughness. Fast-growing trees such as *Populus tremuloides* (quaking aspen, Salicaceae) have weak wood, and individual trunks generally live only 50 or 60 years. Consequently, the quaking aspen tends to be an early successional species (see Chapter 13). We know that the toughness of the wood depends on the amount of lignin deposited into the cell walls. Is it possible for a mutation to occor that can lead to cheaper lignin production, or some other mechanism that can strengthen the cell walls? We cannot tell whether the growth rate versus wood density trade-off is unavoidable, or the right mutation has simply not come along.

Thus, we need to know more about variation among individuals in resource availability, as well as about the mechanisms determining resource allocation to different functions. Using a genetic model, David Houle (1991) showed that even when trade-offs do exist, genetic correlations between life history traits could evolve to be either negative or positive, depending on the genetic details. This finding points to the importance of increasing our understanding of the biological basis of life history traits.

## An important trade-off is in the size and number of seeds

The number of seeds per reproductive bout can vary tremendously. There are some plants that mature only a few dozen or so seeds per flowering episode, such as the palm *Lodoicea maldivica* (sea coconut, Arecaceae; Figure 7.2). A mature *Populus fremontii* (cottonwood tree, Salicaceae), on the other hand, may produce tens of thousands of seeds per episode; some orchids may produce hundreds of thousands to millions of seeds. If natural selection favors those genotypes that leave behind the most representatives over time, why don't all plants produce huge numbers of seeds?

Courtesy of S. Scheiner

**Figure 7.2** *Lodoicea maldivica* (coco de mer, Arecaceae) produces the world's largest seed, which typically weighs about 18 kg. The largest known weighs 25 kg. This species grows only on the islands of Praslin and Curieuse of the Seychelles, in the Indian Ocean.

Most plant ecologists studying this question have focused on the idea of a trade-off between the number of seeds a plant can make and the sizes of those seeds. Larger seeds generally contain more endosperm or other nutritional stores than seeds provided with fewer resources, and thus they generally have a greater chance of establishment success. Large amounts of endosperm in the seeds of some arid zone perennials, for example, allow them to sink deep roots quickly, increasing their chances of reaching water and therefore of survival.

The number of offspring a plant leaves through seeds, known as **female fitness**, involves an important trade-off. Producing more (but smaller) seeds probably reduces the chances of success for each seed because smaller seeds typically grow more slowly. On the other hand, there are likely to be diminishing returns in provisioning only one seed. The joint chance of success of, say, two medium-sized seeds can be greater than the chance of success of a single large seed. Natural selection on

maternal plants favors those that can divide resources among seeds so as to optimize the number of descendants left behind. Is there a single best seed size, or is it best to produce seeds in a range of sizes?

John Harper (1977) suggested that variation in seed size is quite limited in most plant species, as was suggested by the data available at the time. However, studies in the 1980s and 1990s revealed much variation in seed size within populations, even when natural selection was acting to constrain seed size because a particular size was optimal for a species (Mazer and Wolfe 1992; Platenkamp and Shaw 1993; Mojonnier 1998). In fact, individual plants often produce seeds of very different sizes (Venable 1985; Winn 1991; see Figure 9.1).

Using a graphical model, Christopher Smith and Stephen Fretwell (1974) showed that there is a single optimal size for seeds (Figure 7.3) if environmental conditions are predictable. Mark McGinley and his associates (1987) showed that variable seed size is favored when the environment varies over time and the geometric mean of fitness (see Chapter 8) is used as the criterion for fitness. The geometric mean gives the long-term average rate of population growth over a series of variable years and is therefore the appropriate measure of fitness averaged among years. Seed size can vary among individuals in a population or within a single individual. The reason for this variation might be that plants have been shaped by natural selection to produce variable seeds. It could also be that plants cannot avoid variable seed size because of the nature of the flowers they produce (McGinley et al. 1987). For example, many species in the sunflower family (Asteraceae) have flowering heads with two kinds of flowers, disk florets and ray florets, which may be constrained to produce different-sized seeds by their physical shape (Figure 7.4) (Venable 1985). Seeds within a legume pod may differ in size because those at the end closest to the maternal plant are able to garner more resources than seeds farther away (Silvertown 1987). It is important to realize that other factors besides natural selection determine the outcome of evolution (see Chapter 9). Even when selection favors a single seed size and when natural selection is the dominant evolutionary force, it may not be possible to eliminate variation in seed size within or between plants, because

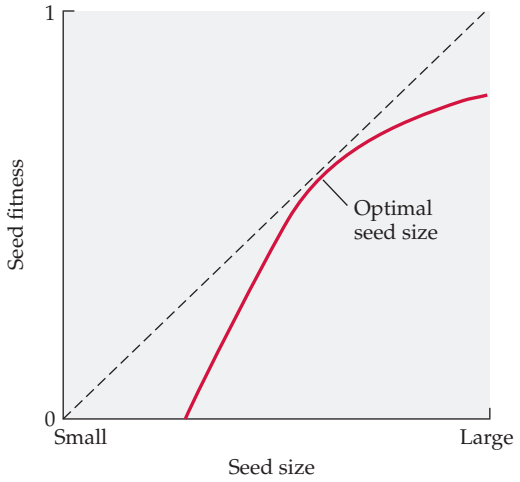



**Figure 7.3** Model for finding the seed size that maximizes the fitness of the maternal plant. The solid curve gives the fitness of seeds of different sizes under particular environmental conditions. Increasing seed size increases fitness, but above a certain size, there are diminishing returns in producing larger seeds. Below a minimum size, the seeds have no fitness. The optimal seed size is the value of the curve tangent to the 45° line. (After C. Smith and S. D. Fretwell. 1974. *Am Nat* 108: 499–506.)

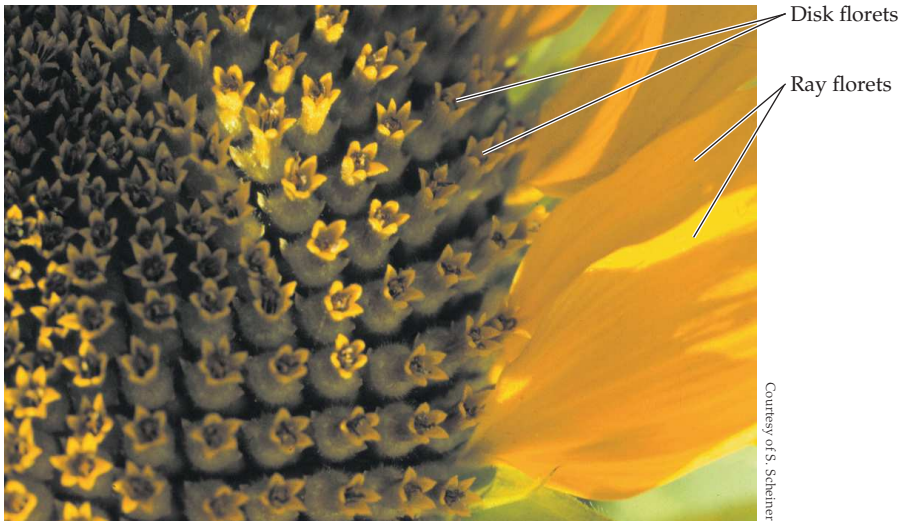



**Figure 7.4** *Helianthus annuus* (sunflower, Asteraceae) is a widespread North American annual plant that is widely cultivated for its seeds but also has many wild populations. Sunflower inflorescences (also called heads) are composed of two types of flowers (florets), each of which can produce a single seed. Disk florets compose the center of the head, while ray florets—which have a single very large petal and four very small petals—occur around the outside of the head. In some species in this family, seeds produced by different kinds of florets can differ in size, shape, and dispersal and dormancy properties.

this variation is also determined by environmental and developmental factors.

In determining seed size, there is an inherent conflict between the fitness of the maternal sporophyte and the fitness of the individual seeds. Natural selection on maternal plants favors those that leave the largest number of descendants, but it acts differently on seeds: the favored seeds are those that have the best individual size—often, the largest. An analogy may be helpful: If two parents have a large sum of money to leave their children, they may do best to ensure the success of the largest number of descendants by dividing the money evenly among the children. However, an individual child may best ensure his or her own success by getting as large an individual inheritance as possible. This sort of competition among offspring—and conflict between parents and offspring—also occurs in plants.

Individual seeds may vary in their efficiency at acquiring a share of the maternal sporophyte's (see Chapter 6) resources. Genetic differences, positional effects, differences in the timing of fertilization, or environmental variation can lead to differences in seed size. Selection on seeds favors increased ability to garner maternal resources whenever the fitness gained by doing so exceeds the inclusive fitness lost by depriving siblings of those resources. An individual's **inclusive fitness** is a measure of how well that individual passes on its genes as well as how well its relatives pass on the genes they share with that individual, weighted by how closely they are related. For example, if every seed in a fruit has the same pollen parent, then selection favors an individual increasing its seed size if doing so increases its fitness by more than the fitness lost by two siblings, because, on average, the seeds share half their genes with one another.

Classic studies of seed size were conducted by Edward Salisbury (1942) in England and Herbert Baker (1972) in California. Both focused on the ecological correlates of seed size. Baker, for example, concluded that plants with larger seeds were characteristic of more open habitats, such as deserts and coastal dunes. These studies have stimulated much research, partly because the methods they used made it difficult to distinguish the effects of selection from the effects of common descent. For example, seed size may have different ecological consequences in different environments, as Baker concluded. But it may also be that open habitats have many large-seeded plants simply because the plants that colonized those habitats are from taxa with large seeds. In an extensive study of the flora of the Indiana Dunes at the southern end of Lake Michigan, Susan Mazer (1989) was able to control for these two kinds of effects statistically. She found that habitat accounted for only 4% of the variance in seed size by itself. In the Indiana Dunes, the relationship between seed size and habitat largely depends on the history of the taxa that occur in each habitat.

Another factor that can affect selection on seed size and number is **granivory**, or seed predation. If a substantial fraction of seeds are eaten by granivores, there can be selective advantages to reducing seed size and increasing seed number. While tiny seeds may have reduced chances of establishment, they may also be too small to be worth the attention of granivores, especially if they are scattered widely, and they may just be difficult to find.

Seed size may also be dictated by dispersal strategy. If seeds are dispersed by wind, then there may be an optimal size to maximize the distance traveled. It is not just the resources invested in the seed that matter, but also the resources invested in the diaspore (the dispersing unit) as a whole (see Figure 6.15). Variation in diaspore size may ensure that seeds are dispersed different distances away from the parent plant, resulting in less competition among siblings. Eugenio Larios and Lawrence Venable (2015), studying the Sonoran Desert annual *Dithyrea califorinica* (spectacle pod, Brassicaceae), found that plants growing in more crowded, competitive locations produced smaller seeds that dispersed farther.

In this discussion of the ecology and evolution of seed size and number, we have assumed that there is a fixed total amount of resources allocated to reproduction at any time. When this is true, the problem reduces to the trade-off between number and size of seeds. This assumption is useful in gaining initial insight. But it is clear that both the timing of reproduction and the amount of resources allocated to reproduction are often genetically variable and have evolved over time. This realization leads to the more general problem of the evolution of life histories.

## 7.2 Evolution Acts on the Schedule of Survival and Reproduction

We now consider the more general problem: How does natural selection act on the schedule of survival and reproduction throughout a plant's life? Martin Cody (1966) suggested that throughout their lives, organisms must apportion resources among competing demographic functions: survival, growth, and reproduction. How an organism does this determines its fitness. Many studies have been conducted using the assumption that these three demographic functions really are competing for resources. It has become clear, however, that this is not always the case, because the functions of plant body parts do not correspond in any simple way to the three demographic categories. However, there is no doubt that life history traits are subject to natural selection.

Courtesy of K. Scheiner

**Figure 7.5** *Arabidopsis thaliana* (mouse-ear cress, Brassicaceae) is a small annual mustard. Native to central Europe, it has become naturalized in much of the temperate world. *A. thaliana* has one of the smallest genomes known in any angiosperm and is therefore widely used in genetic studies.

*Arabidopsis thaliana* (Brassicaceae, Figure 7.5) is a European plant that is now widely naturalized throughout the temperate world. It is perhaps best known as the "fruit fly" of plant molecular biology—that is, as an important model organism—because it has a small genome and reproduces quickly. Under some circumstances, some *Arabidopsis* genotypes can begin flowering 3 weeks after germinating. There is also remarkable variation in *Arabidopsis* life-spans. Individuals from some wild European populations live as long as a year. *Arabidopsis* from other populations can have two generations within a year—one of the few plants that do so. In some places with mild winters, such as the mid-latitude U.S. states of Kentucky and North Carolina, this plant is a **winter annual**—germinating in autumn, growing slowly through the winter, and flowering in spring. Where winters are colder, as in the northeastern United States, it is a **spring annual**—germinating in early spring, growing quickly, and flowering by late spring or early summer. There is a genetic basis for much of this variation within *Arabidopsis* (Kuittinen et al. 1997). Those genotypes with the shortest life histories have been the focus of most research because of their convenience, but evolutionary ecologists are now exploring the molecular basis of life history variation in this species.

Work by Gabriela Auge and her colleagues (Auge et al. 2017a, 2017b, 2019) shows one way that different life history traits can evolve together. In *Arabidopsis*, some of the genes that regulate flowering time also control seed germination, especially how both traits respond to temperature. These genetic connections ensure that plants that flower and set seeds in the fall produce seeds that will not germinate until conditions in the spring are suitable. However, it also means that if climate change alters the optimal time of flowering, the evolution of that trait may also change the timing of seed germination, which may or may not be optimal for the plants.

Several species with highly variable life histories have provided opportunities to study the biological basis of life history variation as well as natural selection on life history traits. Henk van Dijk and his associates studied life history variation in western European beets (Van Dijk et al. 1997; Van Dijk and Desplanque 1999; Hautekèete et al. 2001, 2002a, 2002b). *Beta vulgaris maritima* (sea beet, Amaranthaceae) is iteroparous, but its life-span varies from 2 to 10 years. There were marked geographic patterns to this variation. Long-lived populations were located in the Netherlands, Great Britain, and north Brittany, while short-lived populations occurred along the Mediterranean coast and in inland southern France. A large fraction of plants in inland populations flowered in their first year, but plants in the northern regions required **vernalization**, a period of cool or cold temperatures, before flowering. Several closely related taxa of *Beta* are usually or always semelparous. Much of this variation is genetically based, and there is much evidence that these traits are under strong natural selection.

Even if life histories are not highly variable, the actual performances of individuals within a population or a species can still vary. If there are no genetic differences among individuals, by chance alone some plants might be eaten by an herbivore while neighboring plants growing under the same conditions might not. So while it is very common to talk about the life history of a given species, keep in mind that it is a generalization that will not hold for all individuals of that species. This variation among individuals in life histories has important effects on the demography of a population (see Chapter 8) and provides a necessary ingredient for natural selection to operate (see Chapter 9).

## How long a plant lives and when it does its growing is part of its life history strategy

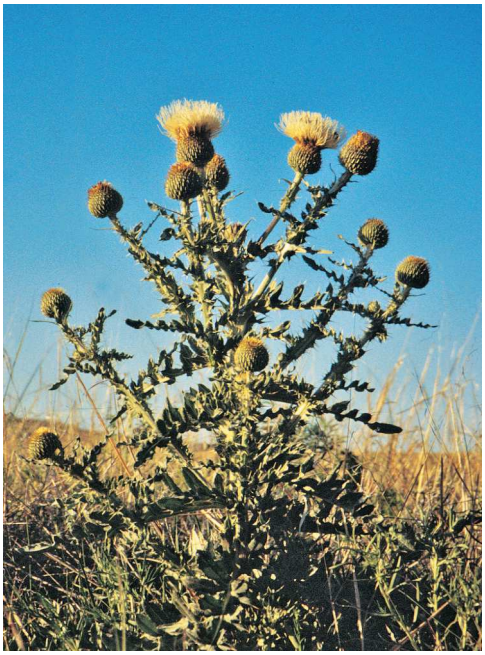
Many botanists use the term *life history* narrowly, to refer only to a plant's longevity and the number of times it can reproduce. Annuals, as we saw above, are plants in which the vegetative life cycle is completed in less than a year. Individual plants are "born" as seeds; since most annual plants have seed banks, this means that most annuals that you see growing are actually more than a year old, since they may have lived as seeds in the soil for many years before germinating. In temperate zone environments, many annuals can be categorized as either winter annuals (like *Arabidopsis* in the southern United States) or summer annuals. These terms can be confusing. Winter annuals germinate in autumn, overwinter as

vegetative plants, and flower and die in spring. **Summer annuals** complete their aboveground life cycle in the warm months. Winter and summer annuals in the same area can differ considerably in their developmental and physiological characteristics (Mulroy and Rundel 1977). Annuals are especially common in certain environments, including recently disturbed sites, deserts, and many dune environments. Most common herbaceous weeds are annuals.

Semelparous plants have a range of life-spans. **Biennials** (Figure 7.6; see also Figure 7.1) are semelparous plants that flower after 2 or more years. Actually, it takes most biennial species longer than 2 years to reach flowering size, and individuals in most biennial populations vary in how long they take to flower. **Semelparous perennials** are plants that live for a number of years before their single bout of flowering. There is no firm boundary between biennials and semelparous perennials (Young and Augspurger 1991), although biennials often die back to ground level in winter (or the unfavorable season), persisting only underground, while semelparous perennials have substantial aboveground structures year-round.

Both theoretical and empirical studies point to two factors that might select for semelparous life histories: a critical size for reproduction and diminishing returns from retaining resources for subsequent bouts of reproduction. Biennials and short-lived semelparous perennials often occur in early successional habitats, where canopy closure after several years causes a decrease in the quality of the habitat. Long-lived semelparous plants are characteristic of relatively unproductive but persistent habitats, such as deserts, many grasslands, and alpine regions. Factors such as competition for pollinators may be more important in selecting for semelparity in these settings, because larger plants with larger flowers and flower stalks may attract more pollinators.

It is important to be clear about whether it is the genet or the ramet that is semelparous. Monocarpic ramets may be favored by selection when production of inflorescences is very costly and it takes plants a long time to acquire enough resources to flower. This can be the case when plants are pollen limited (see Chapter 5) and large inflorescences receive disproportionately many pollinator visits (Schaffer and Schaffer 1979). It is important not to confuse monocarpic ramets with true semelparity: a genet with monocarpic rosettes, such as an *Agave* (Figure 7.7) or a bromeliad, can reproduce many times in its



Courtesy of S. Louda

**Figure 7.6** *Cirsium canescens* (Platte thistle, Asteraceae) is a "biennial."

© Bob Gibbons/Alamy Stock Photo

**Figure 7.7** *Agave* (century plant, Agavaceae) has semelparous rosettes that flower only once, mobilizing all the energy stored during a life-span of a few decades to produce a large inflorescence with hundreds of flowers. The genet itself is iteroparous, however, because vegetative reproduction precedes flowering.

life, while a semelparous plant (such as most Hawaiian silverswords) has a much riskier life history in that it can reproduce only once.

In most habitats, it is the perennial plants that give the landscape its characteristic appearance. The category "perennial" includes a great diversity of plants. Some live for only a few years, while others live for centuries or even longer. While most woody plants live for at least a few years, the reverse is not necessarily true. Some herbaceous plants, such as the cushion plant *Silene acaulis* (Caryophyllaceae), live for centuries (Morris and Doak 1998). Among forest trees, there are species that live for decades and others that live for centuries, even in the same type of habitat. In pine flatwoods in the southeastern United States, for example, *Pinus palustris* (longleaf pine, Pinaceae) can live for 300 or more years, while *Cersis canadensis* (eastern redbud, Fabaceae) lives for only a few decades. Typically, long-lived perennials have short-lived seeds and a small or nonexistent seed bank. Most perennials, however, have the ability to become dormant as vegetative plants during winter or the dry season. Perennials have more ability to withstand variable and harsh environmental conditions in their aboveground phase than do annuals; this ability increases with the longevity of the species. Long-lived temperate trees, for example, are more likely to withstand a hard freeze or a drought than a short-lived herbaceous perennial would be.

How do we explain all of these different life histories, as well as the variation within each? What are their consequences for population growth? Early efforts at such explanations focused on the theory of *r*- and *K*-selection, while recent work has centered on demographic approaches.

## 7.3 Several Theories Address Life History Strategies

There have been several attempts to pull together all of these ideas and observations about life history traits and strategies into more comprehensive theories. The development of those theories has approached the problem from different perspectives, both evolutionary and ecological.

### Demographic life history theory is based on evolutionary principles

The first approach to developing a life history theory based on evolutionary principles focused on demography, examining how population growth is affected by variation in the survival and fertility of individuals as a function of their size, age, or other life history traits. Its utility can be seen in its solution of "Cole's paradox." In studying the consequences of schedules of reproduction, Lamont Cole (1954) developed a model that compared the fitness of a perennial phenotype with that of an annual phenotype of the same species. If a fraction, $p$, of the plants survive to reproduce each year and produce $F$ seedlings if they do reproduce, then the perennial's fitness is $\lambda_p = p(F + 1)$, while the annual's fitness is $\lambda_a = pF$. For example, imagine two species, one that is semelparous and has six offspring and then dies, and a second species that is iteroparous and each year produces five offspring per individual. These two species would have populations that grow at the same rate, with $6p$ individuals reproducing each year (see Chapter 8 for a discussion of lambda [$\lambda$] and population growth rates). Cole concluded that annuals need only increase their reproductive output by one new seedling per year to do as well as perennials. He pointed out that this finding was paradoxical because most organisms are perennial, not annual, suggesting that there must be other advantages to being perennial.

Using a demographic approach, Eric Charnov and William Schaffer (1973) pointed out a flaw in Cole's reasoning: His result depends on the assumption that rates of survival and fertility do not depend on age. If his model is modified slightly to give first-year plants a different chance of surviving than older perennials, the paradox disappears. Charnov and Schaffer demonstrated this as follows: If the fraction of first-year plants surviving is $c$, and the surviving annuals produce $F_a$ seeds while the surviving perennials produce $F_p$ seeds, then $\lambda_p = cF_p + p$ and $\lambda_a = cF_a$. Thus, the population growth rates of both phenotypes are given by the chance of individuals surviving to reproduce times their fertility if they do survive; for the perennials, this rate is increased by the chance of surviving from year to year. Therefore, the annuals have greater fitness only when $\lambda_p < \lambda_a$—that is, when $F_a > F_p + p/c$—or when their fertility exceeds the perennials' fertility plus the survival chances of adults relative to those of juveniles. This means that where survival of adults is much greater than survival of juveniles, the annuals' reproductive output needs to exceed that of the perennials by a considerable amount for the annuals to have a greater lifetime fitness—thus resolving the paradox.

### *r*- and *K*-selection theory was influential in earlier thinking about life histories

Robert MacArthur and his students and colleagues developed their ideas about life history evolution, focused on the concept of *r*- and *K*-selection, during the 1960s. This work stimulated much of the early research on the evolution of life histories, although the strict concept of *r*- and *K*-selection plays little role in current research on life histories. Nevertheless, it continues to be influential in popular writing on ecology, in textbooks, and in some aspects of community ecology, so we explain it briefly here.

Under the logistic model of population growth, the per capita growth rate is

$$\frac{1}{N}\frac{dN}{dt} = rN\left(\frac{K-N}{K}\right)$$

where $r$ is the intrinsic rate of population growth, $K$ is the carrying capacity, and $N$ is the population size. Robert MacArthur and Edward Wilson (1967, MacArthur 1972) suggested that at low population densities, natural selection will be strongest on traits that increase the intrinsic rate of population growth, $r$. At high densities, they suggested, selection will be strongest on traits that increase population size, or the carrying capacity, $K$. These conclusions about the strength of selection come from asking a simple question: When the population is small (or large), what changes would increase the per capita population growth rate the most? MacArthur and Wilson argued that there is stronger selection for productivity than for efficiency at low densities, and vice versa at high densities. Selection for increasing $r$ and selection for increasing $K$ were widely interpreted as counterposed to one another. More recent researchers have pointed out that there is, however, no reason why there must be a trade-off between $r$ and $K$. Increasing both parameters always increases the per capita growth rate, so selection under logistic growth always favors increasing both $r$ and $K$ (Boyce 1984; Emlen 1984).

### *r*- and *K*-strategy theory was extended to the ecology of plants

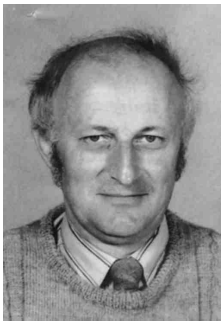
Using the assumption that high population densities lead to reduced survival of juveniles, Eric Pianka (1970) developed detailed predictions for animal populations, which were later extended to plants. Because this version of the theory is about the characteristics of organisms likely to be found in particular types of habitats, we refer to it as *r*- and *K*-strategy theory. A key assumption in Pianka's work was that a given population tends either to be colonizing newly disturbed habitat (so $r$ is under stronger selection) or to be near carrying capacity (so $K$ is under stronger selection). Broadly, Pianka predicted from this assumption that *K*-selection should favor delayed reproduction and a reduced number of offspring, while *r*-selection should favor early reproduction and an increased number of offspring. These predictions are what most people mean when they refer to *r*- and *K*-selection: weedy plants, for example, are taken to be *r*-selected, while forest trees are taken to be *K*-selected. As Boyce (1984) put it, references to "*r*-strategists" generally imply small organisms that reproduce quickly and are poor competitors and thus should be favored in frequently disturbed habitats, while references to "*K*-strategists" generally imply that organisms have the opposite characteristics, although these characterizations were not derived from the original model.

Perhaps the greatest problem with *r*- and *K*-selection theory, or its extension as an *r*- and *K*-strategy theory, is that its predictions are qualitative and relative, rather than quantitative, and so are difficult to test. Plants that are *r*-selected, for example, are predicted to evolve an earlier age at first reproduction and greater seed numbers than those that are *K*-selected. But what can one measure to test this? Plants within one community might be classified as *r*-strategists or *K*-strategists based on which reproduced at an earlier age, but this classification might be different for the same species in another community, or in a comparison of one of the species with another species. Furthermore, such a classification says nothing about whether the differences among the species are actually a result of selection on the age of first reproduction, and it certainly does not tell us whether they are part of an entire life history strategy.

Our contemporary understanding of the importance of stochastic changes in population dynamics (see Chapter 8) and disturbance (see Chapter 13) helps us to understand a deep conceptual flaw in *r*- and *K*-selection theory. Its emphasis on the distinction between plants that live mainly in disturbed habitats (*r*-strategists) and those that live in stable habitats (*K*-strategists) now seems to have been misplaced. Communities in many "stable" habitats, such as mature forests, have been shown to depend on disturbance for the maintenance of the community composition. Previous notions of climax communities and stable equilibria have changed dramatically in recent years, reducing the utility of this equilibrium-based theory. Another major problem was that *r*- and *K*-strategies seemed to many plant ecologists to be inadequate for describing plant life histories.

### Grime's triangular model focuses on the ecological conditions favoring different life history strategies

Philip Grime (1977, 1979) proposed an extension of *r*- and *K*-strategy theory that focused on plants. His theory differed from most others by focusing on the environmental conditions that might favor one life history strategy over another. His central argument was a simple one: he contended that plant populations tend to face consistent selective pressures mainly from physiological stress and disturbance. Stress was defined as all the factors external to a plant that limit growth; most of these factors are abiotic, but he included shading by other plants as



Courtesy of J. Philip Grime

J. Philip Grime

a stress. Plants could adapt to three combinations of these pressures, Grime suggested: environments imposing low stress and little disturbance would select for competitive ability (*C* selection), environments imposing high stress and little disturbance would select for stress tolerance (*S* selection), and environments imposing low stress and much disturbance would select for a **ruderal** (weedy) habit (*R* selection). Grime argued that there was no viable strategy for plants in high-stress, high-disturbance environments. Although he explicitly stated that his theory was a generalization of *r*- and *K*-selection theory, it was not derived from any particular model of population growth or evolution. Instead, it is an ecological theory based on how species with different traits will be sorted into different environments.

The *C*-*S*-*R* model is often described by the diagram shown in Figure 7.8. Trade-offs are implicit in this diagram. Imagine a population initially at the point on the *C* axis for 25% importance of competition; this point also has the values of 75% on the *S* axis and 75% on the *R* axis. Selection for increased competitive ability would move the population along the *C* axis, but that would necessarily reduce its stress and disturbance tolerance. That said, under the *C*-*S*-*R* model, plant populations can have any combination of traits associated with *C*, *S*, and *R* selection. Grime contended that one could map broad groups of plants—such as bryophytes or trees and shrubs—into regions of the triangle (Figure 7.9).

One well-known set of predictions of *C*-*S*-*R* theory is shown in Table 7.1. These predictions, like those of *r*- and *K*-selection theory, are qualitative and relative and therefore do not lend themselves to critical testing. Indeed, if the plants in a community perfectly matched these descriptions, that would not constitute evidence that selection had acted as Grime supposed in developing his theory. Like *r*- and *K*-selection theory, *C*-*S*-*R* theory has a particular emphasis on the relative importance of competition in different habitats (see Chapter 10). Consequently, while this theory received a fair amount of attention and generated considerable controversy, it did not have much influence in the study of life history evolution and has lost influence in plant community ecology more recently.

Both the *r*- and *K*-selection model and the triangular model are examples of typological thinking, which reduces a complex set of traits and organisms to just a few categories. Such models can be useful as a way of organizing one's thinking as long as one realizes that the world is more complex and that any of the models' categories apply to only a subset of species and traits. There are many species that mix and match pieces of the syndromes predicted by these models, and others do not fit into any scheme. One alternative is not to look at whole organisms and syndromes, but to focus, instead, on specific traits, such as seed size and number. When we do, we find evidence for the trade-offs that formed the bases for these models.

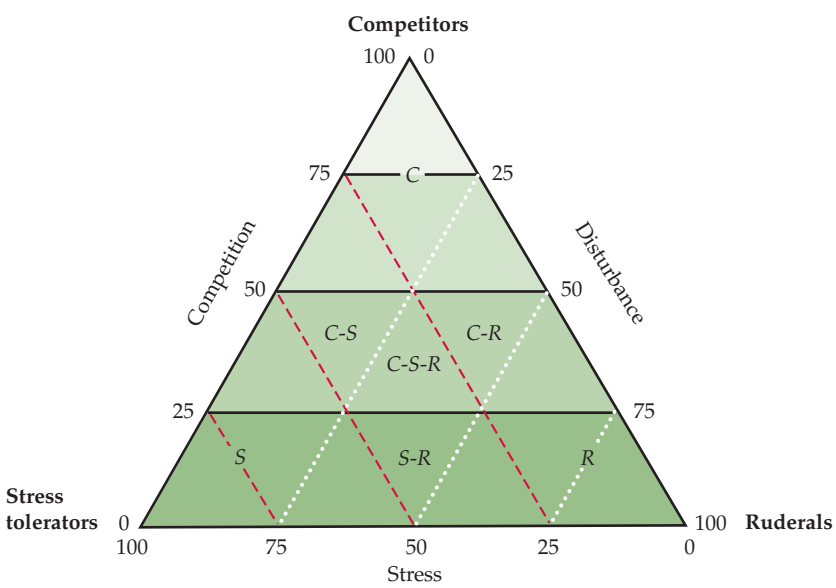



**Figure 7.8** Grime's (1977) triangular or *C*-*S*-*R* model. Each side of the triangle measures the importance of a major factor (competition, stress, or disturbance), relative to its maximum possible value. Each line inside the triangle gives equal values of the importance of one of these selective factors. Plant populations at extreme values of one of these factors (the vertices of the triangle) are identified as competitors, stress tolerators, or ruderals. The initials in the interior regions are descriptive: for example, plants near the center have strategies reflecting selection on all three factors, while those to the left (*C*-*S*) have been subject to *C* and *S* selection. (After J. P. Grime. 1977. *Am Nat* 11: 1169–1194.)

## Demographic life history theory has been tied to patterns of reproductive allocation

If it matters when a plant reproduces, and how much it reproduces at each life stage or age, how does natural selection act on the allocation of resources at each stage? If we assume that there is a trade-off between reproduction and survival, what does that tell us about how evolution shapes life histories? This question has provided the focus for much research in life history evolution.

Key to addressing this problem is the concept of reproductive value ($v$; see Chapter 8). William Schaffer and Michael Rosenzweig (1977) showed that fitness is

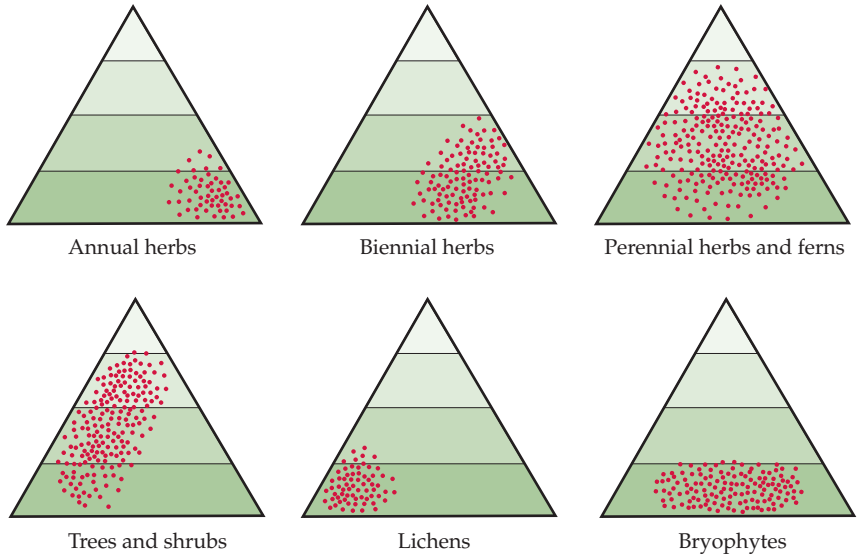



**Figure 7.9** Grime's (1977) characterizations of large groups of plants according to his *C-S-R* model. He argued that lichens had mainly evolved as stress tolerators, and annuals as ruderals, while trees and shrubs reflected every possible combination of selective forces other than the most extreme (the vertices). (After J. P. Grime. 1977. *Am Nat* 11: 1169–1194.)

maximized at every age if the sum of present reproduction plus future reproduction weighted by relative reproductive value,

$$F_i = \frac{p_i v_i + 1}{v_o}$$

is maximized at every age. The fraction of resources allocated to reproduction at age or stage *i* is the **reproductive effort** at that age or stage. Schaffer and Rosenzweig analyzed selection on life histories graphically by plotting current and future reproduction against reproductive effort.

Figure 7.10 shows two simple alternatives. In Figure 7.10A, the curves for both present reproduction ($F_i$) and future reproduction $[(p_i v_i + 1)/v_0]$ are convex, so their sum is maximized when reproductive effort is either 0 or 1 at any given age or stage. These curves can differ for different ages or stages. Since a reproductive effort of 1 means putting all resources into reproduction and none into survival, this graph describes selection for a semelparous life history. In Figure 7.10B, both curves are concave, so selection favors intermediate reproductive effort at every age after the onset of reproduction—an iteroparous life history. Combinations of curves of different shapes can thus describe a wide variety of life histories.

Testing this kind of theory has proved difficult. One problem, pointed out by Schaffer and Rosenzweig (1977), is that in many cases, more than one life history is optimal. But there is a deeper problem in actually measuring reproductive effort: How does one assign the fraction of resources used for reproduction? Little is known about the underlying developmental and biochemical processes of reproduction. An early study (Hickman and Pitelka 1975) found that the dry weight of plant parts, except for some unusual tissues, is highly correlated with their energy

**TABLE 7.1** Some characteristics of competitive, stress-tolerant, and ruderal plants, as described in Grime's triangular model

| | Competitive | Stress-tolerant | Ruderal |
|---|---|---|---|
| Growth forms | Perennial herbs, shrubs, or trees | Lichens, perennial herbs, shrubs, and trees | Annuals |
| Seed production | Small | Small | Large |
| Maximum potential growth rate | Rapid | Slow | Rapid |
| Leaf litter | Abundant, often persistent | Little, often persistent | Little, not persistent |
| Leaf longevity | Short | Long | Short |
| Flowering phenology | Flowering near time of maximum productivity | No pattern | Flowering at end of favorable period |
| Vegetative phenology | Leaf production coinciding with maximum productivity | Evergreens; various patterns | Brief period of leaf production at time of maximum productivity |
| Life span | Long | Long | Short |

*Source*: J. P. Grime. 1977. *Am Nat* 111: 1169–1194.

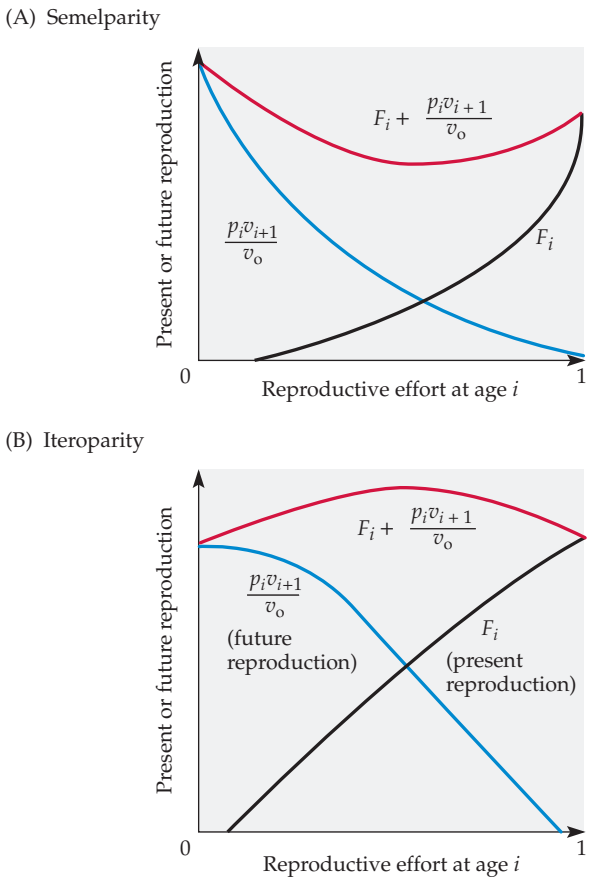



**Figure 7.10** Selection for semelparity versus iteroparity. Fitness is greatest when the sum (red curves) of present reproduction plus future reproduction is maximized at every age. (A) Semelparity is favored when the curves for both present and future reproduction are convex—fitness is greatest when reproductive effort is always either 0 or 1. (B) Iteroparity is favored when both curves are concave—fitness is greatest at intermediate reproductive effort once maturity is reached. (After W. M. Schaffer and M. L. Rosenzweig. 1977. *Ecology* 58: 60–72.)

content. This result led to hundreds of studies comparing the dry weights of plant parts under varied circumstances.

But how does one assign a plant part to "current" versus "future" reproduction? For flowers and fruits, the choice is clear, but not for organs such as roots. One attempt at handling this problem was a study by Edward Reekie and Fakhri Bazzaz (1987). These investigators measured the amount of root mass present during flowering of a grass and the amount present at other times, and they found no difference. They concluded that none of the root mass counted toward reproduction.

Another problem with testing this theory is that the parts of a plant do not relate in any simple way to the demography of the organism. Because reproductive effort is a ratio, it cannot describe unique life histories (Schaffer 1983) because there are many different life histories that could give the same allocation to reproduction at a particular time. For example, one population might grow vegetatively for 10 months and then allocate 90% of its photosynthate to flowers and fruits for 2 months, while another might grow for 8 months and allocate 50% of its photosynthate to flowers and fruits for 4 months—yet when measured at 12 months, both could have 75% of their biomass in reproductive structures. This observation led to models using more sophisticated mathematics, such as optimal control theory, which can handle the technical problem of optimizing reproductive allocation throughout the life-span. However, little additional work has been done along these lines, because meaningful predictions require that the models do something that we do not know how to do yet: to mimic the specific biological processes underlying plant growth, development, and resource allocation. These problems are still under active investigation by molecular and cellular biologists, and progress in this approach to life history theory will depend on progress in better understanding these processes.

Matrix models (see Chapter 8) play an important role in both theoretical and empirical studies of life histories. In the simple case of the Charnov-Schaffer model of perennial life histories, the original model stated that the numbers of juveniles and adults ($J$ and $A$, respectively) next year will be $J_{t+1} = cF_p J_t + pF_p A$ and $A_{t+1} = cJ_t + pA_t$. Rewritten in matrix form, this is

$$\begin{bmatrix} J_{t+1} \\ A_{t+1} \end{bmatrix} = \begin{bmatrix} cF_p & pF_p \\ c & p \end{bmatrix} \times \begin{bmatrix} J_t \\ A_t \end{bmatrix}$$

Analysis of this model tells us that the population will eventually grow at the rate $\lambda_p = cF_p + p$, the dominant eigenvalue of the matrix. At that time, the population will be composed of $F_p$ times as many juveniles as adults. In this case, the consequences of varying the survival and fertility rates become clear. For example, increasing juvenile survival, $c$, increases $\lambda_p$ at the rate $F_p$. With more realistic models, it becomes more difficult to analyze this sort of model without a computer. However, the methods described in Chapter 8 make it possible to completely analyze the effects of varying any survival or fertility rate on population growth or composition.

## Other theories of life history strategies are based on examining plant traits

A variant of *r*- and *K*-strategy theory is the notion of a fast-slow continuum such that slower-growing species tend to reproduce less and at later ages (Stearns 1983; Reich 2014). This theory has been developed using a different approach from those previously described. All of the previous theories start with evolutionary and ecological principles from which mathematical or graphical models are developed. An alternative is to start by

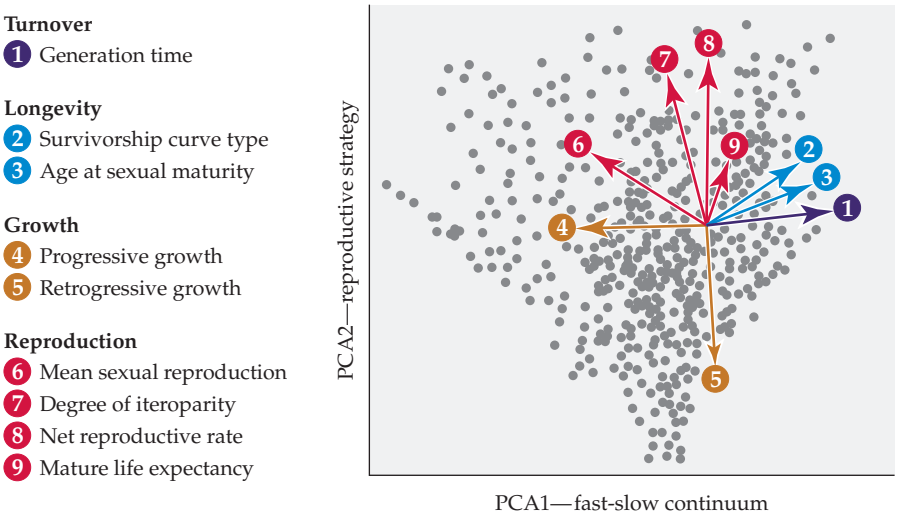



**Figure 7.11** A principle components analysis of the life history traits of 625 species. Each point represents one species. The arrows indicate the direction for increasing values of each trait. The first, horizontal axis relates to the fast-slow continuum. Individuals that grow quickly, have short generation times, have high rates of juvenile mortality, mature quickly, and have high rates of reproduction are located to the left on the graph; those that grow slowly, have long generation times, have low rates of juvenile mortality, reproduce later in life, and have low rates of reproduction are located to the right. The second, vertical axis relates to reproductive strategy. Species toward the top of the graph produce many seeds, often, and for a long time. Species toward the bottom produce fewer seeds and not as often; they also tend to show regressive growth where they can get smaller at times. (After R. Salguero-Gómez. 2017. *New Phytol.* 213: 1618–1624.)

looking at how life history and other traits are distributed among species and to distill general patterns from the analysis of traits. Such bottom-up approaches have become more common in recent decades as the volume of biological information has increased, as the information has become easier to access, and as computational tools have become easier to use.

Rob Salguero-Gómez and his colleagues (Salguero-Gómez et al. 2016; Salguero-Gómez 2017) approached this question by assembling data on 625 plant species from across the globe representing all growth forms and major terrestrial biomes. The data consisted of information on nine life history traits, including generation time, longevity, growth rate, and reproductive pattern, estimated from the COMPADRE database (see Chapter 8). The data were analyses using principle components analysis (see Chapter 12), which summarizes the variance structure among the traits. They discovered that the pattern was more complicated than a simple fast-slow continuum (Figure 7.11). The first axis did align with the predictions of the fast-slow continuum. But it explained only 35% of the variation among species. The second axis explained another 25% of the variation and was related to reproductive strategy. The graph shows that high reproduction can be associated with both fast and slow growth. There is no necessary trade-off between growth and reproduction at this level of analysis. An interesting feature of the data is the relative lack of species in either the lower left or lower right portions—species with low scores on the slow-fast continuum and either very low or very high scores on the reproductive output axis. Species with those combinations may not exist, either because such combinations of traits are not developmentally feasible, or because they are unable to compete successfully. These last thoughts are speculative but show how such analyses can generate new hypotheses.

Courtesy of Rob Salguero-Gómez

Rob Salguero-Gómez

## 7.4 Year-to-Year Variation in the Environment Shapes Life Histories

### Among-year demographic variation reduces fitness

Demographic variation among years reduces the long-term fitness of plants. How can this demographic variation among years be reduced when the variation occurs randomly among years? Many plants appear to achieve this by "hedging their bets" (spreading the risk among years; Starrfelt and Kokko 2012). To see why variation among years changes fitness, consider a hypothetical annual species that always produces an average of 1.5 seeds every year (that is, half of the plants

in the population produce 1 seed, and half produce 2). Compare it with another annual species that produces an average of 0.5, 1, and 3 seeds in dry, average, and wet years, respectively. If these types of years occur with equal frequency (1/3), the average number of descendants of the species that varies its output among years will be $0.5^{1/3} \times 1^{1/3} \times 3^{1/3} = 1.14$, which is 0.36 (24%) fewer than the number produced by the species that does not vary its output. In general, increasing the variance in fitness among years decreases its long-term geometric mean. This implies that there is a trade-off between mean and variance in fitness: gaining added fitness in "good" years can come at the expense of reducing long-term average fitness. For matrix models, the problem is somewhat more complex because the order of multiplication matters. Analyses of both structured and unstructured population models (see Chapter 8) suggest that natural selection often acts to increase fitness by reducing among-year variation in fitness.

### Bet-hedging strategies can reduce the variance in fitness

Bet hedging occurs when an individual trades off a larger payoff for a greater chance of some payoff. For example, if you have money to invest in the stock market, you could invest all of that money on a new tech start-up, hoping that it becomes the next Amazon or Facebook. Of course, like most start-ups it is mostly likely to fail, and you would lose your entire investment. Alternatively, you could invest your money in a fund that spreads that money across many different companies. That investment will not make you wealthy, but if you leave the investment alone for many years, you are quite likely to see it grow at an average rate most people would find reasonable. Biologically, bet hedging means trading off reduced mean fitness against a reduced variance in fitness. Bet-hedging life histories involve issues of allocating resources among related individuals—either between parents and their offspring or among siblings.

Bet hedging may result in a higher return when the environment is variable in time (say, wet versus dry years) or space (good versus bad patches) and that pattern of variation is unpredictable. Consider the evolution of seed dormancy in an annual plant. We can measure the fitness of that plant by the number of offspring that survive to reproduce. If the amount of rainfall is the same every year and suitable for survival and reproduction, the fitness of the plant is maximized when all seeds germinate the following year. But what happens if the amount of rainfall varies from year to year such that some years have sufficient rainfall for growth and other years are very dry and most or all individuals die before they are able to reproduce? In that case, the long-term fitness of the plant may be greater if it produces some seeds that delay germination until later years, reducing the possibility of many (or all) descendants having a low (or zero) probability of surviving to reproductive age in some years. Although a delay in germination (and therefore reproduction) reduces the mean fitness in any given year, having some seeds remain dormant reduces the variance in fitness among years. Similarly, bet hedging can occur through dispersal: if some patches are favorable and others unfavorable in a given year, offspring dispersing to a mixture of patch types can increase long-term fitness by reducing the risk of all offspring being in unfavorable patches (Venable and Lawlor 1980).

Examples of bet hedging include spreading reproduction more evenly among years, increasing the area over which seeds are dispersed, and increasing dispersal time by means of seed banks. Seed banks are perhaps the best-studied of these mechanisms (see Chapter 8). Most annual plants have seed banks. Without environmental variation, having seeds remain dormant for one or more years as opposed to germinating immediately is a bad strategy, and it would be acted against by natural selection, since seeds that remain dormant in the soil cannot reproduce while those that germinate can do so. A number of studies (e.g., Brown and Venable 1991; Kalisz and McPeek 1993) suggest that seed banks play an important role in buffering populations from catastrophic losses due to environmental variation.

Lawrence Venable, Jennifer Gremer, and colleagues (Venable 2007; Gremer and Venable 2014; Gremer et al. 2016) showed how such bet hedging happens in annual plants growing in the Sonoran Desert. Starting in 1982, they monitored permanently marked plots, collecting information about every month for over 30 years. In each plot, individuals of the 15 most abundant annual species were mapped upon germination and followed until death to measure lifetime survival and reproduction (see Figures 8.13 and 8.14). Starting in 1989, once a year they also took samples of the soil and estimated the density of viable, nongerminating seeds, which allowed them to estimate the fraction of seeds germinating each year. Some species were more sensitive than others to the year-to-year variation in rainfall, and more-sensitive species were more likely to die without flowering in dry years. They found that, as predicted by bet-hedging theory, more-sensitive species had a smaller percentage of seeds germinate in any single year and their seeds lived longer in the soil (Figure 7.12).

© Davis Silpher, UC Davis
Jennifer Gremer

Incorporating year-to-year variation into demographic models of life history evolution can change the conclusions about the role of seed banks in buffering populations. Mark Rees and his collaborators (1999) studied the

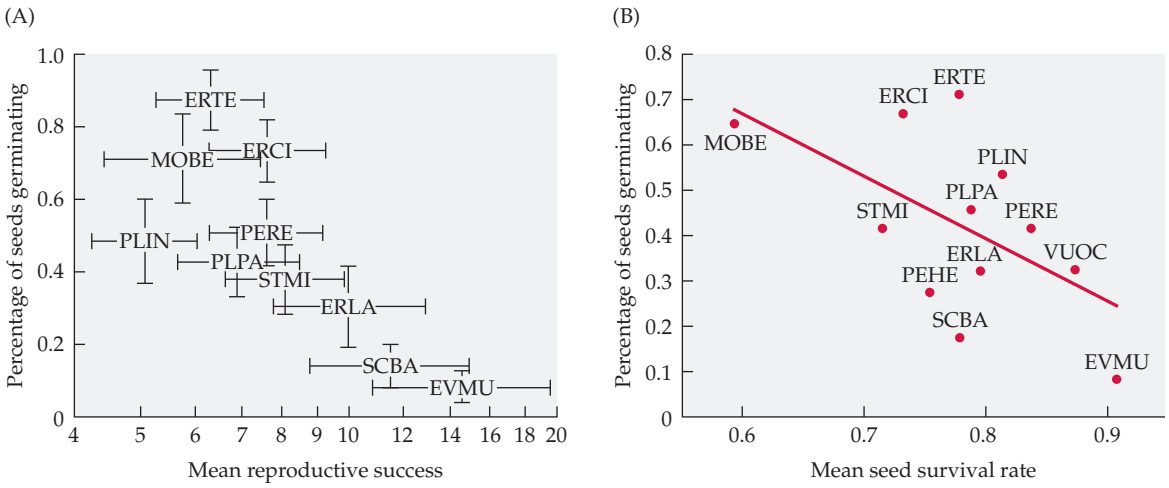


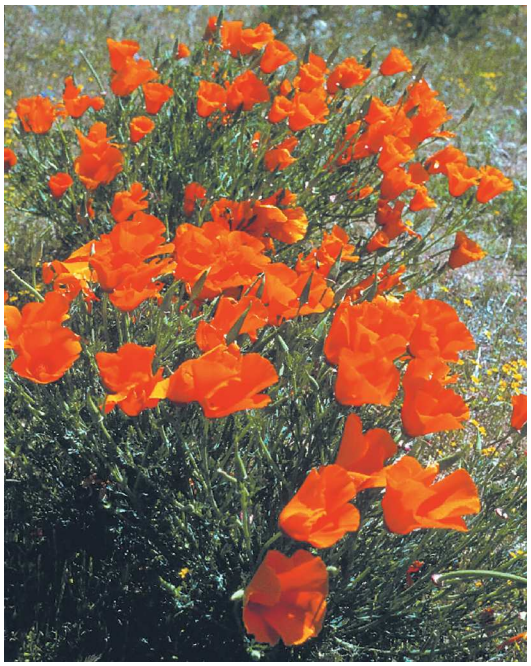

**Figure 7.12** The percentage of seeds that germinate each year for ten species of Sonoran Desert annuals as a function of (A) the mean reproductive success of each species (the average number of seeds produced per germinating seed = the probability of survival from germination to reproduction × the average number of seeds of reproducing individuals) and (B) the mean rate of seeds surviving in the soil from one year to the next. (A after D. L. Venable. 2007. *Ecology* 88: 1086–1090; B after J. R. Gremer and D. L. Venable. 2014. *Ecol. Lett.* 17: 380–387.)

effect of year-to-year variation in vital rates on selection on size-dependent flowering in the semelparous plant *Onopordum illyricum* (Illyrian cottonthistle, Asteraceae) in southern France. They found that temporal variation increased the risk of delayed reproduction and thus selected for earlier flowering than one would predict with a deterministic model.

A key insight is that almost all plant life histories have mechanisms that buffer populations against environmental stochasticity (see Chapter 8). Because populations recruit new individuals in some years but not others, different life histories are different ways of storing successful recruits as a buffer against those bad years. For example, almost all annual plants maintain an interannual seed bank as a store of recruits, while the long-lived individuals of a population of trees serve as a similar sort of storage mechanism. Variation in the conditions under which populations can successfully recruit is thought to play an important role in the coexistence of competing species (see Chapter 10).

## Seed germination is triggered by many factors

A number of different environmental factors help to trigger germination. In many plants, temperature plays a key role in regulating germination. For example, winter annuals in the Sonoran Desert do not germinate during the summer rainy season, and the reverse is true of summer annuals. In at least some plants, some cell membranes in the seeds undergo temperature-dependent phase transitions that render the cells either more or less permeable to water, thus allowing or preventing **imbibition** (uptake of water by seeds), depending on the season (Bewley and Black 1985). Light also plays a key role in many species. For example, many weedy species require light to germinate; that is why many of these species germinate following soil disturbance. There are also species, such as *Eschscholzia californica* (California poppy, Papaveraceae; Figure 7.13), in which germination is inhibited by light. But the induction and breaking of seed dormancy can be highly complex processes,



Courtesy of K. Whitley

**Figure 7.13** *Eschscholzia californica* (California poppy, Papaveraceae) has two unusual, but unrelated, features of its life history: germination is inhibited by light, and it is an annual in some areas and a perennial in others.

regulated genetically and physiologically. Seeds may have to age ("mature") before they are capable of germinating. They may have to experience vernalization (stratification)—days or weeks of exposure to cold temperatures before they will germinate when experiencing warm conditions. They may require cycles of alternating cold and warmth. Seeds also may be able to detect changing daylengths as cues to the right season in which to germinate; it is the seed coat (maternal tissue) that detects daylength changes.

When should a seed germinate? This might seem to be obvious: it should germinate when conditions are favorable. But can plants use environmental conditions at germination time to predict subsequent conditions—that is, can plants evolve predictive germination? The answer depends critically on the predictability of the environment. In a maritime climate (see Figure 16.19C), for example, summers are often warm and wet. Germination in late spring usually provides a fairly good chance of survival to maturity. In contrast, rainfall in desert environments can be highly unpredictable (see Figure 16.22), and many plants that germinate do not survive to maturity (Fox 1989; Venable and Pake 1999).

In a study of the annual *Plantago insularis* (desert Indianwheat, Plantaginaceae) growing in the Sonoran Desert, Maria Clauss and Lawrence Venable (2000) found only a slight correlation between rain at germination time and subsequent rain during the growing season. This correlation was greatest at mesic sites (that is, the most favorable microhabitats). However, populations at xeric sites (the most unfavorable microhabitats) would be expected to experience the strongest selection for being able to predict when adequate rainfall would be available. Clauss and Venable pointed out that predictive germination might still be occurring in these plants, but if so, it is a complicated phenomenon. For example, heavy rain at germination time does predict subsequent rain in xeric (but not mesic) sites in El Niño years (see Chapter 16).

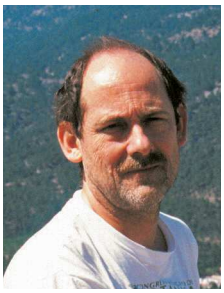
Predictive germination may involve factors other than climate. Some annuals in fire-prone shrublands in California are stimulated to germinate by compounds in smoke (Keeley and Fotheringham 1997). Fire reduces aboveground biomass and the density of competitors, and minerals in the ash add nutrients to the soil, so conditions following fire are favorable for growth of annuals. Many species classified as weedy (ruderal) use various cues signaling recent disturbance as stimuli for germination.

## Masting can result in synchronization of flowering among individuals

Many plants—especially woody plants—have large and erratic variation among years in the size of the seed crop produced, and this variation is generally synchronized among most plants in a population. This phenomenon, called **masting**, has been widely explained as an adaptation to seed predators (granivores; see Chapter 11). Some ecologists have hypothesized that during mast years, the large numbers of seeds overwhelm the capacity of granivores to eat them all, allowing at least some seeds to survive. This explanation has been questioned recently by other ecologists, who argue that other factors may be more important in selecting for masting. They argue that masting can be more easily explained as an adaptation to wind pollination, which is increasingly efficient at high pollen densities (Smith et al. 1990; Kelly et al. 2001).

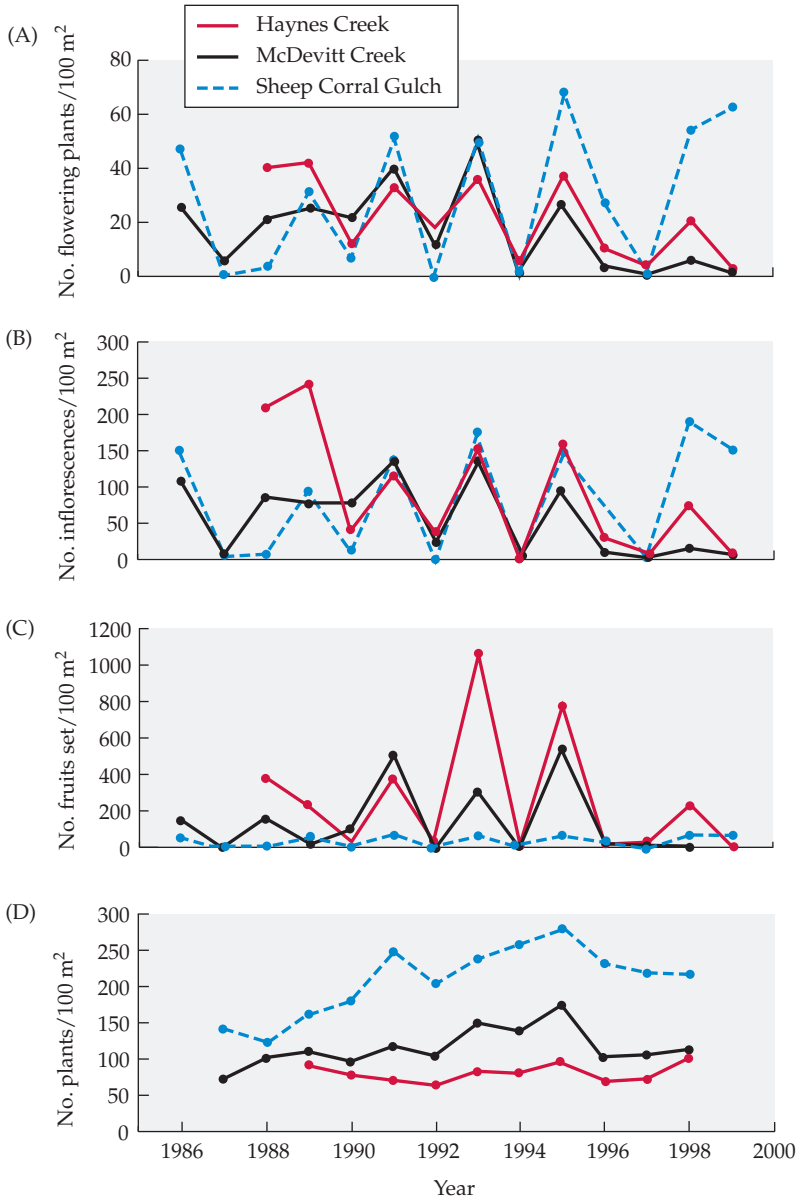
While the two hypotheses need not be mutually exclusive, there is mounting evidence that many wind-pollinated species are indeed pollen limited and that masting may be an evolutionary response to this pollen limitation (Kelly and Sork 2002; Koenig and Ashley 2003). In a survey of 59 populations of 24 different species, Walter Koenig and coworkers (2003) were able to estimate the separate effects of among-year variation in the seed production of individuals and of the tendency of those individuals to flower in synchrony. They found that these two quantities varied independently and that population-level variation among years depended more strongly on individual variation than on among-individual synchrony. This finding suggests that selection can act separately on individual variability and synchrony.

In a large survey of studies on variation in seed production, Carlos Herrera and associates (1998) argued that the distinction between masting species and nonmasting species is false because there is a continuum of variation. Most iteroparous woody plants produce more seeds in some years than in others. Even when phylogenetic relationships are controlled for, there is slightly more variation in wind-pollinated species than in animal-pollinated species. However, the researchers also found more among-year variation in species dispersed by granivores or by abiotic mechanisms (such as wind or water) than in those dispersed by frugivores. The study by Koenig and colleagues (2003) also found that among-year variation was a continuum, rather than falling into just two classes.


Courtesy of Carlos Herrera

Carlos M. Herrera

What are the physiological causes of among-year variation in reproduction, and how does a population come to flower synchronously? Elizabeth Crone and Peter Lesica (2004) studied synchrony in the iteroparous perennial herb *Astragalus scaphoides* (Bitterroot milkvetch, Fabaceae) at three sites in Idaho and Montana. They found



Courtesy of E. Crone

**Figure 7.14** Reproductive synchrony in *Astragalus scaphoides* in three populations in Idaho and Montana. The densities of flowering plants (A) and inflorescences (B) are strongly synchronous both within and among populations; fruit set (C) is less so. The marked fluctuations in reproduction do not reflect variation in the densities of plants (D) in these populations. (Graphs after E. E. Crone and P. Lesica. 2004. *Ecology* 85: 1944–1954.)

that there were advantages to synchronous flowering in these populations, but the advantages varied. All populations experienced at least one of three effects as a result of synchrony—decreased floral herbivory, increased fruit set, and decreased granivory by weevils—but the importance of these effects differed among populations. Despite the strong reproductive synchrony of these plants (Figure 7.14), there was no correlation between weather and flower or fruit production. However, the researchers pointed to models by Akiko Satake and Yoh Iwasa (2000, 2002) that suggest an explanation: If individuals need more resources to flower than can usually be acquired in a year, plants could easily be forced to reproduce in alternate years or at longer (and sometimes irregular) intervals. Synchrony could then be achieved by any mechanism that causes low reproductive output in a given year—either poor weather or pollen limitation—because if few plants flower in a given year, they expend fewer resources, and more plants will reproduce in the following year. Under this scenario, the synchrony itself need not be advantageous to the plants. Mark Rees and his collaborators (2002) reached similar conclusions in a study of the alpine grass *Chionochloa pallens* (snow tussock, Poaceae) in New Zealand. Models of flowering that depended only on internal resource levels or only on weather cues failed to predict this species' synchronous flowering, but a model incorporating both internal resources and permissive weather events successfully predicted the flowering pattern of *C. pallens*.

## 7.5 Phenology Is the Within-Year Schedule of Growth and Reproduction

**Phenology**—the timing of growth and reproductive activity within a year—can vary greatly among species, populations, and even individuals. In temperate climates, there are usually some plants flowering at all times between the last and first frost. Some woody species are evergreen, while others are **deciduous**. Even within groups that we usually think of as all of one type—such as conifers, which we think of as

(A)

Courtesy of S. Scheiner

(B)

Courtesy of S. Scheiner

**Figure 7.15** *Larix* (the larch, Pinaceae) is a deciduous conifer found in taiga around the world. (A) *L. kaempferi* (ochudo) at 3500 m elevation on Mt. Fuji, Japan. (B) A branch of *L. laricina* (tamarack) in a bog at 200 m elevation in northern Michigan.

evergreen—there may be exceptions, for example, the deciduous conifer species *Larix laricina* (the larch or tamarack, Pinaceae; Figure 7.15). In general, plant phenologies are constrained by seasonality—mainly by temperature or moisture availability. Phenologies have a more complex relationship with animal pollinators, seed dispersers, and herbivores because the plants and animals can be agents of selection on each other.

## The timing of growth is driven by both abiotic and biotic factors

In temperate deciduous forests, many herbaceous plants on the forest floor expand their leaves and flower before the canopy trees begin leaf expansion. The consequence is that these plants do most of their growing and reproducing in temperatures that are far colder than those experienced by the canopy trees during the same life stages. Many studies have shown that this timing is actually advantageous to the forest floor herbs. Once the canopy closes, very little sunlight—sometimes as little as 1%—gets through to the forest floor. Thus, if the forest floor herbs delayed their growth and reproduction until later in the season, they would be even more limited by light availability than they are by cool temperatures. Six herbaceous spring species were studied by Carol Augspurger and Carl Salk (2017). They experimentally manipulated the timing of the start of growth in the spring—from early to late—and examined the effects on survival and flowering. They showed that the brief growth period of all of the species was an adaptation to make use of the short window between the end of winter cold and the onset of canopy closure in the late spring.

In most temperate species, it appears that temperature and photoperiod (day length) are the main factors determining vegetative phenologies. Temperature plays a critical role in a particular way. It is generally the number of **degree-days**—the sum of temperatures experienced over some period of time—and not the temperature on a particular day that determines the timing of leaf expansion. In cold years, then, leaf expansion (sometimes called leaf flushing) is delayed. Many plants, like many other organisms, use photoperiod (see Chapter 2) as a reliable predictor of the average temperature. If they relied exclusively on temperature, a warm spell in midwinter would cause many plants to expand their leaves. By using temperature as a cue, plants can respond to an early spring by flushing their leaves early, but because they also use photoperiod as a cue, this response is limited. These plants, in other words, respond to environmental fluctuations, but they do so cautiously.

As the example of temperate forest floor herbs suggests, there are at least two main types of selective factors that act on vegetative phenologies. Abiotic factors limiting growth—such as the timing of killing frosts or seasonal droughts—often determine the beginning or end of growth episodes (or both). Biotic factors—in particular, competition for light or water—are also important in limiting growth.

This observation raises an interesting question: Why do the canopy trees in temperate forests wait so long before expanding their leaves? Why don't they use the early

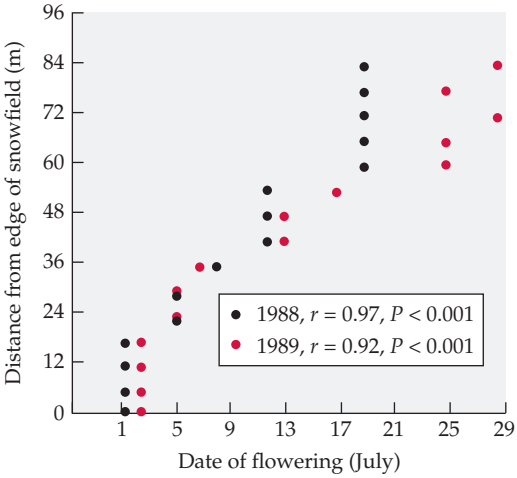
spring to add to their growth, as the forest floor herbs do? There are several selective forces affecting the timing of leaf expansion. First, the canopy is elevated, and temperatures in the treetops can be considerably colder than those at ground level. Thus, leaf expansion for trees occurs later on the calendar but may not really be much later in terms of degree-days. Second, late frosts are a common occurrence in many temperate environments. Leaves on trees are much more vulnerable to frost damage than those on forest floor herbs because the latter are partly sheltered by the trees (longwave radiant heat loss to the trees is far less than from the trees to the night sky; see Chapter 3) and because the ground-level temperatures are higher. Third, all metabolic enzymes have defined temperature ranges over which they can operate and are most efficient at particular temperatures. Enzymes adapted for peak functioning at warm temperatures are unlikely to be efficient in the early spring, so it is possible that earlier leaf expansion might reduce the total annual growth of the trees, not increase it. Finally, many temperate zone trees are wind-pollinated, and pollination in most of these species occurs in early spring while the trees are still leafless. The presence of leaves earlier in the season would be likely to limit pollen transfer.

The phenologies of temperate forest floor herbs are actually more complex than we have implied so far. Most of these plants in northeastern North America have phenologies like those we have described, but there are others that are able to capture and use light at other times. For example, some species can use sunflecks on the forest floor to grow during the warmer months (see Figure 2.27). Still others use the additional light available in autumn, when some canopy species have begun shedding their leaves (Mahall and Bormann 1978).

In warmer habitats, it is usually the availability of moisture that has the strongest effect on vegetative phenologies. In tropical deciduous forests and in deserts, as well as some tropical rainforests, rainfall is seasonal. Where it is strongly so, many species are drought-deciduous, shedding their leaves during dry periods (see Chapter 3). Is the timing of leaf expansion and leaf fall in these species a plastic response to the environment or an evolved strategy based on the average environmental conditions? To a great degree, the answer is, Both. There are some species that retain their leaves if given additional water, and others shed them at the usual time of the onset of the dry season, regardless of current water availability. There are species that expand their leaves only after abundant rainfall and others that do so weeks before the rains begin. Casey Ryan and colleagues (2017) used remote sensing data to track green-up across the woodlands and savannas of southern Africa, finding that green-up happens, on average, about a month before the start of the rains, most likely as a response to day-length changes. All of these strategies are evolved traits. Perhaps the most extreme case of a drought-deciduous phenology is that of *Fouquieria splendens* (ocotillo, Fouquieriaceae; see Figure 3.9A). This desert plant of the southwestern United States and northwestern Mexico produces new sets of leaves whenever water is available, as long as temperatures are high enough. Thus, ocotillos can have four or five sets of leaves over a year.

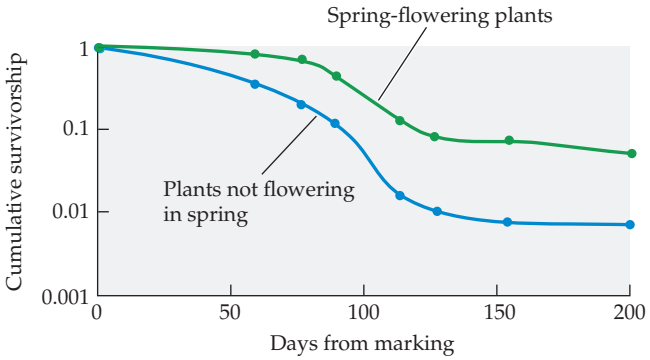
## The timing of reproduction may be due to abiotic factors

While we often think of flowers as something we see in spring, plants can flower under an amazing range of conditions. *Ranunculus adoneus* (snow buttercup, Ranunculaceae) is a plant of alpine snowfields in the Rocky Mountains of Colorado (Figure 7.16). This perennial herb begins flowering while still covered with snow (Galen and Stanton 1991). The North American desert annual *Eriogonum abertianum* (Abert's buckwheat, Polygonaceae; Figure 7.17) has a remarkably variable reproductive phenology: individuals may flower in either or both of two rainy seasons separated by as much as 6 months of drought (Fox 1989). In both of these plants, flowering time is affected by physical factors: temperature and light availability in *Ranunculus* and temperature and moisture availability in *Eriogonum*. Both plants are unable to reproduce in winter: *Ranunculus* is covered by deep snow



Courtesy of M. Stanton

**Figure 7.16** *Ranunculus adoneus* (snow buttercup, Ranunculaceae) is found in alpine snowfields in the Rocky Mountains; it begins flowering while still covered by snow. Flowering time depends on the plant's position in a snowfield: plants on the edge of the snowfield tend to flower earlier. (After C. Galen and M. L. Stanton. 1991. *Am J Bot* 78: 978–988.)



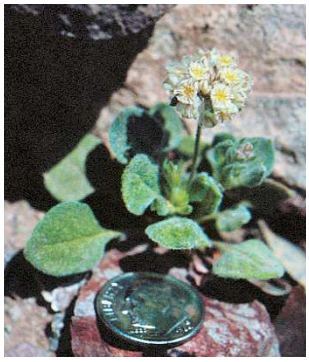


Courtesy of K. Whitley

**Figure 7.17** *Eriogonum abertianum* (Polygonaceae) is an annual plant of the Sonoran and Chihuahuan Deserts of the southwestern United States and Mexico. Its flowering time depends on plant size and moisture availability. Therefore, spring-flowering plants are more likely to survive to the summer rainy season, when most seed is set. These data are from a population at Organ Pipe Cactus National Monument, Arizona, in 1985. (Data from G. A. Fox. 1989. *Ecology* 70: 1294–1306.)

then, while *Eriogonum* germinates only during the cool desert winter, so it is usually much too small to flower in winter. In both cases, there is apparently strong selection to flower at what appear to be unusual times.

*Eriogonum* may flower during either the spring or the summer rainy season, or both. Like most plants, *Eriogonum* produces more flowers and fruits the larger it gets—and it gets very large only after the summer rains. The fact that many or most individuals, depending on site and year, die before summer does not necessarily provide sufficient selection for earlier flowering. For example, at one site, well under 1% of the plants alive in spring managed to survive and flower in summer, but they accounted for about 25% of the seeds produced (Fox 1989).

In both the *Eriogonum* and *Ranunculus* species, there is substantial variation in flowering time within a population. Much of this variation appears to be due to phenotypic plasticity (see Chapter 9) rather than genetic variation within populations (Fox 1990a; Stanton and Galen 1997). This is especially clear in *Ranunculus*. Plants on the edge of the snowfield flower much earlier, mainly because the snow melts earlier there, so they receive more light earlier in the year (see Figure 7.16). With *Eriogonum*, small differences in soil moisture lead to big differences in reproductive timing. Plants that reach sufficient size and have access to sufficient moisture begin flowering in spring; these plants have a greater chance of surviving to the summer rainy season than do plants that do not flower in spring (see Figure 7.17). Despite this plasticity, it is clear that differences among populations also reflect local adaptation. Populations that experience a greater risk of dying before summer have more spring-flowering plants than those at a lesser risk (Table 7.2), and the populations differ genetically for the time to flowering, size at flowering, total seed set, and time to completion of flowering (Fox 1990a).

It is widely believed that the initiation of moisture limitation at the start of the dry season—and perhaps other types of physiological stress—tends to induce flowering in many plants. That is, plants "decide" to start flowering and producing seed when they detect that the end is near. However, there is no actual evidence for this effect in annual plants (Fox 1990b). In perennials, the problem is more complicated because flower buds are often formed one or more years before they open. The effect of drought or other stress therefore probably depends on its timing. Drought that occurs when buds are forming may have a different effect than drought that occurs just as the buds are opening.

How tightly is reproductive phenology coupled with vegetative phenology? Evidence for temperate forest trees suggests that their coupling is generally weak (Lechowicz 1995). The main reason is that plants can store resources such as photosynthate when conditions are favorable and remobilize them at other times. This buffering mechanism is what allows so many temperate trees to flower when they are leafless.

## The timing of reproduction may be due to biotic factors

Biotic factors may also influence reproductive phenology. Because of the inherent difficulties in studying multispecies interactions, many studies of these influences have relied heavily on circumstantial evidence—but the circumstantial evidence in some cases is quite strong.

If pollinators are only briefly available, there may be strong selection for flowering while these animals are present and active. Although this seems clear in principle, it is important to realize that such selection will be strong only if the plants require animal pollination and if seed set is limited by the availability of pollen. Some plants that are mostly self-fertilizing still require animals for pollination.

A number of authors in the 1970s inferred that flowering schedules of entire communities were structured by competition for pollinators. They concluded that this competition was a crucial factor in shaping phenologies,

**TABLE 7.2** Frequencies and sizes of spring-flowering classes of *Eriogonum abertianum* in several populations

| Site and year | Plants alive in spring that flowered in spring (%) | Mean (standard error) height to highest node | | *n* |
|---|---|---|---|---|
| | | Spring flowerers | Nonspring flowerers | |
| Chihuahuan Desert | | | | |
| Portal, 1985 | 2.6 | 4.4 (0.58) | 2.2 (0.05) | 494 |
| Portal, 1986 | 2.6 | 4.8 (0.43) | 1.7 (0.03) | 1404 |
| Sierra Vista, 1986 | 0.0 | — | 1.3 (0.04) | 200 |
| Sonoran Desert | | | | |
| Tucson Mts., 1986 | 56.7 | 2.8 (0.10) | 1.7 (0.06) | 201 |
| Waterman Mts., 1986 | 31.0 | 4.2 (0.43) | 2.0 (0.05) | 200 |
| Organ Pipe, 1985 | 21.3 | 2.5 (1.9) | 1.5 (0.06) | 150 |
| Organ Pipe, 1986 | 23.1 | 3.1 (0.7) | 1.7 (0.02) | 1382 |

*Source*: G. A. Fox. 1989. *Adaptation, history, and development in the evolution of a desert annual life history.* PhD. Dissertation, University of Arizona, Tucson, AZ.

*Note*: More plants flower in spring in the Sonoran Desert (which receives less rain, and less predictable rain, in the summer) than in the Chihuahuan Desert. Spring-flowering plants are larger on average and have improved chances of surviving to summer (see Figure 7.17), even when compared with same-sized plants (Fox 1989). Moisture availability, as well as size, strongly influences phenological variation within populations; the differences between the Portal and Organ Pipe populations, at least, have a strong genetic basis (Fox 1990a).

causing the evolution of separation in phenologies among overlapping species. Other authors suggested that many plant species act as mutualists by having similar phenologies and thus jointly attracting more pollinators than any would by flowering alone.

Clearly both phenomena are possible, but there is remarkably little evidence for either supposition. Analyses by Robert Poole and Beverly Rathcke (1979) showed that many of the claims that observed phenologies showed strong correlations among plant species were based on inappropriate data analyses. For example, many of the effects that had been claimed disappeared when Poole and Rathcke took into account the fact that the seasons limit when plants can flower. If one analyzes temperate zone data without accounting for winter, or tropical data without accounting for dry seasons, it is easy to wrongly conclude that many species' flowering times are clustered more than one would expect by chance.

Frugivores (fruit eaters) and granivores (seed eaters) may also exert considerable selection on phenologies. In most cases, however, it appears that the animals' phenologies are more strongly dependent on plant phenologies than the reverse. Animals may also be important causes of variation in reproductive phenology. For example, herbivory (like many other stresses, including crowding and drought) may cause an individual's flowering time to be delayed (Lyons and Mully 1992).

Finally, other plants may influence reproductive phenology, either directly or indirectly. Crowding by neighbors, for example, may act to delay flowering (Lyons and Mully 1992). Such delays, in turn, may result in selection for faster growth or earlier flowering.

## Summary

- A plant's life history is its schedule of birth, mortality, and growth.
- Life histories can vary considerably.
- The key to understanding this variation is trade-offs, which occur because of basic constraints in resource availability.
- An important trade-off in plants is between seed size and the number of seeds.
- There have been several attempts to develop broad theories to describe how trade-offs shape syndromes of life history traits.
- These theories have focused on demography, patterns of reproductive allocation, possible trade-offs between growth and competition (r- and K-selection theory),

(*Continued*)

## Summary (*continued*)

and the environmental conditions that might favor one life history strategy over another (the *C-S-R* model).

- Examination of life history traits among species shows that various growth rate strategies and various reproductive strategies are largely independent of each other.
- Environmental uncertainty creates variation in fitness among years, which reduces long-term fitness.
- Many plants appear to reduce among-year fitness variation by bet hedging—spreading risk by evolving traits such as iteroparity and seed banks.
- Masting—large-amplitude fluctuations from year to year in reproductive output—may evolve as a consequence of granivory, pollen limitation, or large demands of reproduction on resources.
- Phenology, the within-year timing of growth and reproductive events, is also affected by various trade-offs.
- Many plants living on forest floors must either experience most of their growth when the canopy trees are bare or evolve the ability to use small amounts of light efficiently when the canopy is in leaf.
- The timing of flowering involves trade-offs between the availability of abiotic resources, such as water and light, and biotic factors such as pollinators and frugivores.

Courtesy of G. Fox

8

# Population Structure, Growth, and Decline

Most populations in nature are always changing. But is a population changing just a little bit, or a lot? Slowly or quickly? Is a population of plants increasing, decreasing, or staying about the same size over time? Is it spreading, or occupying a smaller area? These questions are central to both basic and applied ecology. Evolutionary biologists want to understand how natural selection on plant traits shapes and is shaped by population change, while foresters aim to maximize the number of forest trees surviving attacks by bark beetles (see Chapter 11), land managers attempt to control the numbers and spread of invasive weeds, conservation ecologists might seek to prevent the decline of a rare plant species, and other ecologists track the expansion of plant species to higher elevations in response to climate change. All of them are asking questions that rely on understanding **population dynamics**—the study of changes in numbers, composition, and the spatial dispersion of a population.

Changes in population numbers are referred to as population growth, which (unlike in ordinary language) can mean either increase or decrease in abundances—much as acceleration means increase or decrease in velocity. A key factor affecting population growth is whether those dynamics are **density-dependent** or **density-independent** (referring to individuals of the same species). Some factors do not depend on the **density** (number of individuals per unit area) of the population. For example, if 90% of seedlings die every year because of seasonal droughts, regardless of the number of plants around them, ecologists would say that these droughts are a density-independent factor. If mortality is increased beyond this level in high-density sites (because the seedlings are competing for water), ecologists would describe the additional deaths as density-dependent mortality. The distinction

Above: *Trichocereus pascana* and *Larrea divaricata*, in Parque Nacional Los Cardones, Salta Province, Argentina.

between density-dependent and density-independent factors is often useful. In any case, density independence is a good starting point for thinking about population growth, because it is the simplest scenario.

Most demographic studies of plant populations assume density independence. Since some **vital rates**—the reproductive rate and chance of surviving—undoubtedly depend on density at least under some conditions, does this assumption of density independence make sense? Several arguments suggest that it does. First, the estimates found by assuming density independence are always correct as initial descriptions. Imagine that 83% of juvenile plants die during a year. That 83% figure is a legitimate estimate of mortality regardless of whether some process causing mortality is actually density-dependent. Second, growth rates derived from density-independent data can still be very useful as indicators of the population's current status, even if they are not always useful as predictions of future population size. Hal Caswell, a leading demographic theorist, provides a useful analogy: the speedometer of a car imparts useful information about its current speed, but few would confuse that number with a prediction that the car will have traveled a particular distance in the next hour.

The concept of density must be used cautiously when considering plants. John Harper (1977) pointed out that the *average* density over a whole plant population is not the important quantity to measure when assessing how a plant experiences density, for example, in understanding the effects of competition among plants. Because plants are rooted in place, they usually compete only with their neighbors, not with more distant plants in the same population. When considering competition, think of density dependence on a very localized scale. But it is even more complicated: competition for different resources may occur at different spatial extents. Competition for phosphate, which is fairly immobile in soils (see Chapter 4), occurs over centimeters, while competition between trees for light can occur over meters or more. Individual plants may compete for pollinators with neighbors at some distance (negative density dependence), but dense patches of flowering plants can attract pollinators from far away (positive density dependence). Other spatial scales may be important for other phenomena; for example, because many pathogens can disperse over tens or hundreds of meters, or more, the average density over that larger area may be a useful quantity for understanding the progress of disease in a plant population. A more general way to categorize these different kinds of phenomena is to distinguish between changes in population size that depend on interactions among individuals (such as competition) and changes in which such interactions do not play a role.

In this chapter, we focus on the basics: density-independent population dynamics in which density does not play a role or its role is not considered. There are no general models for density-dependent population growth (Buckley et al. 2001; Caswell 2001). The models that do exist often involve fairly complex mathematics; density-independent models are sometimes good approximations of those models (Grant and Benton 2000; Buckley et al. 2003). We discuss density-dependent changes in populations more generally in Chapter 10 and Chapter 11. Because population growth is all about numbers and changing numbers, it necessarily involves mathematics, and in this chapter we will introduce some of the mathematics used to understand how plant populations grow and decline. Without the tools of mathematics, we could not understand or predict these essential things about plant populations.

## 8.1 Plant Biology Creates Special Issues for Population Studies

What controls the growth of plant populations? Change in population size is determined by the numbers of births and deaths, plus the number of individuals immigrating into the population, minus the number emigrating from the population. Ecologists usually get data on changes in population size by conducting regular censuses—say, by counting individuals every year or every week. If we think of the first census as occurring at time $t$, and the next census as occurring at time $t + 1$, we can describe the change in population size as

$$n(t + 1) = n(t) + B(t) - D(t) + I(t) - E(t) \qquad (8.1)$$

This is the basic equation of demography. Here the $n$'s are the size of the population at each sequential census. For example, $n(t)$ might be the number of individuals of *Pinus ponderosa* (ponderosa pine, Pinaceae) in the first year of a census, and $n(t + 1)$ the number the following year. $B(t)$ is the number of births between the censuses, $D(t)$ the number of deaths, $I(t)$ the number of individuals arriving (immigrating) from other populations (in this case, as seeds), and $E(t)$ the number being dispersed away (emigrating) from the population (also likely as seeds).

The relationship in Equation 8.1 above may seem to be a truism. Despite its simplicity, it is a useful starting point because it helps us to focus on some major factors affecting population growth. It also points to some major differences between plant and animal populations, and to some major conceptual issues in plant ecology. These issues first of all have to do with what we consider an "individual" when we study plant populations (Box 8A).

## BOX 8A Genets and Ramets: What Is an Individual?

When ecologists study vertebrates, this problem rarely arises. Each common wombat (*Vombatus ursinus*) or brown pelican (*Pelecanus occidentalis*) is genetically unique, and they are not physiologically connected to one another. While some vertebrates may produce identical twins or triplets, asexual reproduction as a major factor in producing offspring is uncommon among vertebrates (although it is much more common in many groups of invertebrates).

In plants the situation is quite different. Many plant species rely upon asexual reproduction as a major means for producing offpring. Multiple "individuals" may have the same genotype, and they may or may not be physiologically connected to one another. Moreover, the plant life cycle (see Chapter 6) means that, depending on one's purpose, it is correct to think of a mature oak tree, a seedling oak, an acorn, the pollen grain of an oak, and an oak ovule as all being "individuals." (Oak pollen grains and ovules are the individuals of the haploid gametophyte generation of angiosperms; see Chapter 6). But before your head spins because of this complexity, we promise you that it is manageable; we just need to be clear on a few ideas and on what our goals are for any particular study.

Genetically distinct individuals are called genets, while individuals that may be physiologically independent but not necessarily genetically distinct are called ramets (see Figure 9.10). One useful way to keep the distinction straight is that a new (potentially) physically independent individual (a new ramet, also sometimes called a clone) is not necessarily a new genetic individual (a new genet). New genets are formed by fertilization of an ovule and maturation of the resulting seed (see Figure 6.5). In plant species that never spread vegetatively—such as most annual herbs (and most vertebrates)—each genet is a single ramet.

Many plant populations can grow when individual plants produce entirely new plants that are genetically identical to the parent plant. This is often called vegetative or clonal reproduction (in contrast to sexual reproduction; see Chapter 6). Thus, discussing birth and death rates in these plant populations means that we must first decide whether we are studying the change in the number of genets or ramets. If we wanted to know only about genet number, we could consider birth to be the production of new (sporophytic) genets: newborns would consist only of new seeds. Studies of ramet number would count all new individuals, including both new seeds and new plants produced from shoots, because all of them are new ramets. Which definition is best? That depends on what the plant demographer is trying to find out.

This distinction between ramets and genets affects the way we think about dispersal between populations—the $I(t)$ and $E(t)$ terms—as well. Most mature plants are sessile, but seeds are certainly plants, and they are usually mobile. Similarly, pollen grains are (haploid gametophyte) plants that are usually mobile. (A pollen grain is the male gametophyte of gymnosperms and angiosperms, is multicellular at maturity, and is genetically different from the sporophyte plant that produced it; see Chapter 6). It should not be surprising, therefore, that the ecology of seeds and pollen and their dispersal has been and continues to be the subject of a lot of ecological research. But some plants can disperse by clonal fragmentation as well. For example, many cholla cacti (*Opuntia* spp., Cactaceae; see the figure) lose pieces of stems (called joints) that are then moved by animals or runoff from storms; these joints may root and grow into mature plants. At the other end of the moisture gradient, dispersal of clonal fragments by sea grasses (*Zostera* spp., Zosteraceae) may be important in the establishment of new populations (Barrett et al. 1993).

Dispersal by clonal fragmentation has another implication: when a seed disperses to another area, this individual is immigrating from one population to another population. Clearly this is not so with dispersal by fragmentation (if one is counting genets): while the fragment itself leaves one population, the rest of the genet from which it was derived can stay put. This does not mean that dispersal by fragmentation is a "free lunch," however. Indeed, it has an obvious consequence: the fragmenting genet is smaller than it would have been had the dispersing ramets been retained. Since the survival and fertility of ramets usually depend at least in part on their size, it is possible that this sort of clonal fragmentation might eventually affect the survival of the genet as a whole.

Seed germination is not analogous to birth, because seeds are already new individuals—physiologically independent and (usually) genetically distinct from their parents—long before they germinate. Moreover, they often spend a substantial amount of time, sometimes years, living in the soil before germination. Plant populations in which ungerminated seeds persist (usually in the soil) from year to year are said to have a seed bank (see Chapter 6).

Courtesy of S. Scheiner

Many *Opuntia* cacti, especially the chollas, which have cylindrical joints, disperse partly by clonal fragmentation; some of the fragments root and establish new ramets. *Opuntia fulgida*, shown here, is known as jumping cholla because its joints detach very readily; unwary people walking by, including the authors, have been unpleasantly surprised by this plant.

## 8.2 Plant Populations Are Structured by Age, Size, and Developmental Stage

Equation 8.1 implies several equivalent processes that can affect plant population size. For example, the same population growth could occur by an increase in the number of births or by a decrease in the number of deaths. As we will see in much of the rest of this chapter, however, life is not generally so simple. This equation assumes that all individuals in a population are demographically equivalent—have the same probabilities of giving birth, dying, or migrating—and that is rarely true. Large plants usually have different effects on a population than small ones, and old plants usually have different effects than young ones. Usually, adding $x$ births is not equivalent to reducing the number of deaths by $x$. Thus we need to think about a new idea, the **population structure** (its composition in terms of the numbers of individuals of different sizes, ages, or life stages; Figure 8.1).

Structure refers to the frequency of different kinds of individuals; for example, if a population has 1000 individuals and 950 are seedlings, we expect its growth to be different than if 950 were reproductive adults. Most plant demographic studies define population structure in terms of size (e.g., 0–2 cm, 2–4 cm, >4 cm tall), age (e.g., 5–10 years old), developmental stage (e.g., seed, seedling, sapling, mature tree), or combinations of these (e.g., small but old trees) rather than calendar age. Size can be measured in many ways: as mass, as height, or as the number of modules (e.g., branches, tillers in a grass, or leaves) a plant has. Tree size is often measured as **diameter at breast height** (**DBH**), which, curiously, is defined as 1.3 m above the ground in some countries and 1.4 m in others. Other sources of population structure might include morphology or physiological status (rosettes vs. reproductives for herbaceous perennials, for example). In any population in which evolution is occurring by natural selection, there is also structuring by genotype.

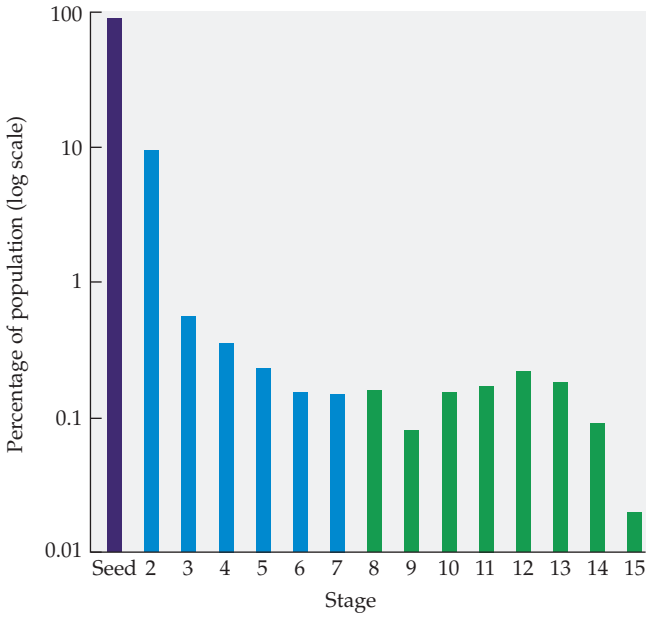



**Figure 8.1** Stage structure of a population of *Pentaclethra macroloba* (Fabaceae), a tropical tree, on Barro Colorado Island, Panama. Stages 2–7 are in increments of 50 cm in height. Subsequent stages are based on diameter at breast height (DBH): stage 8, 2–5 cm; stage 9, 5–10 cm; stage 10, 10–20 cm; stages 11–14 are in 20 cm increments; and stage 15 plants are >100 cm. Thus, 88% of the population are seeds, and 9% are seedlings 0–50 cm in height. (After G. S. Hartshorn. 1975. In *Tropical Ecological Systems*, F. B. Golley and E. Medin [Eds.], pp. 41–51. Springer-Verlag: New York.)

Currently, there are real limits to our abilities to measure these differences, to tell how important they are in the dynamics of a particular population, and to perform appropriate computations. Consequently, in most ecological studies, we focus on only one or a few of these potential sources of population structure. Like most things in science, reality is more complex than our demographic analyses.

While there is a large list of factors that can structure populations, categories such as size, morphology, and age are important for population studies only if we can measure demographic differences among stages. For example, it may be easy to see the difference in size between 30, 60, and 90 cm DBH trees of the same species. It would be a waste of time to use these size differences in a model of population dynamics if size had little effect on survival, growth, or reproductive output.

Most scientific publications on plant demography use the term *stage* to refer to size as well as developmental status, because both kinds of structure are incorporated into models in the same way. (The approach to using stages instead of ages in demographic studies of natural populations goes back originally to studies of insect populations, where *stage* really does mean stage, such as larvae or pupae; plant ecologists adopted the approach and terminology later.) In this book, we say "size" or "developmental stage" when we need to refer specifically to these characteristics, but we use the more general term **state-structured** when we refer to both terms jointly or to other factors like physiological status or genotype that may also structure populations.

Population structure can be important in the demography of any biological population. But structure is central to plant demography because individual plants can vary over orders of magnitude in size, shape, physiological status, and consequently, in their importance for population growth. Animals certainly vary in their

size and contribution to population growth as well. But to a good approximation, we can predict the vital rates of many mammals or birds if we just know their age. Many animal populations can thus be thought of as **age-structured**. To describe such populations, we need to know how many individuals are at each age. The methods used for the study of these populations are **age-based**, requiring only information on the population's age structure. And age-structured populations have a convenient property: a bear that is $x$ years old now will either be dead or $x + 1$ years old next year.

In most plant populations, size plays a stronger role in determining demographic performance than age alone; in other words, plant populations are generally **size-structured** (see Figure 8.1). This should be obvious in the case of reproductive output, because the numbers of flowers and fruits usually depend on plant size. Survival and subsequent growth usually depend on plant size more than on age as well. Older plants are usually larger, but in most plant species there is so much variation in the sizes of plants of a given age (Figure 8.2) that if we need to use only one factor, it is often more useful to use size-based methods than age-based methods to study plant demography.

Chronological age can be important to know for several reasons. Age is quite important in an evolutionary context; for example, if we want to understand how the genetic composition of a population changes over time, we need information about the ages of individuals. Similarly, we often need such information if we want to study the evolution of life history traits (see Chapter 7). As we will see in several examples below, age is sometimes an important determinant of survival and reproduction. Finally, a tree that becomes a local dominant in a forest canopy at 20 years of age is likely to have a different wood structure (and therefore a different chance of surviving) than one that becomes a dominant at 100 years of age. Rapid growth of wood usually results in much less dense, weaker wood than the dense wood that is produced slowly; in this case, we may need to characterize individuals by their age as well as their size.

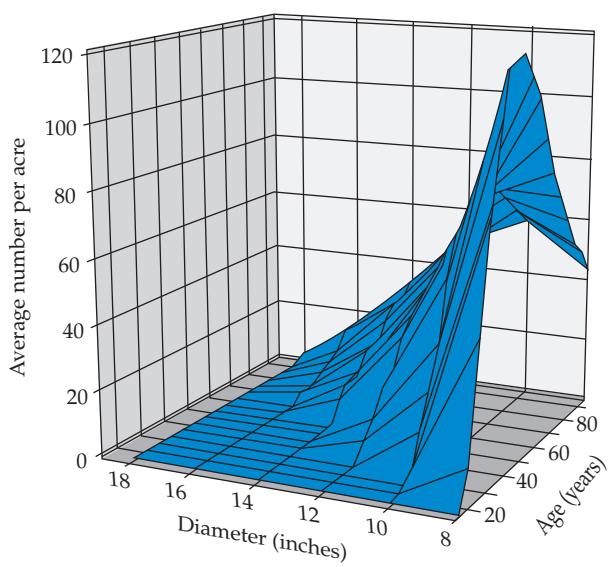



**Figure 8.2** The average relationship between age and size in *Pinus palustris* (longleaf pine, Pinaceae) in the southeastern United States. The graph shows the average number of plants per acre of a particular size, given the age of the stand. Longleaf pine frequently grows in such even-aged stands. (Data from R. D. Forbes. 1930. *USDA Tech Bull* 204, U.S. Department of Agriculture.)

Statistical analyses by Caswell (1988) of data on several plant and animal species pointed to the conclusion that in many plants (and some animals) size alone is a better predictor of demographic performance than age alone. Many plant ecologists drew the conclusion that age in plants is unimportant; this is very often an exaggeration. If size alone is a better predictor than age alone, that does not imply that the two together might not be even better predictors. Chengjin Chu and Peter Adler (2014) studied 19 populations of perennial grasses in the western United States, in which both sizes and ages of individuals were known, so they were able to ask when adding age to statistical models predicting survival and growth improved the models. The answer was "usually"—age improved the fit of 18 of the 19 survival models and 14 of the 19 growth models. Recent theoretical work has involved analyses of models including both size and age (Caswell et al. 2018), although this work has yet to be applied to plant populations. But there is another problem with using plant age in demographic models: it is frequently difficult or impossible to determine the ages of plants. Although many temperate zone (and some tropical) trees do produce reliable annual growth rings (see Figure 13.4), most other plants, both herbaceous and woody, do not. Unless we are working with marked individuals of known ages, then we usually do not know the ages of plants.

## Plant population structure is complicated because plants can change size or form at variable rates

Why are plants so much more variable than animals? A major reason is their modular structure: An individual plant is a system of repeated units, as we saw in Chapter 6. This modularity means that plants have very flexible growth patterns. It also means that they can lose large portions of their bodies but still survive. In other words, plants can actually shrink from year to year. Plants can also go through a year or more of dormancy. For example, in *Cypripedium* (Figure 8.3), some adults in a

Courtesy of J. Delphia

**Figure 8.3** *Cypripedium acaule* (pink lady's slipper, Orchidaceae), a terrestrial orchid from eastern North America. After germination, a corm develops underground for at least 2 years; subsequently, each plant produces zero, one, or two leaves every year it is above ground, but it may go through dormant periods for several years at a time.

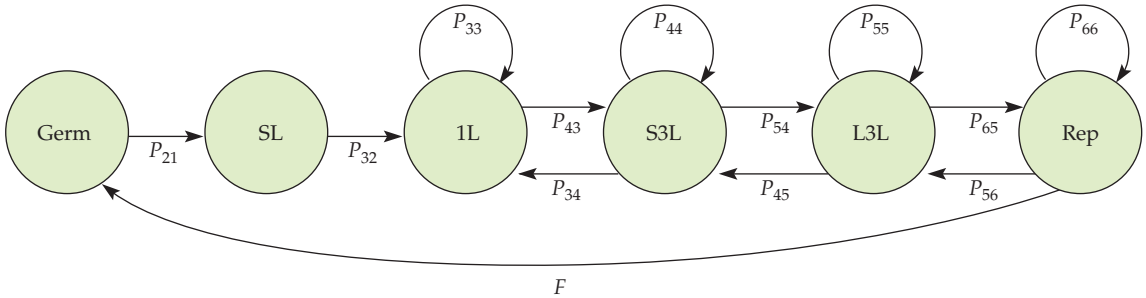
population may not appear above ground in a given year (Cochran and Ellner 1992; Kéry and Gregg 2004).

This is why using size structure and stage structure makes life complicated for plant ecologists. You can be certain that in a year, all individuals now $x$ years old will be either dead or a year older. But in most stage-structured animal populations, individuals can never move to an earlier stage—frogs do not become tadpoles. There are few such prohibitions in plants. Once germinated, a plant can never become a seed, but in most plant species, established plants can grow, stay the same size, or shrink. In a study of the forest floor herb *Trillium grandiflorum* (snow trillium, Liliaceae), Tiffany Knight (2004) found that reversions to earlier stages were common, and several of those reversions were twice as common when the plants were subjected to herbivory from deer (Figure 8.4). Even trees lose branches and sections of their trunks. Consequently, plant ecologists need to keep track of sizes, multiple stages, and the possible transitions between them.

Courtesy of J. Chase

*Trillium grandiflorum*

**Figure 8.4** Life cycle graph for *Trillium grandiflorum* (snow trillium, Liliaceae), a forest floor herb from eastern North America. Such forest floor herbs are among the earliest plants to emerge and flower in the spring (see Chapter 7). The stages used are germinants (Germ), seedlings (SL), one-leaved (1L), small three-leaved (S3L), large three-leaved (L3L), and reproductive (Rep). The arrows show the possible transitions; for example, Germ plants become SL plants $P_{21}$ of the time, and they die $(1 - P_{21})$ of the time. The first two stages are also ages—only first-year plants can be in the Germ class, and they either survive and grow into the SL stage, or they die. However, plants in the S3L, L3L, and Rep classes sometimes move to the next lower class. (After T. Knight. 2004. *Ecol Appl* 14: 915–928.)



## 8.3 Studying Population Growth Usually Involves Models of Changes in Population Structure

Addressing questions about population growth usually requires using models that involve changes in population structure. This might seem surprising. Why can't we take a simple approach: count the number of plants this year, call it $n(t)$, and then count the number next year and call it $n(t + 1)$? The rate of change in population size is then $n(t + 1)/n(t)$. If this ratio is greater than 1, the population size is increasing over the time studied, and if it is less than 1, it is decreasing over that time. In fact, this is how many conservation and wildlife studies estimate population change in animal and plant populations. But it can be a very inaccurate way to evaluate the rate of population growth. In a structured population individuals in different states make different contributions to future population growth. This means that their short-term effects can differ from their long-term effects on the population, and we need to be able to ask about both. We need to model the population's changes in numbers and kinds of individuals, its demography. Such models also can shed light on questions such as why a particular population is growing rapidly or slowly. For example, it might be that survival of large adults is critical for this population—something that would not be evident if we just divided next year's population size by this year's.

Consider a population of a long-lived tree. The oldest trees have a period of senescence that lasts up to several decades, in which they do not produce seeds but slowly lose limbs and experience heartwood rot. Young trees cannot flower until they are several decades old. Thus, neither old nor young trees are reproducing, and a group of 100 very old or very young trees will simply decrease in number for some period of time as some die and none reproduce. Now imagine that year-to-year survival rates are the same among the young and old trees, and compare a population composed entirely of old trees with one composed entirely of young trees. In the short run, both populations will decrease at the same rate. But over time the surviving trees in the "young" population will reach maturity and begin reproducing, and the population can potentially grow in number. By contrast, the "old" population will eventually disappear. This is exactly what can happen during the process of range shifts in plants responding to climate change: old trees may survive for many years after a population has essentially become doomed by failure to produce young. In a different situation, it is occurring in many forests of northeastern North America, where deer overpopulation is resulting in elimination (by deer herbivory) of any new seedlings or saplings, completely eliminating forest regeneration. Therefore, studies of population growth in plants must take into account the structure of the population and must look at both short-term and long-term population growth.

You may find it helpful to think of models in two different ways: they are explicit statements of more general theories, and they are also simplifications of more complex phenomena in nature. Thus they are a way to examine the consequences of our understanding of nature, because models make predictions. Because they are simplifications, no model is "true." The statistician George Box (Box et al. 2005, p. 440) commented that "the most that can be expected from any model is that it can supply a useful approximation to reality: All models are wrong; some models are useful." Much experience with ecological models has shown that, under many circumstances, simple models often lead to more useful insights than more complex, seemingly realistic models (Walters 1986). Obviously a model must be complicated enough to capture important features of the process being studied. The difficulty with models that are too complex is that many more parameters must be estimated, each with some statistical error; the model predictions tend to grow multiplicatively as a result. Think of a model airplane. The simplest ones generally look like a real plane, from the outside. They may be more or less detailed; in any case they can be informative about the shapes of planes. Models that can actually fly are not useful for understanding how the engines or steering mechanisms work in real planes, because they use different sorts of mechanisms, but they can be informative about some aspects of aerodynamics. To understand what is likely to occur with full-size planes operating at much greater speeds, engineers use mathematical or computer simulation models for some issues and may use physical models or actual parts in wind tunnels. All of these are models, and they can be useful—but they are still approximations. Empirical studies with real planes are needed to identify the limitations of the models themselves.

Similarly, population models can take a number of forms; they can be written as equations, graphs, computer programs, or sets of verbal statements. In each case, they are descriptions of assumptions made, based on one or more theories (see Chapter 1). Using rules—such as the rules of mathematics—one can then examine the consequences of those assumptions.

The simplest sort of mathematical model for population growth uses the assumptions that all individuals contribute equally to growth and that they do so in a way that does not change over time. Mathematically this leads to exponential growth, $N(t) = N_0e^{rt}$, or geometric growth, $N(t) = N_0\lambda^{\tau}$, the difference depending only on whether growth is modeled as a continuous process (exponential growth) or as one occurring at fixed intervals like years (geometric growth).

## Life cycle graphs are useful models of plant demography and its relationship to data acquisition

Incorporating population structure requires more complicated models, because individuals of different types may contribute differently to population growth. It is often easiest to begin with a **life cycle graph**, a diagram of how different sorts of individual plants contribute to population growth. A life cycle graph is actually a mathematical model, though it does not involve equations or simulations.

Each node (shaded shape) in the graph in Figure 8.4, as well as those in Figure 8.5, represents a stage class. Some stages here are defined by developmental status (seeds, or dormant individuals), while others are defined by their size within a developmental stage (small juveniles). Arrows between stage classes describe transitions—survival and reproduction—for every interval between censuses. Thus, for the cactus *Coryphantha robbinsorum* (Figure 8.5), the arrow from the large juveniles to the adults represents the fraction of large juveniles at one census that become adults by the next census. The self-loops—arrows from a node to itself—refer to the fraction remaining in the same stage. Thus, the arrow from small juveniles to small juveniles represents the fraction of small juveniles at one census that will still be small juveniles at the next one.

Courtesy of William T. Radke

*Coryphantha robbinsorum*

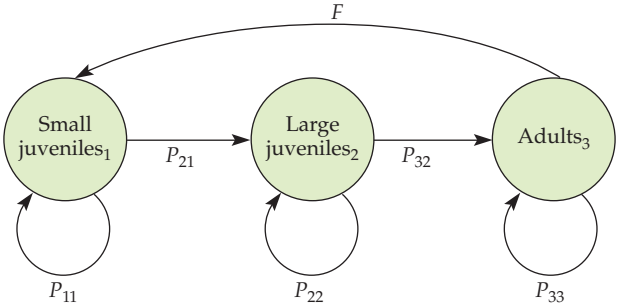



**Figure 8.5** Life cycle graph for *Coryphantha robbinsorum* (pincushion cactus, Cactaceae), an endangered cactus from Arizona and Sonora, Mexico. As this species has no seed bank and plants were censused annually, the graph does not include a seed stage. The transition between adults and small juveniles is thus the product of (average number of fruits per adult) times (average number of seeds per fruit) times (chance of a seed surviving and germinating) times (chance of a seedling surviving to the first census) (see Box 8F for analysis). (After R. J. Schmalzel et al. 1995. *Madroño* 42: 332–348.)

The arrow from adults to small juveniles in Figure 8.5 needs a more careful interpretation, and it points to an important lesson about life cycle graphs and their corresponding demographic models. This arrow refers to the number of small juveniles produced by the adults. We know that in a sexually reproducing species like *C. robbinsorum* every individual starts as a seed and that there is generally some time between seed maturation and germination. A complete biological diagram of the life cycle would include these steps. But we are describing the demography of the life cycle as examined in a real field study. Since this species has no seed bank and censuses are annual, no individuals in the seed class would be counted unless the census were timed to include them. To represent *C. robbinsorum* demography as it is actually measured, then, we need

$F$ = effective fertility = (number of seeds produced per surviving adult) × (chance of seed survival) × (chance of germination) × (chance of seedling surviving to the first census) = (the rate at which adults one year produce small juveniles the next)

If we wanted to study the seed stage per se, we would need censuses that were timed differently, and this would be very difficult in practice because there aren't many seeds, and they are very hard to find and follow once they leave the plant. (People who study seed fates have a number of tricks up their sleeves that depend on the species they are studying, but studies like this are not all that common.)

Life cycle graphs are also useful if we model the different sorts of individual plants as varying continuously rather than in discrete steps. For example, plant size varies continuously; it is often more natural to think of it that way than to force plants into arbitrarily chosen categories like small, medium, and large. Moreover, vital rates like survival or flowering probabilities often depend on size. Then the graph we draw is a description of when reproduction, mortality, and growth (as well as censuses) occur, which dictates how they are combined to make estimates. For example, in Figure 8.6 the life cycle of *Cirsium* (thistle) species is modeled as reproduction, followed by mortality, and then growth. Censuses might occur at any point, but two simple possibilities are that they might occur just before or just after reproduction. The choice between pre- and postreproduction censuses affects how one calculates various quantities; either can be made to work, but it is often simplest conceptually to use prereproductive censuses.

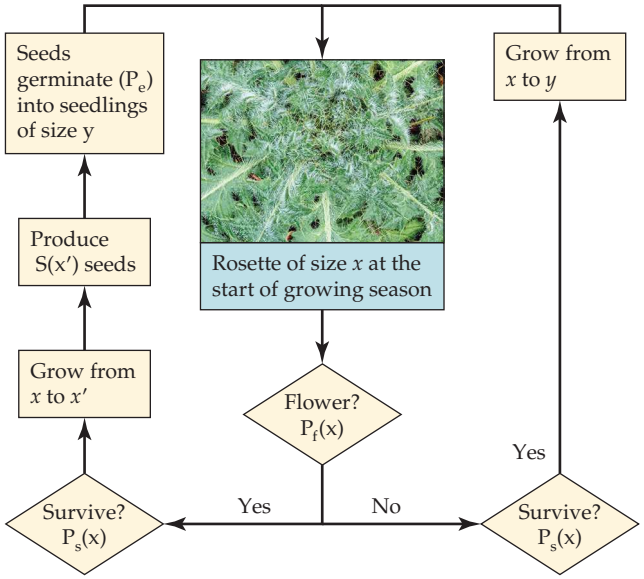



Photo credit: Dominicus Johannes Bergsma/CC BY-SA 4.0

**Figure 8.6** Life cycle graph for *Cirsium* species, used for construction of integral projection models. The diamond-shaped nodes are for probabilistic events that either do or do not happen to each individual (flowering, surviving), while the ovals are for processes (growth and seed production) that are described with regression models. (After B. Tenhumberg et al. 2015. *Ecosphere* 6:1–18. CC BY 3.0.)

## Estimating vital rates can be done several ways

With a life cycle graph in hand, we can use field data to estimate values for the various transitions (see Figure 8.4, Figure 8.5, and Figure 8.6). There are several ways to estimate survival and growth probabilities. The easiest way (conceptually) is to mark or map a large number of individual plants in a way that allows the tracking of individuals, a technique pioneered by Charles Darwin (Harper 1977). These individuals are revisited repeatedly to see which have died and which have survived. If you mark 500 plants at one time and there are 400 surviving at a later time, the survival probability over that time is 400/500 = 0.8. If the initial individuals are of different sizes (you probably wouldn't know their ages), you might want to calculate size-specific survival probabilities. For example, if 300 of the original plants were large and 200 small, and the numbers of survivors were 250 and 150, respectively, large plant survival would be 250/300 = 0.83, while small plant survival would be 150/200 = 0.75. If the plants came in many sizes rather than just two, so size varied on a continuous scale, you could use a technique called logistic regression to estimate the probability of surviving a particular time interval as a function of size.

If you have data on a **cohort** (a group that germinated, reached a particular size, or entered a study at the same time) you can use **life table** methods (Box 8B) to estimate not only survival over a time interval, but cumulative survival over repeated intervals. Originally developed as the basis for life insurance, methods related to life tables are widely used in ecology. There are complications to these methods that are beyond the scope of this book. For example, if you accidentally step on some of your plants, or if some tags have fallen off and you cannot tell what happened to those individuals, some further techniques will be needed.

One important practical problem in estimating vital rates for plants is that many species are hard to find even when mapped and marked, and some life stages—such as the dormant corms of the *Cypripedium* orchids mentioned above—cannot be found at all. In a study by Joslin L. Moore and her associates (2011), potted individuals of *Hieracium aurantiacum* (orange hawkweed, Asteraceae), an invasive plant in Victoria, Australia, were placed in known, randomly selected spots. Searchers (ignorant of the locations) showed highly variable abilities to locate these plants, depending on the length of time they searched, but (surprisingly) not on their personal experience with this species. Accounting for imperfect detection of individual plants may often be important in demographic studies. In recent years, plant ecologists have begun to address this problem by borrowing a statistical technique long used by animal ecologists, the mark-recapture method (Box 8C).

In certain settings, it is possible to estimate survival for larger trees by using standing dead trees ("snags" in forestry jargon). If they can be aged using tree rings, we can infer their age and size at death, and if the rate of decay is low (such as in dry regions like the mountains in western North America), the standing dead and live trees together provide a complete sample of fates. (By examining the pattern of the rings and comparing these with the rings of living trees, their date at death and sometimes their date of germination can also be estimated.) Small trees in the same population probably cannot be studied in the same way, because the decay of these snags is much more rapid.

There is one method for estimating survival that we strongly recommend you not use: estimating survival by using ratios of unmarked individuals. In one version of this, you might count 600 seedlings on your first visit, 400 on your second, and conclude that the seedling survival probability was 400/600 = 0.67. But unless you knew that there was certainly no germination between your visits, this estimate might be quite wrong! In the other version, you might count 300 small plants, 200 medium plants, and 100 large plants and conclude that the chance of a medium plant growing and surviving to large size was 100/200 = 0.5, and the chance of small plants reaching medium size was 200/300 = 0.67. Wrong again! Doing this assumes that rates are constant over time and that the population growth rate is constant and

## BOX 8B How to Construct a Life Table

The main calculations in life tables are relatively straightforward. If we mark $n(t)$ plants at time $t$ and then recensus them 1 month later (time $t + 1$) and find $n(t + 1)$ plants, we can estimate the survival probability as $p_1 = n(t + 1)/n(t)$. For example, if we mark 500 plants initially and find 450 when we recensus, we estimate the survival probability as $450/500 = 0.9$.

The real power of the life table approach—and the part that sometimes causes students to make mistakes—comes from the ability to calculate additional quantities over multiple census periods. Suppose we recensus our plants after still another month and find 425 still alive. Then we can estimate the month 2 survival probability (from month 1 to month 2) as $425/450 = 0.94$. But now we can also estimate the **cumulative survivorship probability**—the chance of surviving from the initial time to our new census. The cumulative survivorship probability to time $x$ is often written as $l_x$ but is also often written as $S_x$; we use the former symbol. For the first time interval, $l_1 = p_1$, but for the second interval, $l_2 = p_1p_2$; in this case, $l_2 = 0.9 \times 0.94 = 0.85$. In other words, we have estimated that 85% of the original 500 plants are still alive.

You might notice that there is another way to perform this calculation that seems simpler: if we just divide the number alive (425) by the initial number (500), we will also get 0.85. *Warning: This is a dangerous way to proceed!* The reason is that in almost all studies, some individuals are simply not located again (e.g., because their tags fall off), die of some cause irrelevant to the study (e.g., because an ecologist stepped on them), or die after the end of the study. Such an individual is represented by a censored data point—we know the individual did not die (in any sense relevant to the study) before some time $c$, but we do not know the actual time of death. We can include these data points in our life table calculation up to the interval in which $c$ occurs. For example, if the ecologist accidentally sat on 10 study plants just after censusing them at the end of the first month, we would still have the same estimate for $p_1$, but now $p_2$ would be estimated as $425/440 = 0.966$. If we used this $p$ value to calculate cumulative survival, we would get the estimate $l_2 = 0.869$, rather than the lower estimate of 0.85 we get by simply dividing the number alive by the initial number. It is always wise to calculate cumulative survival this way, rather than using the seemingly easier (but usually incorrect) method.

To calculate a full life table, it helps to define several variables (the subscript $x$ refers to the time interval $x$):

$n_x$ = number at risk of dying (the number of individuals, after we have accounted for censored data points)
$d_x$ = number dying
$p_x$ = probability of surviving
$q_x$ = probability of dying

These definitions let us write the probability of dying as $q_x = d_x/n_x$, and the probability of surviving as $p_x = 1 - q_x$. The cumulative survivorship probability is $l_x = p_1p_2 \ldots p_{x-1}$.

We can use this equation to estimate the average contribution to the reproductive rate $R_0$ (see Equation 8.3) made by individuals of age $x$—it is simply $l_xF_x$, where $F_x$ is the average number of seeds produced by individuals of age $x$. We can also calculate quantities like the average life expectancy of individuals age $x$.

An example of the life table approach is Susan Kalisz's (1991) study of *Collinsia verna*. She was able to estimate the probability of a plant's surviving each of several stages (seed, seedling, overwintering plant, flowering plant, and fruiting plant), as well as the contribution of each stage to the population's reproductive rate:

Life table methods are used when the researcher defines the time between censuses (and thus determines when deaths are recorded). If the record is continuous, or nearly so, a closely related method called the Kaplan-Meier estimator can be used. There is a very large literature on life tables and their uses. An introduction to statistical comparisons of life tables and related topics can be found in Fox (2001).

| | Duration of stage (month) | $n_x$ | $l_x$ | $q_x$ per month | Number of seeds produced by stage | Average number of seeds produced by stage ($F_x$) | Contribution to reproductive output ($l_xF_x$) |
|---|---|---|---|---|---|---|---|
| Seed | 5 | 13,742 | 1.000 | 0.119 | 0 | 0 | 0 |
| Seedling | 2 | 5,593 | 0.407 | 0.046 | 0 | 0 | 0 |
| Overwintering | 5 | 4,335 | 0.316 | 0.025 | 0 | 0 | 0 |
| Flowering | 1 | 2,582 | 0.189 | 0.112 | 0 | 0 | 0 |
| Fruiting | 1 | 1,056 | 0.077 | 0.077 | 22,725 | 21.520 | 1.657 |

## BOX 8C Borrowing the Mark-Recapture Method from Animal Ecology

In a typical study of mobile animals (such as birds), a scientist captures a number of individuals, marks them, and releases them. At a later date, more animals are captured, some of which will be already marked. The new individuals are marked, and all are released again. Over time, the scientist collects capture histories of a number of individuals, usually represented as a string of ones and zeros: 1 0 1 1 0 0 1 would mean that an individual was captured at times 1, 3, 4, and 7. Obviously, we know that even though we did not capture that individual at times 2, 5, and 6, it must have been alive then. As a result, we can calculate the probability of recapturing an individual if it is alive, as well as survival probabilities over all time intervals. Older studies of animal populations emphasized the use of this approach to estimate population sizes rather than survival probabilities.

Using the mark-recapture method in plants would seem to be trivial. After all, plants do not usually move after you mark them. But, as we have seen, individuals of some species can disappear for a time by remaining underground and dormant, or we can simply fail to find them. We can use the mark-recapture method in these instances to estimate the number and survival rate of those hidden individuals. Using this method for plants is really no different from using it for animals, but plant ecologists have only recently begun to do so. Helen Alexander and her coworkers (1997) first used the mark-recapture method to estimate the population size of a rare prairie plant in Kansas, *Asclepias meadii* (Mead's milkweed, Apocynaceae). In any given year, the number of flowering patches (aggregations of flowering stems) ranged from 15 to 105 at the study site, but the mark-recapture approach allowed the researchers to estimate the actual population size at 175–302 patches. In more recent studies, Marc Kéry and Katharine Gregg (2003, 2004) estimated rates of survival and the fraction of the population that was dormant in several populations of the species *Cleistes bifaria* (rosebud orchid, Orchidaceae) and *Cypripedium reginae* (showy lady's slipper, Orchidaceae).

While the actual estimation of these quantities using mark-recapture methods is a fairly technical subject in applied statistics, there is a large and sophisticated literature on these methods (Lebreton et al. 1992; Kendall and Nichols 2002). We expect mark-recapture methods to play a growing role in plant ecology.

equal to 1. While occasionally we hear of this method still being taught in some ecology classes, we look forward to the day we can pronounce it truly dead. Some of the issues involved in obtaining data to study survival are discussed in Box 8D.

An innovation by James Kellner and Stephen Hubbell (2017, 2018) used remote sensing techniques (see Box 15A) to study both survival and **recruitment** (the combination of processes leading to new individuals) of adult trees of *Handroanthus guayacan* (Bignoniaceae) in Panama. This approach requires that species and individuals be distinguishable. That *H. guayacan* populations occur at low densities made this much easier. Small individuals living below the canopy still cannot be identified with satellite or airplane-based sensors, but you can combine remotely sensed data for canopy individuals and ground-based data for smaller trees. A less obvious (but perhaps more serious) challenge was that Kellner and Hubbell had to develop statistical methods for estimating the numbers of dead or newly recruited canopy trees. An important advantage to using remote sensing for demography is that one can sample large areas, thereby relating variation in demographic parameters to environmental factors. Recently ecologists have begun to collect demographic and other data using drones, which are of course much closer to the ground, much less expensive than photography from planes or satellites, and controlled by the scientist.

Simple estimates of seed survival in seed banks can be made by mixing known numbers of seeds with soil in small fabric seed bags, burying the bags, and digging them up the following year. The soil with the seeds in it is then spread in flats in a controlled environment to see how many seeds germinate; it can be sifted to find the remaining seeds to test them for viability. It can be challenging to study year-to-year changes in actual seed banks. One approach involves simultaneously estimating the fraction of seeds surviving from year to year and the fraction germinating. For example, Margaret E. K. Evans and her collaborators (2007) studied germination and among-year survival in the annual *Oenothera arizonica* and a closely related perennial, *O. californica* ssp. avita (California evening primrose, Onagraceae). They collected multiple soil samples from sites with populations of these plants, timing collection so that germination for the year was completed but new seeds had not yet been produced. They estimated the fraction germinating as

$$\frac{\text{density of seedlings}}{\text{seedling density + density of viable, ungerminated seeds}}$$

## BOX 8D Obtaining Data for Survival Studies

How do plant demographers mark individuals in plant populations? Differing goals will require markers that vary in size, durability, and other characteristics. Large trees can have numbered metal tags hammered into them, and you can expect them to last for many years. With small herbs, you may need to place some sort of marker in the ground next to the plant, for example, metal tags secured to spikes or nails hammered into the ground next to the plant. This may be problematic in dense populations (which plant was the tag referring to?) and in rocky sites. Newly emerged tree seedlings can be marked with prenumbered plastic poultry bands meant for tagging the spindly legs of small chicks quickly and efficiently. The only general rules are that you use a tag that minimizes the effect on the plant itself, and that the markers stay put—they are durable and don't fall off or get moved away. Brightly colored tags can be snatched by crows! Fires will melt plastic tags. PVC pipe cut into smaller lengths can be a good temporary marker for plot boundaries where fire is unlikely, but steel rebar with a numbered endcap is more durable and permanent. GPS (so far) does not have fine enough resolution to map individual plants in many populations. Carefully positioned and leveled photographs may be used for repeated censusing. You may position small low-legged tables with acetate sheets over quadrats and mark plants with permanent markers on the acetate sheets, which you then scan and digitize. There are some settings in which it may not be possible to identify different individuals. If you want to study the demography of a plant that grows in very dense clusters (e.g., a grass with clonal offshoots, or many kinds of shrubs with indistinct boundaries between individuals), you may not be able to mark or map different individuals within a cluster. In such a case, it may be necessary to reformulate your research questions and strategies for marking plants and modeling population change.

How many plants should be marked? For statistical purposes, if we have a choice in determining the sample size, the rule of thumb for survival studies is "Mark more than you think you need, and then mark at least twice as many." For long-lived trees, the probability of death in any given year might be very low, necessitating a large sample size for accuracy. If half the sampled individuals die, the variance of the estimated survival is the largest it can be, and therefore we again need a large sample. For newly germinated seedlings, most might die in the first year. A large sample size may ensure enough individuals among the survivors for estimates of juvenile or adult survival. Also, if all plants in a plot are marked, 99% of the effort may go into marking small recruits, when the most important aspect driving population growth may be the large adults, which may be sparse in number. Strategies that change sampling methods—to reduce the effort on marking small recruits and identify and mark more large adults—may be much more efficient. Sometimes the plants decide the sample size for us. If we are working with a rare or endangered species, there may only be a small number of individuals.

from the samples. Having this estimate and data on the density of seeds produced each year, they were able to estimate the fraction that survived from the prior year. One important caution about this kind of sampling is that seed densities typically vary considerably among samples, even over short distances, so many samples may be required.

Other estimates of survival and germination in among-year seed banks have used more specialized approaches. Susan Kalisz and Mark McPeek (1992) did so destructively: they excavated soil and fumigated it to kill all seeds in the soil. By then replacing the soil and including known numbers of seeds of *Collinsia verna* (blue-eyed Mary, Scrophulariaceae), they were able to estimate the fraction germinating by dividing the number of seedlings by the number of seeds. By excluding new seeds, and by excavating some of their plots, they were able to estimate the fraction surviving to and germinating in a subsequent year. One interesting result was that plants originating from 2-year-old seeds were demographically different from plants originating from 1-year-old seeds. Seed banks may also be structured by other factors; for example, Eugenio Larios and his collaborators (2014) found that in the desert annual *Dithyrea californica* (spectacle pod, Brassicaceae) larger seeds were more likely to germinate than smaller ones, and the resulting plants had greater chances of surviving, and greater fertility, than plants originating as small seeds. They were able to draw these conclusions because in *D. californica* the seed coat stays attached to the root of the resulting plant; one can dig up the plant and measure the seed size.

Estimating reproductive output in plants presents multiple challenges. Most demographic models are on an annual scale, and typically they use estimates either of size-specific seed production (combining everything from flower formation to seed maturation into a single model component) or of size-specific seedling production (further combining seed survival into the reproductive model component). Estimating the seed output of

individuals can be done by following marked individuals (as with survival). Often it is necessary to sample from within individuals: certainly large, long-lived trees often produce many more seeds than can be efficiently counted (or even reached). Even small plants may produce too many seeds to count, particularly for an entire population of plants. In many plants the numbers of seeds per fruit vary; it may be necessary to sample several fruits to know how much they vary and thus how many to include per plant in a sample.

Many demographic studies do something much simpler: they count the number of seedlings and divide by the number of reproductive adults (possibly weighting the large adults as contributing more than small adults). This approach—while widely used—makes the assumption that adults do not vary much in reproductive output, but in numerous studies this has turned out to be untrue. It is also possible to use genetic markers to assign parenthood to individual plants (Grivet et al. 2005); this is particularly useful for understanding dispersal of offspring away from the parent plant as well as for strict demography.

## There are several approaches to building models for structured populations

How do survival and fertility affect the growth of a population as a whole? How do these quantities depend on important environmental variables (such as herbivory or fire), and how does this translate into effects on population growth? Which stages, sizes, or ages have the strongest effects on population growth? Questions like these are important in many contexts, including conservation biology (On which stages should we concentrate protection efforts? What is likely to happen to a population if we change the frequency of fire?), population ecology (Does seedling survival limit population growth?), and evolution (On which sizes of plants can natural selection have the greatest effect?). Understanding these answers is important because, for example, while we might be alarmed by the huge numbers of seeds produced by a plant we think is becoming invasive, it might be that their mortality is so high that few survive to reproduce, and they contribute little to population increase. A disease that kills large numbers of seedlings of another plant species might also make no difference to population change, or it might drive the decline of the species. We need to know the demographic details. One of the strangest things to consider is that, on average, each plant in a stable population produces only one surviving and reproducing offspring over its entire lifetime (think about that for a tree in a forest that produces hundreds or thousands of seeds per year!). Of course, that is on average, and some individuals produce more than one offspring, and many never produce any.

To answer questions like these, we use our estimates of survival, fertility, and the growth rate of individual plants to model the population's growth. With such a model, we can ask how population growth could be expected to respond to changes in different vital rates. Although it can take a fair bit of mathematical savvy to do it right, the basic ideas behind this kind of study are fairly easy to understand.

**Fertility** and **fecundity** are terms that refer to the amount of reproduction. Demographers use *fecundity* to describe the capacity to reproduce, while *fertility* means the amount of actual reproduction. In ecology, the terms are often used inconsistently. We use the demographic definitions in this book.

Ecologists primarily use two types of models for these studies: matrix projection models, and integral projection models (IPMs). IPMs were developed to overcome a particular problem in using matrix models (which are older, and thus are better developed mathematically), and they share many basic ideas (Caswell 2001; Ellner et al. 2016). The basic idea underlying both is to take the population at some time $t$ (say, this year) and project it forward at the next time interval (often called a time step), for example, next year. Matrix models treat all the stages of the life cycle as being discrete (e.g., seeds and seedlings); IPMs allow some parts of the life cycle (usually individual body size) to vary continuously. The important difference is that in predicting the number of individuals, a matrix model predicts the number in a particular class, while an IPM predicts the new continuous distribution of individuals.

There are two principle advantages to IPMs. First, for size-structured populations, IPMs do not require that one make a choice about what size classes to use; instead, survival and fertility are continuous functions of size. This leads to the second advantage: generally, IPMs require many fewer parameters than matrix models. A matrix model requires separate parameters for each class; an IPM just needs the parameters of the function (two for a straight line, three for a quadratic function, etc.). There is no general rule about how the numbers of parameters compare, but in a study of the thistle *Onopordum illyricum* (Asteraceae), Stephen Ellner and Mark Rees (2006) pointed out that their IPM required estimation of 17 parameters (regressions for growth, survival, fertility, flowering probability, probability of seedling establishment, and distributions for seedling size and seedling quality, an unobserved quantity); a matrix model using only two quality classes would require 144! The fact that these functions in IPMs are normally estimated directly from data is an advantage, but (as with any statistic) larger samples lead to more stable estimates (Merow et al. 2014). Matrix models are sometimes constructed with smaller samples, but those estimates may not always be statistically defensible.

While both types of models require some mathematical (and computational) skills, matrix models are easier to understand and simpler to analyze. For this reason, we focus on matrix models. But both types of models are likely to be important parts of our toolkit for some tifme to come.

All of the information needed for building a matrix model is intrinsic to the life cycle graph, as shown in Box 8E.The **population vector** is a list of the numbers of individuals in each stage class at a given time; multiplying it by the matrix of vital rates projects the population one step into the future. If we assume (for the moment) that the birth and survival rates stay constant (so each year we multiply the population vector by the same matrix), we can find several important properties of the population, as we will explain below:

- The short-term and long-term population growth rates
- The population structure at any time in the future
- The reproductive value of each age or stage class
- The sensitivity and elasticity of population growth to changes in the specific probabilities of survival and reproduction
- The relationship between the ages of individuals and their stages, and several other ways of looking at stage structure over time

Roughly speaking, the **reproductive value** of an individual in class $x$ is the expected contribution to future population size of an individual in that class. In other words, reproductive value is a way of evaluating the relative demographic contributions of the different classes. **Sensitivity** tells us how absolute changes (e.g., how adding 0.01 to a survival term) in the survival and fertility of each class affect population growth rates; **elasticity**

## BOX 8E Constructing Matrix Models

To write a matrix model from a life cycle graph, consider *Coryphantha robbinsorum* (see Figure 8.5). The graph tells us that adults generate $F$ small juveniles each year. $F$ is composed of several factors, including the number of seeds produced per adult, the chance that a seed survives until the germination time, the probability of germinating for a surviving seed, and the chance that the new young plant survives until the census. For convenience, we number the stages from 1 to 3 so that every symbol we use can be interpreted. Thus, $n_1$ means the number of small juveniles, $n_2$ the number of large juveniles, and $n_3$ the number of adults. Each transition has two subscripts—the first refers to the stage next year, and the second to the stage this year. Thus, $P_{11}$ is the probability of remaining as a small juvenile, $P_{32}$ is the probability of a large juvenile becoming an adult, and so on. Using these symbols, we can write an equation for the number of small juveniles next year ($t + 1$):

$$\underbrace{n_1(t+1)}_{\text{Number of small juveniles next year}} = \underbrace{P_{11}n_1(t)}_{\text{Number of small juveniles remaining from last year}} + \underbrace{Fn_3(t)}_{\text{Number of new small juveniles}} \qquad (8E.1)$$

Similarly, the number of large juveniles is

$$\underbrace{n_2(t+1)}_{\text{Number of large juveniles next year}} = \underbrace{P_{21}n_1(t)}_{\text{Number of small juveniles from last year that survived and grew to be large}} + \underbrace{P_{22}n_2(t)}_{\text{Number of large juveniles from last year that survived as large juveniles}} \qquad (8E.2)$$

Finally, the number of adults is

$$\underbrace{n_3(t+1)}_{\text{Number of adults next year}} = \underbrace{P_{32}n_2(t)}_{\text{Number of large juveniles from last year that survived and matured}} + \underbrace{P_{33}n_3(t)}_{\text{Number of adults that survived from last year}} \qquad (8E.3)$$

More generally, the number in each class $n_j$ can be written as

$$n_j(t+1) = \sum_{i=1}^{k} n_i(t)\,p_{ji}(t) \qquad (8E.4)$$

where the $p$ terms mean either survival or fertility and $k$ is the number of size classes.

A **transition matrix model** is a compact way of writing the same thing. The matrix collects the coefficients from the model in order (the $F$s and $P$s) and uses them to multiply the vector composed of the number of individuals of each stage. Consequently, the model can be written as

$$\begin{pmatrix} n_1(t+1) \\ n_2(t+1) \\ n_3(t+1) \end{pmatrix} = \begin{pmatrix} P_{11} & 0 & F \\ P_{21} & P_{22} & 0 \\ 0 & P_{32} & P_{33} \end{pmatrix} \times \begin{pmatrix} n_1(t) \\ n_2(t) \\ n_3(t) \end{pmatrix} \qquad (8E.5)$$

An equivalent way of writing this is

$$\mathbf{n}(t+1)=\mathbf{A}\,\mathbf{n}(t) \qquad (8E.6)$$

where **n** is the vector giving the population's structure (the number in each class) and **A** is a transition matrix. In Box 8F, we show how data from *Coryphantha* populations can be used to analyze this general model. In Box 8G, we show how to multiply the population vector by the transition matrix—this exercise demonstrates that the matrix Equation 8E.6 says the same thing as Equations 8E.1, 8E.2, and 8E.3. (More general rules for manipulating matrices are given in Caswell 2001.)

IPMs are quite similar in their general structure; if we replace the summation sign with an integral in Equation 8E.4 and change notation a bit to make it clear that the $p$ terms are size-dependent and the $n$'s are the size distribution of the population, we have an IPM.

tells us how proportional changes (e.g., how increasing a survival term by 1%) yield proportional changes in the population growth rate. Estimates of growth rates, reproductive value, sensitivity, and elasticity are therefore important tools in evolutionary ecology, conservation biology, and applied ecology; Box 8F shows an application to the endangered cactus *C. robbinsorum*. All of these quantities are part of the basic analysis of an IPM as well.

## Analyzing demographic models gives information on population growth rates and population composition

How do we get all of this information from demographic models? In this section, we introduce some major ideas used in analyzing matrix models and IPMs. The basic method depends on an important observation from matrix models: If we multiply the population vector by the matrix repeatedly (see Box 8G), after a while the population reaches a **stable structure** or **stable stage distribution**, at which point the proportion of individuals in each class stays constant over time, although the population keeps growing (Figure 8.7). This implies that once the population has reached its stable structure, we can multiply the population vector by a number (a scalar) rather than by the entire matrix and get the same result as if we were multiplying by the matrix. Mathematically, this means that we are looking for the value of λ (the Greek letter lambda), based on the equation $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$, where **A** is the transition matrix, **x** is a vector, and λ is an eigenvalue of **A**. (In this book, we use standard mathematical notation: capitalized boldfaced symbols are matrices, like **A**, and lowercased boldfaced symbols are vectors, like **x**.) It is always possible to find the values for λ for transition matrices and IPMs if the model is written correctly; this is almost always done numerically on a computer. Numbers that can play this role are called **eigenvalues** or **characteristic values**. Every such number λ has a corresponding population vector **x**, called an **eigenvector** or **characteristic vector**.

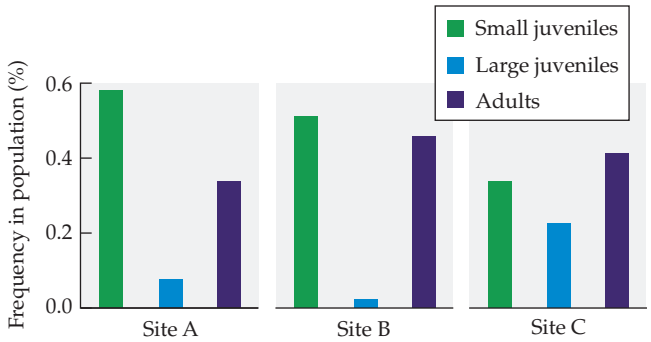



**Figure 8.7** Stable stage distribution for the cactus *Coryphantha robbinsorum* at three sites, based on the matrix models in Box 8F. Sites A and B (on hillsides) have similar stable population structures. The stable structure at site C (on a hilltop) has many more large juveniles. At this site, 33% of small juveniles become large juveniles, compared with only 1%–2% at the other two sites. This is also the principal factor causing λ at site C to be so much larger (1.12) than at the other two sites (where it is just less than 1). (After G. A. Fox and J. Gurevitch. 2000. *Am Nat* 156: 242–255.)

## BOX 8F Demography of an Endangered Cactus

*Coryphantha robbinsorum* is a small, cluster-forming cactus found on limestone outcroppings in southern Arizona and adjacent Sonora, Mexico. Robert Schmalzel and colleagues (1995) marked plants at three sites on a hill and followed their growth, reproduction, and survival for a 5-year period. Because adults (plants that had flowered at least once), small juveniles (plants <11 mm in diameter), and large juveniles had significantly different chances of survival and reproduction, the authors used these and seeds as their stage classes.

The life cycle graph for *C. robbinsorum* is shown in Figure 8.5. Since *C. robbinsorum* has no seed bank and censuses were annual, we reanalyzed the authors' data with a model that does not include seeds. The data can be arranged in matrices for the three sites:

Site A (northeastern exposure)

$$\mathbf{A}_{\text{site A}} = \begin{bmatrix} 0.672 & 0 & 0.561 \\ 0.018 & 0.849 & 0 \\ 0 & 0.138 & 0.969 \end{bmatrix} \quad (8F.1)$$

Site B (southwestern exposure)

$$\mathbf{A}_{\text{site B}} = \begin{bmatrix} 0.493 & 0 & 0.561 \\ 0.013 & 0.731 & 0 \\ 0 & 0.234 & 0.985 \end{bmatrix} \quad (8F.2)$$

Site C (hilltop)

$$\mathbf{A}_{\text{site C}} = \begin{bmatrix} 0.434 & 0 & 0.561 \\ 0.333 & 0.610 & 0 \\ 0 & 0.304 & 0.956 \end{bmatrix} \quad (8F.3)$$

Analysis of these matrices reveals that the population at site C is growing, as λ = 1.12. At the other two sites, the population is stable or declining slowly: λ = 0.998 at site A and λ = 0.997 at site B. The stable stage distributions differ among these plots (see Figure 8.7). Most notable is the predicted scarcity of large juveniles at sites A and B. The sites also differ in the rates at which they approach the stable distribution (see Figure 8.8). Sensitivities are shown in Table 8.2, elasticities in Figure 8.9, and reproductive values in Table 8.1.

## BOX 8G Multiplying a Population Vector by a Matrix

To understand how Equation 8E.6 says the same thing as the equations 8E.1, 8E.2, and 8E.3, you need to know how to multiply matrices. Matrix multiplication is "row by column." To get the first element in the population vector for next year (the number of small juveniles), use the coefficient in the first row, first column ($P_{11}$), to multiply the first element in the vector for this year [$n_1(t)$] to get $P_{11}n_1(t)$. Then use the coefficient in the first row, second column (0), to multiply the second element in the vector [$n_2(t)$] to get 0. Finally, use the coefficient in the first row, third column ($F$), to multiply the third element [$n_3(t)$] to get $Fn_3(t)$. Then sum these three products to get $n_1(t + 1) = P_{11}n_1(t) + 0 + Fn_3(t)$, which is exactly what equation 8E.1 says.

To find the second element in next years' vector (the number of large juveniles), you repeat this process, but multiply each of the vector elements by the coefficients in the second row of the matrix.

Another useful way of thinking of the matrix is that it describes the transitions in "from-to" form. The matrix element in the *i*th row and *j*th column always refers to the transition from the *j*th stage to the *i*th stage—from the "column-number" stage to the "row-number" stage.

But what are these eigenvalues and eigenvectors? There are several eigenvalues, which together are components of the population's growth rate, and several eigenvectors, which are components of the population's structure. A very important result is that one eigenvalue is always larger than the others for this type of matrix. As a result, over time, the population growth rate approaches this value and is therefore called the **stable growth rate**, and the population structure approaches the associated eigenvector (the stable stage distribution; Figure 8.8). The eigenvalues are often ranked by their size, so $\lambda_1$ usually refers to the **dominant** or **leading** (largest) eigenvalue and $\lambda_k$ to the smallest. When authors discuss $\lambda$ without any subscripts, they usually mean the dominant eigenvalue. This number $\lambda$ tells us some essential things about the population. If $\lambda > 1$, the population is increasing; if $\lambda < 1$, it is decreasing. The population remains at a constant size only in the special case that $\lambda = 1$.

Even when the population is not yet near its stable distribution, its growth can be predicted using the eigenvalues and eigenvectors (Caswell 2001). The population's size at any time in the future can be written as a weighted sum of the products of the eigenvalues and eigenvectors:

$$\mathbf{n}(t) = c_1\lambda_1^t\mathbf{x}_1 + c_2\lambda_2^t\mathbf{x}_2 \ldots = \sum_{i=1}^{k} c_i\lambda_i^t\mathbf{x}_i \qquad (8.2)$$

where the *c*'s depend on the initial population structure and $\mathbf{x}_i$ is the *i*th eigenvector. Raising the eigenvalues to a power (equivalent to repeated multiplication by the matrix) makes the smaller eigenvalues decline in relative importance over time. Recent studies have taken a look at matrix methods for studying populations that are far from their stable distribution; methods for these analyses were developed by Gordon Fox and Jessica Gurevitch (2000), and subsequently by Caswell (2007). (Fox and Gurevitch 2000; Caswell 2007). Population

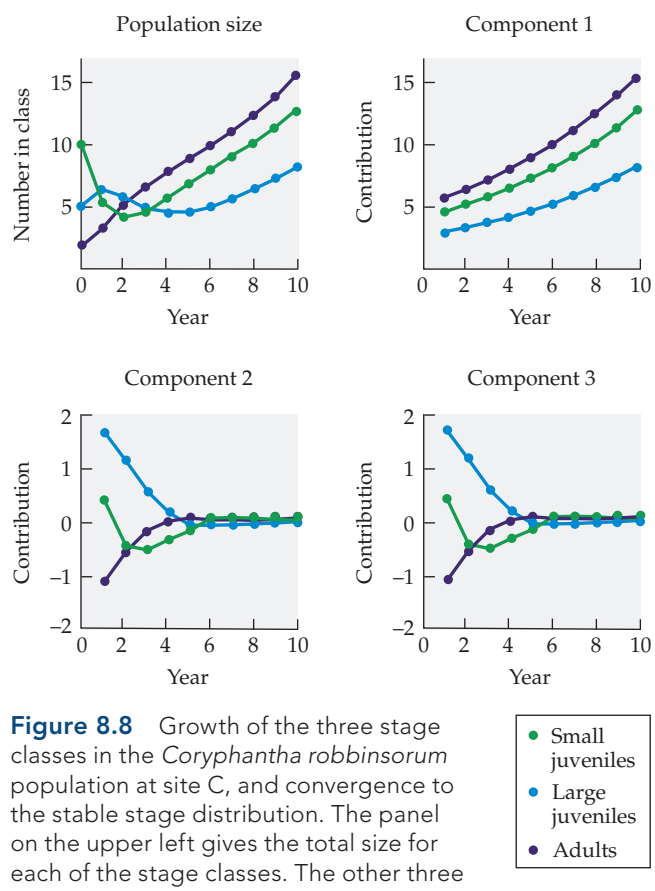
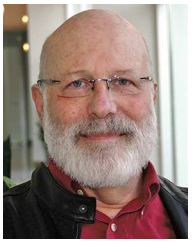



**Figure 8.8** Growth of the three stage classes in the *Coryphantha robbinsorum* population at site C, and convergence to the stable stage distribution. The panel on the upper left gives the total size for each of the stage classes. The other three panels show the contribution of each pair of eigenvalues/eigenvectors. By the year 8, the size change for each class is within 0.01 of $\lambda = 1.12$; only the dominant eigenvalue is making a substantial contribution at this point. (After G. A. Fox and J. Gurevitch. 2000. *Am Nat* 156: 242–255.)


Courtesy of Gordon Fox

Gordon Fox

growth rate can fluctuate before a population gets to a stable distribution. These changes are called transient dynamics and their analysis is called **transient analysis**. It is important because many populations are subject to frequent disturbances and never reach a stable distribution or constant population growth rate. Martha Ellis and Elizabeth Crone (2013) studied 25 matrices of nine perennial plant species and concluded that transient dynamics contributed more to the year-to-year variability in growth rates than did variation in vital rates alone.

## Measuring lifetime reproduction gives us the net reproductive rate of the population

How many offspring can an average individual expect to produce in its lifetime? For age-structured populations, this quantity is easily estimated from the data in a transition matrix, using some simple mathematics. Use $l_x$ (see Box 8B) as the probability of surviving from the first census to the $x$th census. For a group of marked individuals, $l_x$ cannot increase over time—if 1000 individuals are censused starting after their birth, then at each subsequent census, the number surviving must be the same as at the prior census, or smaller. We can use $F_x$ to represent the fertility of an individual that is in age class $x$. Averaging over the whole population, a newborn individual can expect to have $l_xF_x$ offspring at age $x$, $l_{x+1}F_{x+1}$ offspring at age $x + 1$, and so on. Over its entire life, then, an individual can expect to have $R_0$ offspring:

$$R_0 = l_1F_1 + l_2F_2 + \ldots l_xF_x + \ldots = \sum_{i=1}^{\infty} l_iF_i \qquad (8.3)$$

$R_0$ is called the **net reproductive rate**; it is the average number of offspring a newborn can expect to have over its lifetime. In general, it is *not* equal to λ (the rate of growth of the population, per year or other time unit being studied) unless the population is at equilibrium—that is, the population is not growing or declining; then $R_0$ and λ are both equal to 1 and on average an individual replaces itself by leaving one offspring over its lifetime. $R_0$ is obviously a useful quantity to know. We have defined it in terms of age; there are also (more complicated) methods for calculating it from a **stage-structured** matrix (Cochran and Ellner 1992).

## Reproductive value is the contribution of each stage to population growth

Individuals of different stages do not make equivalent contributions to future population growth. In the study of the *Coryphantha* populations we have described, for example, most small juveniles died, while most adults survived for a long time. If a reserve manager, say, wanted to establish a new population, it might be better to introduce 100 adult plants than 100 seeds or small juveniles (if enough plants could be raised to adults in "captivity" and were available, and if they could survive transplanting well enough). But since seeds are generally much easier to work with, one might be tempted to introduce them. Instead of guessing how many individual seeds or adults to introduce, it would be helpful to have a way of measuring the effects of different kinds of individuals on future population growth. The reproductive value of the different stages or ages gives us precisely this information. The reproductive value (RV) of stage $x$ is the expected contribution of an individual now in stage $x$ to the future population size. Reproductive value depends on the rate of population growth as well as the expectations for survival and fertility of individuals at each age (or stage). Most literature uses a relative RV, in which the RVs are all divided by the RV of a newborn. In Box 8H we develop the idea of reproductive value for age-structured populations and then extend these ideas to stage-structured populations.

There are some important differences between age- and stage-structured populations in the pattern of reproductive value through the life cycle. In most age-structured populations, reproductive value at birth is very low (because reproduction is delayed for some time, and many newborns die before they reach maturity). Reproductive value increases to a maximum near the age of onset of reproductive capability, and then it decreases. In stage-structured populations, in which individuals can remain reproductive for a long time—as in many perennial plant populations—a decrease may not occur.

The *Coryphantha* data make this clear. The reproductive values of each stage at each site are given in Table 8.1. At site A, a plant that makes it to the adult stage contributes roughly 19 times as much, on average, to the next generation as an average small juvenile does,

**TABLE 8.1** Reproductive values ($v_x$) of stage classes of *Coryphantha robbinsorum* at three sites

| Site | Small juveniles | Large juveniles | Adults |
|---|---|---|---|
| A | 1 | 17.82 | 19.26 |
| B | 1 | 39.72 | 45.22 |
| C | 1 | 2.06 | 3.44 |

*Note*: These reproductive values are calculated as the dominant left eigenvectors of the transition matrices in Box 8F, standardized so that small juveniles have $v_1 = 1$.

*Source*: G. A. Fox and J. Gurevitch. 2000. *Am Nat* 156: 242–255.

## BOX 8H Reproductive Value

It might seem at first that reproductive value could be estimated by summing the quantities in $R_0$ (see Equation 8.3) over a shorter interval—for example, for an individual of age $x$, one might start the sum at $x$ instead of 1. However, this would not provide the information we need. Summing the average reproduction from age $x$ on will always give a smaller number than summing over an entire lifetime—it provides only the number of offspring a newborn can expect to have from age $x$ on, instead of the number expected over its whole life. Moreover, this sum over the shorter period does not account for the fact that the population will have changed in size by the time a newborn reaches age $x$. Its future reproduction will have a different effect on the population than its present reproduction, because a newborn that begins its life sometime in the future will be part of a larger (or smaller) population. Calculating the effect of individuals of age $x$ on future population growth therefore requires accounting for the chance of surviving to age $x$, as well as for the change in population size in that interval.

This leads us to a second, and more precise, verbal definition of reproductive value (RV):

$$\text{RV} = \frac{\text{proportion of future births in the population to individuals now age } x}{\text{proportion of the population now age } x} \qquad (8\text{H.1})$$

Both the numerator and denominator of Equation 8H.1 are related to $R_0$ (see Equation 8.3), but they also take into account the rate at which the population is growing. To see how this works, consider a population with $n$ individuals, growing at the rate $\lambda$. A single offspring born to a parent of age $x$ now constitutes $1/n$ of the population. A single offspring born in the next time interval will constitute $1/(\lambda n)$ of the population, one born two time intervals later will constitute $1/(\lambda^2 n)$ of the population, and so on. In a growing population, offspring born later in life will contribute less to the next generation than offspring born earlier. The reverse is true in a decreasing population.

Thus, to calculate the numerator in Equation 8H.1—the proportion of future births to individuals now age $x$—we need to sum the expected future reproduction at each age (from $x$ to death), as in calculating $R_0$, but we need to divide each element $l_iF_i$ by $\lambda^i$:

$$\text{numerator} = \sum_{i=x}^{\infty} \frac{l_iF_i}{\lambda^i} \qquad (8\text{H.2})$$

The denominator in equation (8H.1) —the proportion of the present population that is age $x$—can be found using similar reasoning. Individuals now age $x$ were first recorded $x - 1$ censuses ago, and $l_x$ of the original cohort was still alive then. Because the population has grown by a factor of $\lambda^{x-1}$ since then, we have

$$\text{denominator} = \frac{l_x}{\lambda^{x-1}} \qquad (8\text{H.3})$$

Putting the numerator and denominator (see Equations 8H.2 and 8H.3) together, the final expression for the reproductive value (the expected contribution to future population growth of an individual of age $x$) is

$$v_x = \frac{\lambda^{x-1}}{l_x} \sum_{i=x}^{\infty} \frac{l_iF_i}{\lambda^i} \qquad (8\text{H.4})$$

(Goodman 1982). Thus, all of the information needed to calculate reproductive value is contained in a transition matrix or a life cycle graph.

These ideas apply to stage-structured populations as well. We can think of the reproductive value of a tree that is 5 m tall, or that of a dormant seed in the seed bank. But a moment's thought should suggest a problem: in a stage-structured model, it is difficult to calculate the numerator of Equation 8H.1. For example, our 5 m tree may grow continually, stay 5 m tall for years, or shrink for a time.

There is an alternate way of calculating reproductive value that gets around this problem. The reproductive values for each age or stage can be calculated as the dominant **left eigenvector** of the matrix model. A left eigenvector $\mathbf{y}$ of the matrix $\mathbf{A}$ is defined so that $\mathbf{yA} = \lambda\mathbf{y}$. This definition is parallel to that for the (right) eigenvector that we used above in discussing the stable stage distribution. (When biologists refer to an eigenvector of a matrix without any further qualification, they usually mean a right eigenvector.)

but only slightly more than an average large juvenile. At all three sites, the reproductive value of an adult is much greater than that of a small juvenile, but only a bit larger than that of a large juvenile, because relatively few small juveniles survive to reproduce, but most large juveniles do. Reproductive value does not decrease in the last stage, because adults can survive for indefinite periods, as expected in stage-structured plant populations in which there is no senescence. This information is precisely what would be needed if a manager were planning to introduce plants to augment this population. For example, if transplanting nursery-bred plants were practical, the data suggest that introducing large juveniles would be nearly as effective as introducing adults, which would allow the manager to save time, money, and effort in growing the plants. Reproductive value also tells us something about evolution by natural selection (Chapter 9),

because the response to natural selection is greatest for individuals with the highest RV.

### Sensitivity and elasticity indicate how individual matrix elements affect population growth

How does λ change as the individual matrix elements (or coefficients in an IPM) change? We already have all the information we need to answer this question from the eigenvectors of **A** and the reproductive values of each stage. In Box 8I we show how to use this information to calculate the sensitivity of λ to changes in each matrix element.

The sensitivities of λ to changes in the matrix elements in the three *Coryphantha* populations are shown in Table 8.2. The element in the *i*th row and *j*th column represents the rate at which λ changes as that element in the transition matrix **A** changes, with all other matrix elements held constant. The sensitivities at sites A and B are quite similar. In both cases, the largest change in λ would be expected if there were increases in small juveniles' chances of surviving and moving into either of the larger classes and if there were increases in the survival of adult plants. In contrast, increasing the fertility term—the rate at which adult plants generate new small juveniles—would have little effect on λ in either population. Site C (where the population is clearly growing) is a bit different. Increasing any of the terms would have a marked effect on λ, although the largest effects would still be achieved by increasing survival rather than fertility.

**TABLE 8.2** Sensitivities of *Coryphantha robbinsorum* long-term population growth rates (λ) to changes in the matrices

| | | Sensitivities | | |
|---|---|---|---|---|
| Site | | Small juveniles | Large juveniles | Adults |
| A | *Small juveniles* | 0.0695 | 0.0085 | 0.0404 |
| | *Large juveniles* | 1.2391 | 0.1521 | 0.7203 |
| | *Adults* | 1.3390 | 0.1643 | 0.7784 |
| B | *Small juveniles* | 0.0230 | 0.0011 | 0.0206 |
| | *Large juveniles* | 0.9120 | 0.0435 | 0.8200 |
| | *Adults* | 1.0383 | 0.0495 | 0.9336 |
| C | *Small juveniles* | 0.1526 | 0.0999 | 0.1866 |
| | *Large juveniles* | 0.3138 | 0.2054 | 0.3837 |
| | *Adults* | 0.5251 | 0.3437 | 0.6420 |

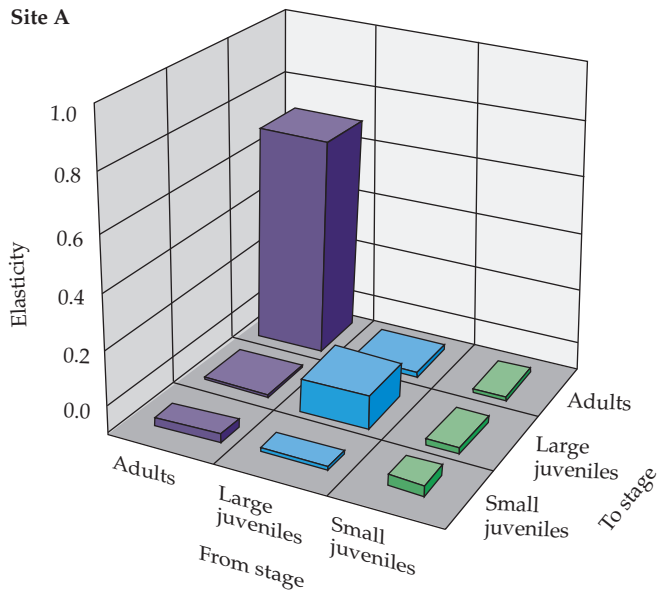
*Note*: Sensitivities and λ are calculated from the transition matrices in Box 8F. Each entry shows the change in λ resulting from a unit change in the corresponding element of the transition matrix while all other matrix elements are held constant. For example, a unit increase in $P_{11}$ increases λ at a rate of 0.0695 at site A.

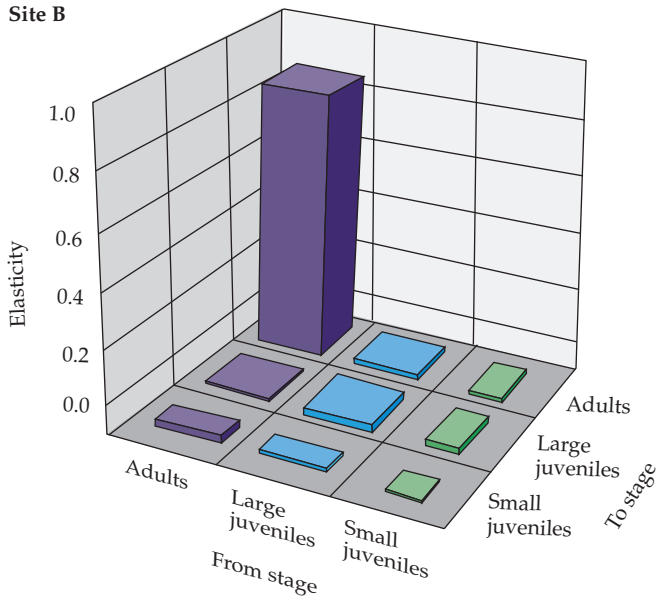
There is a potential problem with sensitivities. They compare things that are often measured at very different scales. For example, the survival terms in a matrix model must be between 0 and 1, but the fertility terms can sometimes be orders of magnitude larger. Similarly, the survival terms for small plants are often much smaller than those for larger plants. When this is true, a small increase of, say, 0.01 actually represents a much larger *proportional change* in the smaller terms than in the larger terms. A related problem with sensitivities is that matrix elements may be zero for some basic biological reason, or by definition. In the *Coryphantha* example, it is impossible for juveniles to reproduce (by definition)—but the sensitivity analysis still tells us that λ would be increased by a certain amount if that transition were increased from zero. To handle these issues, we can calculate elasticity, the proportional change in λ caused by proportional changes in a matrix element (see Box 8I).

For the *Coryphantha* data, the elasticities are given in Figure 8.9. At sites A and B, the largest proportional effects on λ would be achieved by increasing adult survival. All other changes would have very small effects on λ. At site C, λ would still be most strongly affected by increasing adult survival. Increasing the other terms would have smaller, but perhaps not negligible, effects. In any case, it seems clear that at all three sites, an effective plan to protect the population would emphasize protecting the established plants—especially adults—rather than enhancing their fertility. This result is common for organisms with long-lived adult stages.

Elasticities have an additional property that makes them useful—all the elasticities of a matrix sum to 1. Consequently, they can be interpreted as the relative contribution to λ of the corresponding matrix element (given that all other elements stay constant). This means that one can directly compare the elasticities of different matrices that have the same life cycle graph. For example, at *Coryphantha* site C, adult-adult survival was responsible for about 55% of λ. By contrast, at the sites where the population is shrinking or just holding its own, adult survival accounted for a much greater proportion of the population growth rates—76% at site A, and 92% at site B.

Information on sensitivities and elasticities of population growth can be critical in guiding research on and management of endangered and threatened species, fragmented habitats, and invasive species. However, we often don't have the data we need to do this. A survey of research on the demography of plants in



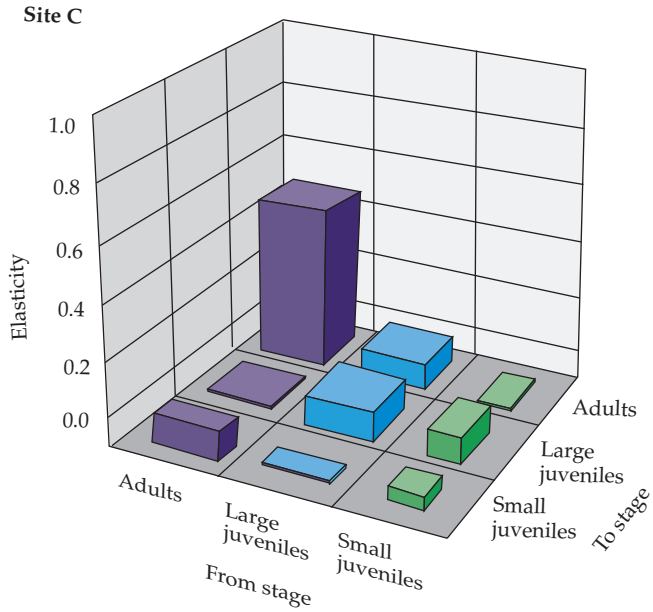




**Figure 8.9** Elasticities for the three populations of the endangered cactus *Coryphantha robbinsorum*. Each bar gives the elasticity for a particular transition corresponding to the matrices in Box 8F. The elasticities describe the proportional change in the long-term growth rate λ resulting from a proportional change in each matrix element, holding all other matrix elements constant. They sum to 1 for any given analysis and can therefore be thought of as describing the proportional importance of each matrix element for λ. For example, at site A, adult-adult survival ($P_{33}$) accounts for 76% of λ, and at site B it accounts for 92% of λ. At site C, where the population is expected to grow rapidly (λ = 1.12), adult-adult survival makes only about a 55% contribution to λ.

fragmented habitats by Emilio Bruna and his collaborators (Bruna et al. 2009) revealed that in general, growth and survival of large, established individuals has the greatest elasticity, but we have the least information on those individuals. Most field research focuses on reproduction and dispersal; the authors suggested that this is because it is relatively easy to conduct manipulative experiments with reproduction and dispersal, but difficult with large individuals.

## Life table response experiments can examine the demographic differences among populations

Ecologists are often interested in asking how environmental differences contribute to changes in λ. For example, we might want to know not only how λ varies between fragmented and unfragmented forests, but also what components of λ are changing. If we have a set of estimated matrices from different environments, we can use **life table response experiments (LTREs)**, a method developed by Hal Caswell (2001), to ask just this kind of question. LTREs are not experiments in the usual meaning of the word; rather, they are a way to examine and analyze matrix demographic data to answer questions like this.

For example, Emilio Bruna and Madan Oli (2005) studied the tropical rainforest herb *Heliconia*

Courtesy of Hal Caswell

Hal Caswell

## BOX 8I How Do Changes in Transition Probabilities Affect the Population Growth Rate?

Call $a_{ij}$ the matrix element in the $i$th row and $j$th column. The sensitivity of λ to changes in $a_{ij}$ is

$$\frac{\delta\lambda}{\delta a_{ij}} = \frac{v_i x_j}{\sum_{i=1}^{N} x_k v_k} \qquad (8I.1)$$

where the $x$'s are the elements of the dominant eigenvector (so $x_j$ is the $j$th element of the dominant eigenvector), the $v$'s are the reproductive values (so $v_i$ is the reproductive value of stage $i$), and δ denotes a partial derivative. In other words, a small change in $a_{ij}$ (the rate at which stage $j$ individuals generate stage $i$ individuals) causes a change in the long-term growth rate. The magnitude of this change is proportional to the relative importance of stage $j$ in the stable distribution ($x_j$) times the reproductive value ($v_i$) of stage $i$ individuals.

If you have forgotten (or never knew) what a partial derivative is, think of it as being like an ordinary derivative (that is, a rate of change over a very small change in something else, such as time), but one in which we temporarily hold everything else (all other variables) constant. For example, if $f = ax + by^2 + cxy$, then

$$\frac{\delta f}{\delta x} = a + cy \qquad (8I.2)$$

In other words, $f$ changes with respect to $x$ by $a$ (because $a$ multiplies $x$ in the first term in the sum) plus $cy$ (because $cy$ multiplies $x$ in the third term; we are momentarily acting as though $y$ is constant). Using this logic, you should see that the partial derivative of $f$ with respect to $y$ is

$$\frac{\delta f}{\delta y} = 2by + cx \qquad (8I.3)$$

The elasticity (proportional sensitivity) of λ with respect to changes in the $ij$th matrix element is

$$e_{ij} = \frac{\delta \ln(\lambda)}{\delta \ln(a_{ij})} = \frac{a_{ij}}{\lambda}\frac{\delta\lambda}{\delta a_{ij}} \qquad (8I.4)$$

(Caswell 2001). In other words, the elasticities are the sensitivities times $a_{ij}/\lambda$. Thus, when a matrix element is zero, its elasticity will also be zero.

*acuminata* (Heliconiaceae) in the Amazon rainforest near Manaus, Brazil, as part of a larger study of the consequences of rainforest fragmentation, the Biological Dynamics of Forest Fragments Project (see Figure 15.14). They found that, on average, in continuous forest, λ is about 1.05, while in forest fragments of 10 ha or 1 ha, λ is about 1. Using the LTRE method, however, they found that different mechanisms caused the reductions in the population growth rate in different-sized fragments. In the 10 ha fragments, the reduced growth rate was caused by reduced reproduction, while in the 1 ha fragments, changes in the rate of individual plant growth also contributed to the reduction in the population growth rate. In another study, Ingrid Parker (2000) studied the invasive shrub *Cytisus scoparius* (Scotch broom, Fabaceae) in a number of sites in and near Seattle. All populations had λ greater than 1, but those in city parks were growing much more slowly than those in prairies. Using the LTRE approach, she found that the principal cause of the difference was the greater chance of seedling establishment in the prairie habitat.

### Ecologists are beginning to study demography at larger spatial scales

Matrix models have been used extensively for plant population studies since the 1980s. A large database (COMPADRE) of all plant demographic matrix data that have been published and meet a set of standards has been compiled and made public (Salguero-Gómez et al. 2015). As of early 2017, COMPADRE included over 7000 matrix models from over 800 studies of nearly 700 species. This large data set promises the possibility of addressing important general questions. Nevertheless, a note of caution is in order; in a recent analysis of animal population models in the related COMADRE database, Bruce Kendall and his collaborators (2019) found that there were common errors in modeling that can lead to incorrect conclusions; there is every reason to expect that the COMPADRE database may involve similar problems. Large databases and "big data" from molecular biology to global data often have many errors embedded in them, and it is worth approaching them carefully!

In 2015, an international group of ecologists including Yvonne Buckley began a large, distributed study of the population ecology of *Plantago lanceolota* (Plantginaceae), an herb native to Asia and Europe but now widespread around the world. As of 2019, there were 56 study sites around the world, ranging from subtropical (Queensland, Australia) to alpine (Switzerland) and subarctic (Sogndal, Norway). The goals of this collaboration (called PlantPopNet; Buckley et al. 2019) are to understand the factors that drive

© Viv Buckley
Yvonne Buckley

demography in this species across both its native and introduced ranges, using a single study protocol. As the project is still young, it has not yet published results, but studies on this large scale are new in ecology, and we expect that it will lead to new understanding of demographic processes and inspire more broad collaborative research networks. Landscape demography (see Chapter 15) is an approach to studying the demography of populations that may be spread across a landscape, and it may be helpful where an ecologist needs to consider population demography at different spatial scales.

### There are additional approaches to modeling plant demography

The matrix and IPM approaches can be much more sophisticated than the basics we have outlined here. For example, they can include random fluctuations in the environment and random variation in individual performance. They can also be made density-dependent. Doing so requires more advanced mathematics than we will use in this book, and the analysis no longer depends simply on the eigenvalues and eigenvectors. It is important to know that these methods exist and can address more complex questions and conditions affecting populations.

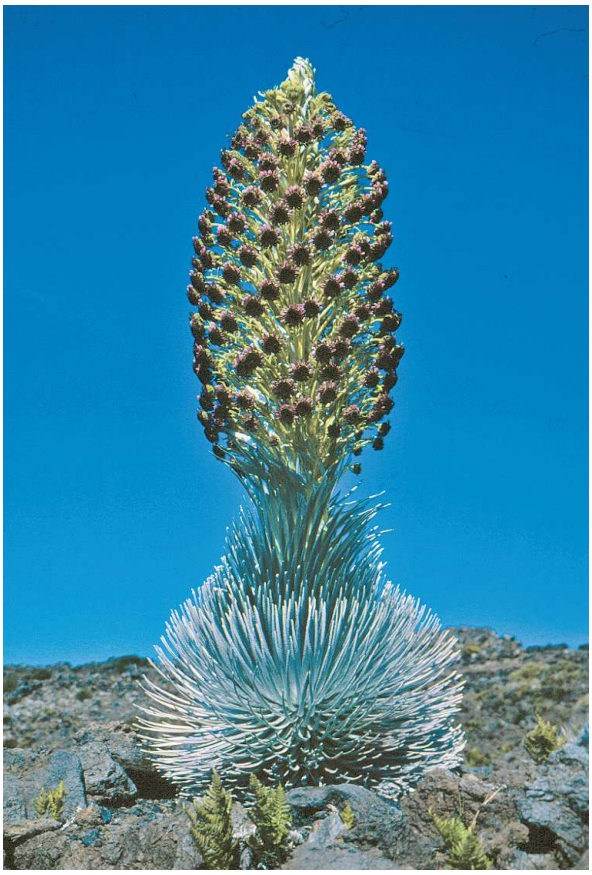
There are also situations in which matrix models or IPMs are not the best choice. For example, if the spatial arrangement of individuals proves to be important in population dynamics, these methods can be misleading. More sophisticated models can take into account the spatial location of individuals and interactions with their neighbors. There is, of course, a catch: to accurately estimate the parameters of such spatially explicit individual-based models, one needs substantially more data than for a matrix model. A series of models developed by Stephen Pacala and associates (Pacala and Silander 1985; Pacala 1986a, b; Pacala 1987) illustrates both the strengths and difficulties of this approach. On the one hand, the authors were able to successfully predict many features of the dynamics of their study population; on the other hand, large quantities of data were needed to estimate model parameters, sophisticated computer programming was necessary to study the models, and analyses of sensitivity and elasticity would be difficult to perform. Individual-based simulation models are another approach (Grimm and Railsback 2019). The idea is to incorporate much of the variability in populations by simulating the responses of each individual. A major difficulty with this approach is that these models also require large quantities of data to estimate the many parameters that are employed. Moreover, because they are simulation based, it is often difficult to derive general results.

### Plant populations are heterogeneous

Plants within populations vary, even after we have accounted for such things as age, size, or other categories used in typical matrix models or IPMs. Many factors can cause variation, but the two most obvious are genetic differences and differences among microhabitats. This variation can affect the demographic performance of individuals. Indeed, if there were no demographic variation within populations, evolution by natural selection could not occur (see Chapter 9). Another important contributor to heterogeneity is competitive hierarchies (see Chapter 10), which lead to some individuals having much greater chances of survival or reproductive output than others (Weiner 1985; Herrera 1991; Schwinning and Weiner 1998).

Several studies (Fang et al. 2006; Coutts et al. 2012; Haymes and Fox 2012) have found that individuals (or plants in particular locations) consistently outperform others in the same population, in terms of reproductive output, survival, or both. We expect that many more studies will find similar variation. But is this variation important for populations? Theoretical studies say that demographic heterogeneity can have important effects on population growth rates and extinction risks (Fox 2005; Kendall et al. 2011; Vindenes and Langangen 2015). The circumstances under which it affects these population-level outcomes are still matters of active research.

In some populations, the variation may be qualitative: individuals may actually have different life histories (see Chapter 7). For example, most individuals of the Mauna Kea silversword (*Argyroxiphium sandwicense*, Asteraceae; Figure 8.10) reproduce once and then die (they are semelparous; see Chapter 7), but a few individuals revert to vegetative growth after flowering or flower repeatedly (Powell 1992). A standard matrix model would show the probability of flowering repeatedly as an average across the entire population, although the reality is that most individuals can never do this. One approach to this problem is to ask how each of the different life history types contributes to the overall population growth rate—an advanced method called loop analysis (van Groenendael et al. 1994). For example, Glenda Wardle (1998) contrasted the importance of annual versus biennial life histories (see Chapter 7) in *Campanula americana* (tall bellflower, Campanulaceae) and showed that the biennials had a much larger effect on population growth. Chi-Hua Lin and collaborators (2016) showed that in *Dicentra canadensis* (Fumariaceae) clonal reproduction contributed



Courtesy of Robert H. Robichaux

**Figure 8.10** *Argyroxiphium sandwicense* (Mauna Kea silversword, Asteraceae) from Hawaii. These plants are mainly semelparous, reproducing only once in their lives; rare individuals, however, revert to vegetative growth after flowering, or they flower repeatedly.

much more to λ than sexual reproduction in forest remnants, but the two pathways contributed about equally in a well-forested site. Currently, loop analysis cannot be done with IPMs.

## 8.4 Demographic Studies of Long-Lived Plants Require Creative Methods

A small fraction of plant species reach very great ages. Some *Larrea tridentata* (creosote bush, Zygophyllaceae) genets in the Mojave Desert are estimated to be about 14,000 years old. Obviously it is difficult to study the demography of populations in which some individuals have been alive for nearly as long as humans have occupied North America! Fortunately, most plants do not present such extreme difficulties for demographic study. More typically, trees and shrubs live for a few decades to a few centuries. But even these life-spans are difficult to study when the average researcher is active for only a few decades.

One approach has been to develop matrix (or similar) models for long-lived plants based on a set of samples from the current population. In her study of *Trillium grandiflorum* (see Figure 8.4), for example, Tiffany Knight (2003, 2004) marked individuals in each class and then recensused them. Following an entire cohort would have been impossible, as individuals can live longer than 70 years. William Platt and associates (1988, 1993) have been studying populations of longleaf pine (*Pinus palustris*) since the 1970s. Year-to-year variation in environmental conditions—sometimes due to large disturbances such as hurricanes—has been a hallmark of these studies, and it appears to play a major role in the dynamics of these populations.

There are a few long-term studies of cohorts of marked trees. Robert Peet and Norman Christensen (1987) studied large plots of trees established in the early twentieth century in the Duke Forest at Duke University. Clearly, such studies can have some limitations. Measurements of survival are often taken at longer intervals than one might like—such as 10 years or multiples of 10 years. Seeds, seedlings, and small saplings are usually not measured. Because of the long intervals between censuses, one generally undersamples the variation between years. But even information obtained with such limited resolution is still very valuable.

This does not mean that one can never infer anything from examining a population's structure at one time point or over a short period of time. For example, if one finds that a population consists of widely separated age classes (as occurs in many forest trees and many cacti), it is reasonable to suggest that opportunities for successful recruitment to the population occur only rarely. If there is also a tendency for plants of a certain age class to be located near one another, it is reasonable to infer that the process creating recruitment opportunities is spatially patchy. For example, many forest trees occur in nearly even-aged stands because wildfires create open patches for recruitment (see Chapter 13). By constructing fire histories—inferred from examining the ages of fire scars on trees and from dating charcoal in the soil—investigators have gained insight on the frequency and severity of these disturbances. Although methods like these do not provide estimates for matrix elements, they produce strong inferences if done well.

Several researchers have gained insight into the demography of long-lived plant populations by finding creative ways to obtain data on survival retrospectively. For example, Deborah Goldberg and Raymond Turner (1986) first used maps of plots established by

(A) 1903

All photos courtesy of R. Turner

(B) 1961

(C) 1996

**Figure 8.11** Comparison of old and new photographs has provided important insights into the demography of long-lived plants as well as vegetation change in arid country. Pioneered by Rodney Hastings and Raymond Turner (1965), the technique requires landmarks for identifying the sites of old photographs. Here, a comparison of photographs shows changes in the population of *Pachycereus pringlei* (cardón, Cactaceae) on Isla Melisas, in Guaymas Bay, Sonora, Mexico. (A) This photograph, taken in 1903, shows a population of old, many-branched cardón. (B) In this photograph, taken in 1961, most of these old individuals have been replaced by much younger plants, although several large individuals can still be seen along the ridgeline. (C) In this photograph, taken in 1996, a dense stand of cardón can be seen; most of the plants are much larger than in the 1961 photograph. Three large, old individuals are still apparent along the ridgeline.

Forrest Shreve in the 1930s to document survival in several species of long-lived cacti and desert shrubs. Perhaps the most unusual studies have been those in which Rodney Hastings and Raymond Turner (Hastings and Turner 1965; Turner 1990; Turner et al. 2003) used landmarks to match old photographs (some dating to the nineteenth century) with new ones and identified surviving desert trees, shrubs, and large cacti (Figure 8.11).

## 8.5 Population Growth Fluctuates Randomly over Time

By now you might reasonably be wondering whether this approach to plant populations makes much sense. After all, we have assumed that the population experiences fixed survival and birth rates, and at least in a demographic sense, we have assumed that they live in a constant environment. Since one of the more obvious facts about ecology is that things change, and factors influencing populations (such as the weather) can be quite variable (see Chapter 16), the assumptions used for these analyses might seem to render any conclusions based on them invalid.

Matrices and life cycle graphs are nevertheless quite useful, but their utility depends on how one interprets the results. We can use these models for two very different purposes: to try to predict actual population growth and composition at some time in the future, or to ask what would happen to the population if present conditions persist. Only the latter purpose is usually strictly valid, but still, the information obtained can be helpful for purposes like comparing populations and noticing when things might not be going the way you had hoped (a threatened species is declining rapidly, an invasive population is exploding). Science often starts from the simplification of "all things being equal" and then elaborates from there.

It follows from this that using matrix models is helpful in studying the year-to-year variation experienced by plant populations. For example, in their study of the annual *Collinsia verna*, Kalisz and McPeek (1992) found that in one year the population increased substantially, with $\lambda$ estimated as 1.80, but in the next year it decreased, with $\lambda$ estimated as 0.41. Hurricanes cause large-scale variation in population growth rates of *Pinus palustris* (longleaf pine, Pinaceae), but the use of matrix models has shown that there is substantial variation in growth rates among even the "normal," nonhurricane years (Platt et al. 1988; Platt and Rathbun 1993).

Finally, concentrating on the dominant eigenvalue and eigenvector does not mean that one is studying things that would be important only in the long run, and only in a constant environment. As noted above (see the discussion of Equation 8.2), short-term population growth can be analyzed as a weighted sum of

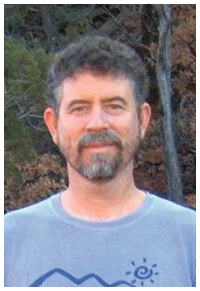
the eigenvalues and eigenvectors of a matrix. The dominant eigenvalues and eigenvectors are part of this sum, so in studying them, one is always studying a major component of short-term population growth (Caswell 2001). Many plant population matrices seem to approach the stable stage distribution fairly rapidly—often it is very close within 5 or 10 years (see Figure 8.8). This convergence can occur only if the dominant eigenvalue is much larger than the others, making it a very important component of short-term population growth.

Thus, one need not know the entire history of a population to understand the reasons for its stage structure. The stage structure reflects the survival and birth rates of the population. This still tends to be the case even when the matrices vary randomly from year to year—for example, due to environmental variation (Tuljapurkar 1990).

In the real world, population growth rates vary because of both **stochastic** (randomly varying) factors (like the occurrence of an unusual freeze when plants are flowering) and long-term trends (like climate change or land use change). Factors affecting population growth—such as temperature, moisture, herbivores, pathogens, competitors—vary from year to year. More dramatic events, such as hurricanes and fires, also have important effects on the growth rates of many populations. Fortunately, many of the tools described above can be modified for use in a variable world. First, however, it is important to be clear about how random variation affects population growth.

## There are two general types of random variation

There are two ways in which population growth can be affected by random variation. First, in the case of **environmental stochasticity**, vital rates can vary as a result of environmental factors that affect all the individuals in a stage class (or in a population) in roughly the same way. For example, desert environments have highly variable rainfall (see Figure 16.22). In a 33-year study of winter annuals (plants that germinate in autumn and flower in spring) in the Sonoran Desert, D. Lawrence Venable and colleagues (Venable and Pake 1999; Venable and Kimball 2013) documented major variation in the **realized fertility** (the chance of surviving to maturity times the fertility of survivors) of 15 species of these plants (Figure 8.12). Germination fraction also



Courtesy of D. Lawrence Venable

D. Lawrence Venable

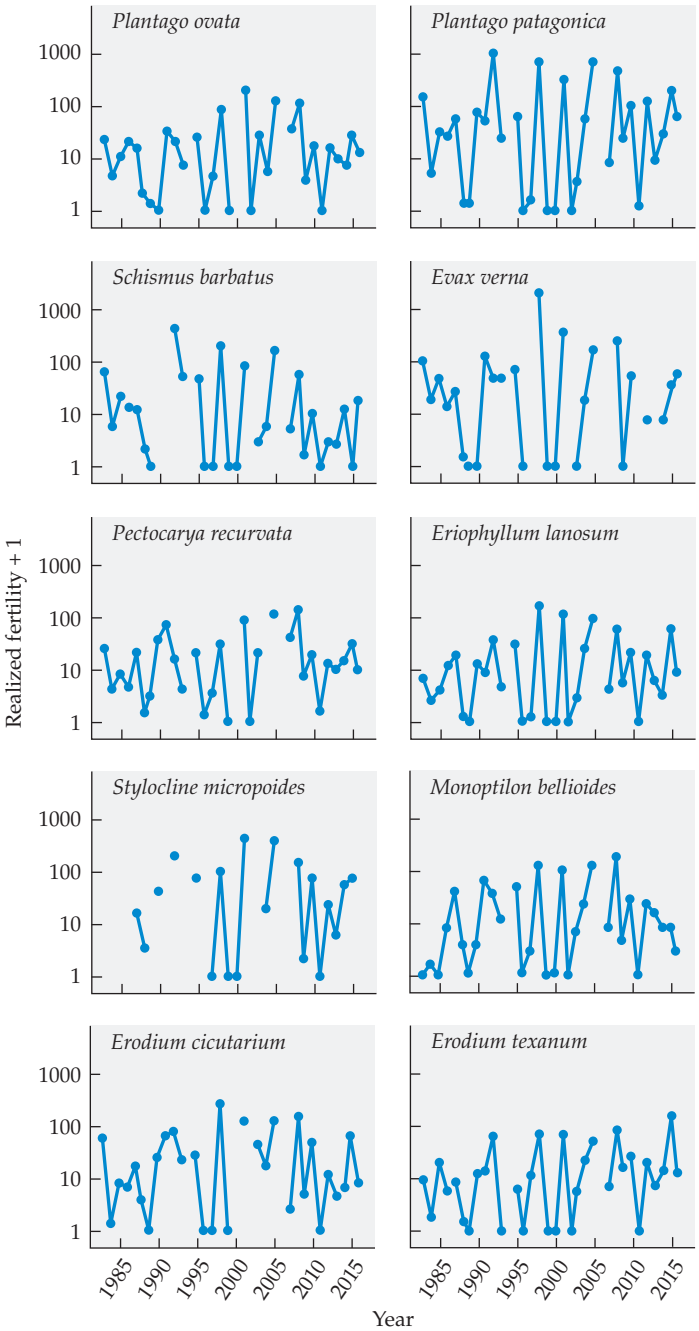



**Figure 8.12** The effect of environmental stochasticity on the realized fertility of Sonoran Desert winter annuals. Realized fertility is the chance of surviving from emergence to maturity times the mean fertility of survivors. Increased rainfall due to the El Niño Southern Oscillation occurred in 1983, 1987, 1991, 1992, 1998, and 2001—all years of high realized fertility for most species. All species except *Erodium cicutarium* and *Schismus barbatus* are native to the study area. (After D. L. Venable and S. Kimball. 2013. In *Temporal Dynamics and Ecological Process*, C. K. Kelly et al. [Eds.], pp. 140–164. Cambridge University Press: Cambridge; additional data courtesy of D. L. Venable.)

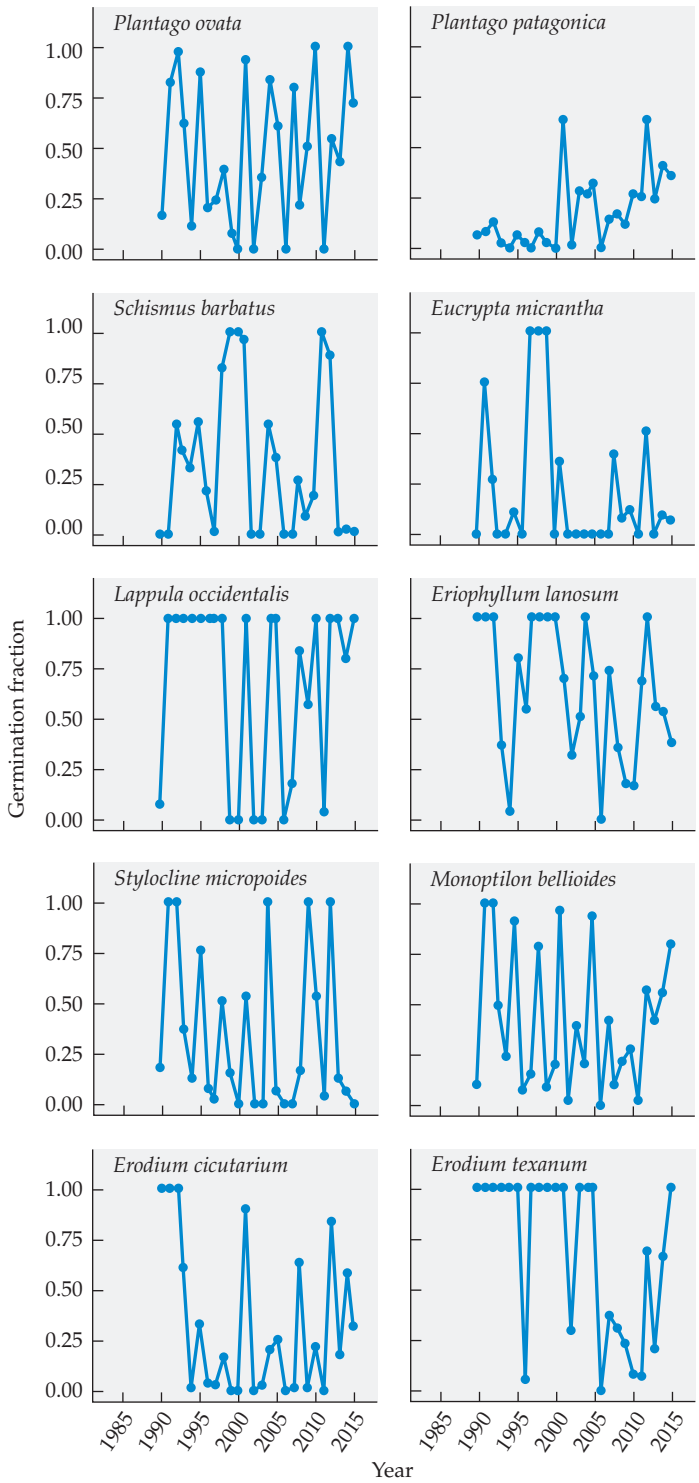



**Figure 8.13** Germination fractions for 10 species of Sonoran Desert winter annuals. The germination fraction is the density of seedlings emerging in natural field sites, divided by the total number of viable seeds in the soil. (After D. L. Venable and S. Kimball. 2013. In *Temporal Dynamics and Ecological Process*, C. K. Kelly et al. [Eds.], pp. 140–164. Cambridge University Press: Cambridge; additional data courtesy of D. L. Venable.)

varies widely between years (Figure 8.13). This variation in germination, survival, and fecundity have all been shown to drive fluctuations in the population sizes of these species (Huang et al. 2016; Figure 8.14). Years in which realized fertility was high appear to be associated with El Niño events, when rainfall in this desert is high (see Figure 16.19).

Second, in the case of **demographic stochasticity**, chance variation in the fates of individuals (while all have the same vital rates) reduces the average long-term growth rate of the population. In other words, this is a case of **sampling error** for demographic rates. Sampling error does not mean that the ecologists made a mistake in the way they sampled! Rather, it refers to the fact that when we only have a small sample, as in a small population, we typically get answers that differ somewhat from what we would see if we could measure a large number of individuals. For example, imagine that a plant's chance of living until next year is $p = 0.25$. But the reality is that this particular plant will either live (so a crude estimate would be $p = 1$) or die (and a crude estimate would be $p = 0$). There is no way that with a sample of one plant, we could ever estimate the true probability of surviving. A little thought about arithmetic should suggest to you that only if we had at least four plants (all with $p = 0.25$) could we possibly estimate the true survival probability. But we know that with four, having all of them die is not unlikely; just as with coin tosses, it is only when the number gets to be fairly large that departures from the number we expect become very small. That is a general characteristic of sampling. Demographic stochasticity might also result in a small population with many juveniles and few reproductive adults; this could also depress population growth for some time. This sampling process is similar to that which causes genetic drift (see Figure 9.13).

Some important features of demographic stochasticity are shown in Figure 8.15. Figure 8.15A shows the observed number of survivors in a hypothetical population, given a probability of 0.25 of surviving. You might notice that there are more extremely low values than extremely high values. There is a reason for this: even in a population of 1000, it is possible that, by chance, no plants will survive, but it is never possible for more than 1000 to survive. Figure 8.15B shows the relative departure of the simulations from the expected number of survivors. The relative departure tends to be greater in small populations.

Both forms of stochasticity occur in all populations. Environmental stochasticity can have substantial effects in populations of any size. As the examples above suggest, demographic stochasticity is most important in small populations.

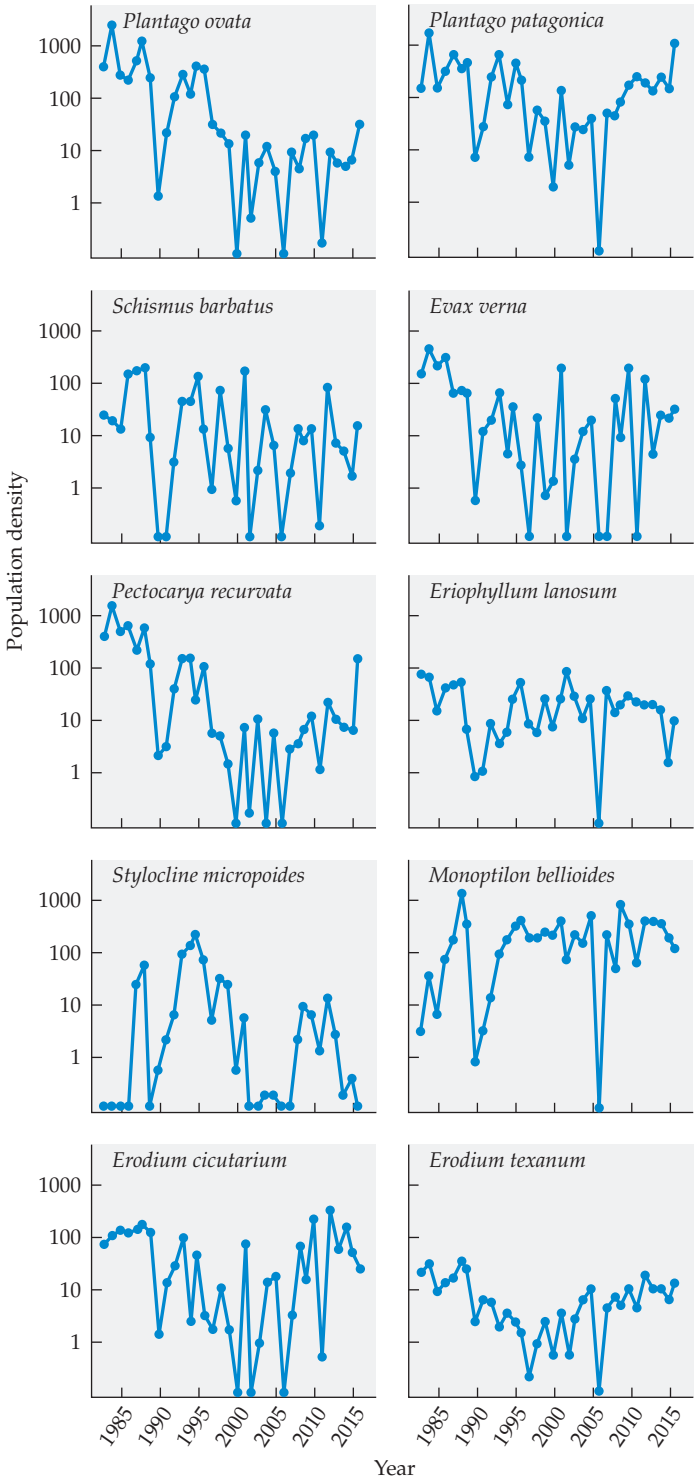



**Figure 8.14** Fluctuations in population density in 10 species of Sonoran Desert winter annuals. (After D. L. Venable and S. Kimball. 2013. In *Temporal Dynamics and Ecological Process*, C. K. Kelly et al. [Eds.], pp. 140–164. Cambridge University Press: Cambridge; additional data courtesy of D. L. Venable.)

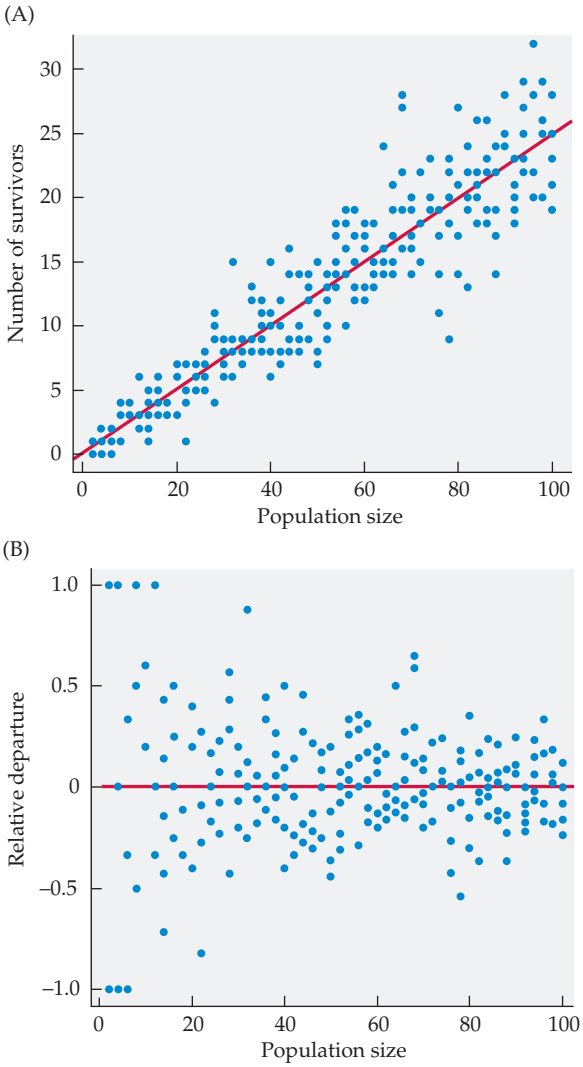



**Figure 8.15** Demographic stochasticity in survival. (A) Each point shows the observed number of survivors in a single sample from a hypothetical population, given a probability of surviving of 0.25. The line represents the expected number of survivors: 0.25 times the population size. (B) The relative size, for the same samples, of the departure from the expected value: (*N* survivors – expected)/expected.

## Random fluctuations reduce long-term growth rates

The most important effect of both of these types of stochasticity is to reduce the average long-term growth rate of a population, compared with a population with the same average growth but no fluctuations. Imagine a population of annual plants (without a seed bank, so it is an unstructured population) that in good years grows at the rate $\lambda_{good} = 1.1$ and in bad years grows at the rate

$\lambda_{bad} = 0.9$. To keep things simple, imagine that good and bad years alternate. It may surprise you to realize that this population is slowly going extinct, even though the average $\lambda$ is 1.0 (Figure 8.16)! How can this be? It is because the long-term growth rate is not an ordinary arithmetic average, but the **geometric mean** of the annual growth rates. The geometric mean of a group of $n$ numbers is based on multiplying all of those numbers, rather than adding them together; it can be calculated as either the $n$th root of their product or the exponential of the average of their logarithms. This makes sense, because we are considering how the population size is multiplied over time, which is what the growth rate is about. It is mathematically guaranteed that if the numbers vary, their geometric mean will be less than their arithmetic mean—so in the case where the average value of annual growth rates is 1 but the growth rates fluctuate, population extinction is (in the long run) assured! One way to see this is to consider what happens if in one year $\lambda = 0$ so no one survives. It does not matter what any of the other values of $\lambda$ are; multiplying them by zero will always give you zero. Thus environmental stochasticity tends to reduce the long-term population growth rate and increase the risk of extinction, as compared with (the unlikely case of) a population having the same average annual growth rate but not varying from year to year. If the fluctuations among years occur randomly, as in Figure 8.17, in the long run the average $\lambda$ also is reduced from what one would expect without fluctuations, and so is the population size.

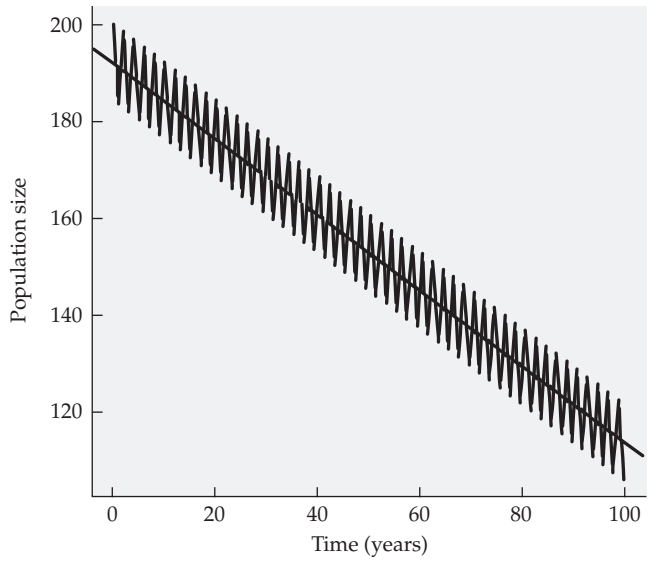



**Figure 8.16** Simulation of population size (on a logarithmic scale) for a population that in alternate years has $\lambda = 1.1$ and $\lambda = 0.9$. The average growth rate is 1, but the population declines in the long run, at a rate depending on the geometric mean: The slope of the line is $-(1 - GM)$.

Both the average conditions and their variation are important in determining long-term rates of population growth. There are few studies on this kind of long-term variation. There are many anecdotal accounts of plants in highly variable environments (such as deserts). They usually have high mortality and low fertility, with an occasional "good year" producing vast quantities of seed. It is possible for a plant population to have an increasing

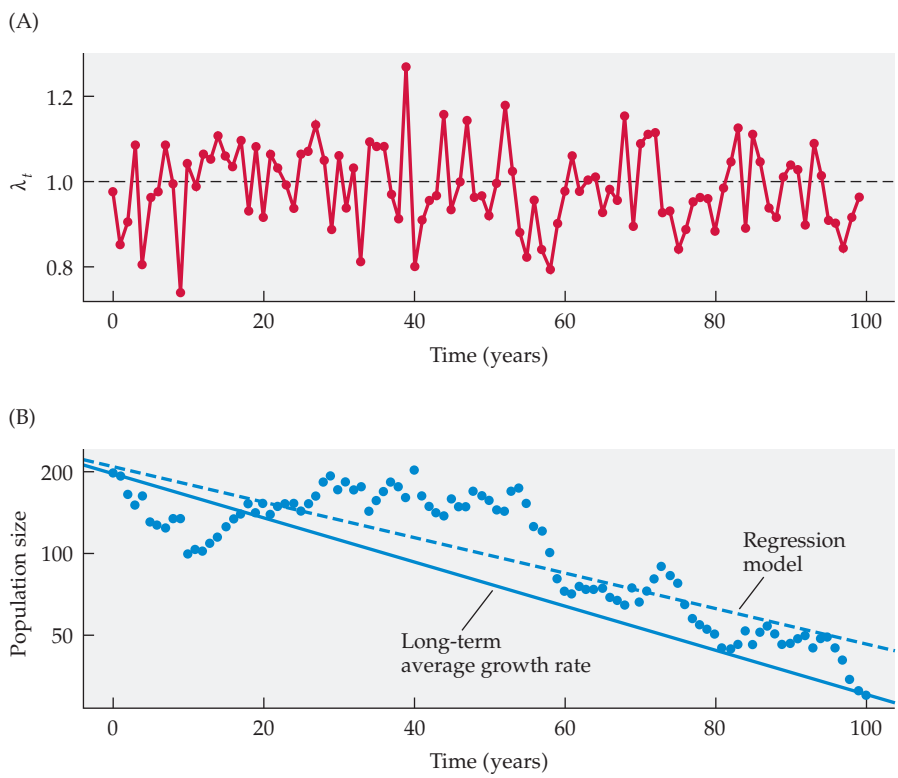



**Figure 8.17** Environmental stochasticity, with $\lambda$ having a mean of 1 and a standard deviation of 0.1, as in Figure 8.16, but with the variation occurring randomly. (A) Values of $\lambda$ over 100 years. (B) Resulting population size. The solid line is for the long-term average growth rate (as in Figure 8.16); the dashed line is from a regression model fitting this particular sequence of 100 years.

size in the long run, even while decreasing in numbers during most years. This sort of year-to-year variation is thought to play an important role in the coexistence of species within communities (see Chapter 10). Both the mean and the variance of the growth rate are also important in the evolution of plant life histories; in Chapter 7 we discuss how some plants can "trade off" the mean and variance of their growth rates.

Calculating the long-term stochastic growth rate for age- or stage-structured populations is more complicated, but it is still true that among-year fluctuations will increase the probability of extinction in such populations (see Caswell 2001 for more details).

## Studying variable population growth requires data recorded over many years

To study the effects of environmental stochasticity, one needs the same kinds of data used to estimate average population growth rates, recorded over enough years to estimate the variances and covariances of matrix elements. Given such data, one can perform computer simulations of population growth that is subject to random variation. There are many studies reporting variation in demographic rates (e.g., Bierzychudek 1982; Horvitz and Schemske 1990), but few are long enough to provide good estimates of variance, and even fewer can provide estimates of correlations among parameters in matrices or IPMs.

A population's **extinction probability** is the fraction of replicate populations that can be expected to go extinct within a stated period of time. This is estimated by computer simulation. Extinction probabilities for several California species of *Calochortus* (Mariposa lily, Liliaceae) are shown in Figure 8.18. Extinction probabilities increase as the amount of environmental stochasticity increases, and this increase is fastest for populations with the smallest value of λ.

A number of demographic models have examined population growth in the context of disturbances such

(A)

Courtesy of G. Fox

(B)

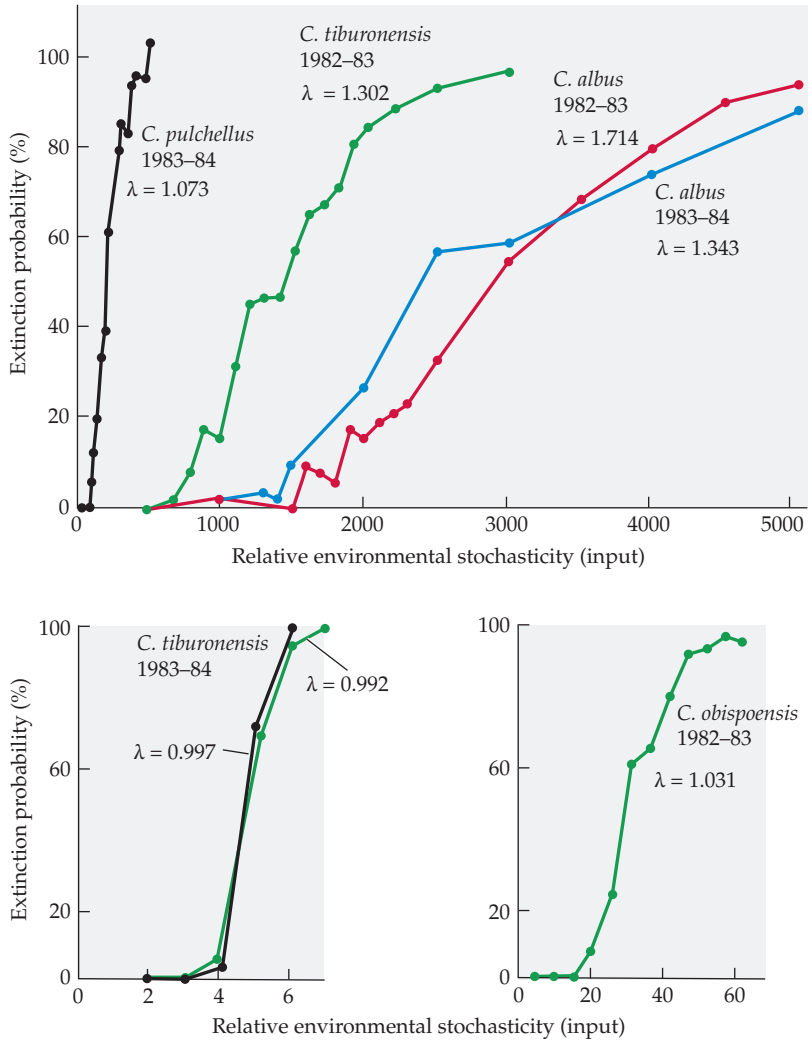



**Figure 8.18** (A) *Calochortus gunnisonii* (mariposa lily, Liliaceae) occurs throughout the Rocky Mountains and adjoining regions of the United States. (B) Extinction probabilities of several California populations of *Calochortus* spp. Environmental stochasticity was modeled by taking 1% of the variance/mean for seed production, and 0.01% of the variance/mean for other vital rates, and then multiplying these values by the numbers shown on the horizontal axis. (B after E. S. Menges. 1992. In *Conservation Biology: The Theory and Practice of Nature Conservation, Preservation, and Management*, P. L. Fiedler and S. K. Jain [Eds.], pp. 253–275. Chapman and Hall: New York.)

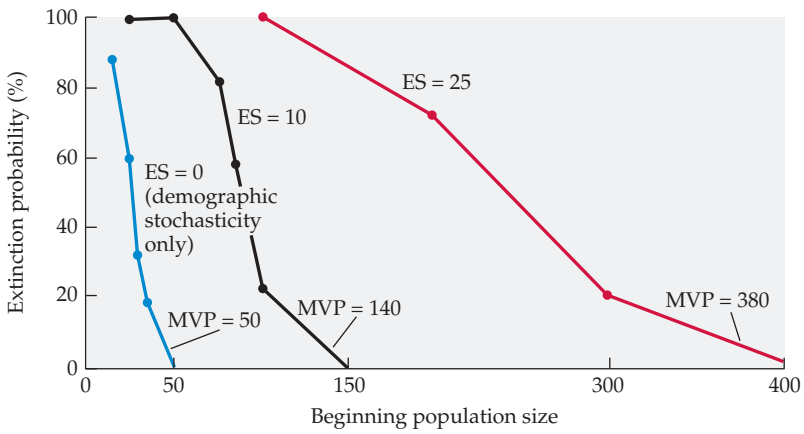



**Figure 8.19** Results of a simulation conducted to estimate the minimum viable population (MVP) size for a population of the Mexican palm *Astrocaryum mexicanum* (chocho, Arecaceae). Environmental stochasticity (ES) was modeled using the methods described in Figure 8.18. MVP is the smallest number giving an extinction probability of 5% or less in 100 years. (After E. S. Menges. 1992. In *Conservation Biology: The Theory and Practice of Nature Conservation, Preservation, and Management*, P. L. Fiedler and S. K. Jain [Eds.], pp. 253–275. Chapman and Hall: New York.)

as fire. For example, Juan Silva and coworkers (1990, 1991) studied *Andropogon semiberbis* (Poaceae), a Venezuelan grass from savannas that burn almost every year. Their analyses suggested that if fire occurs in fewer than about 85% of years, the population will go extinct. Eric Menges and Pedro Quintana-Ascencio (2004) used a similar approach to find a minimum fire frequency for maintenance of *Eryngium cuneifolium* (wedgeleaf eryngium, Apiaceae) in Florida rosemary scrub.

An idea that is closely related to extinction probability is the **minimum viable population** (**MVP**). The MVP is the minimum size necessary to give a population a probability $x$ of surviving a specified number of years. Typically, MVP is calculated using a probability of 95% and a time of 50 or 100 years. Eric Menges (1992) estimated the MVP for the Mexican palm *Astrocaryum mexicanum* (chocho, Arecaceae) using published matrices from the extensive study by Daniel Piñero and colleagues (1984, Figure 8.19). Populations subject only to demographic stochasticity could begin with only 50 plants and still have a 95% chance of persisting 100 years. Populations subject only to environmental stochasticity required larger initial sizes to have this chance of persisting. As environmental stochasticity became larger (that is, there was more year-to-year variation), as the population size became smaller, the chance of a population going extinct rapidly increased. An important problem is that political entities sometimes see the MVP as the goal, rather than treating it as a minimum; as a result, scientists constructing species management plans use it much less frequently than in the past, and we do not encourage its use.

Most demographic studies of plants have lasted only a few years and consequently have focused on average values of growth rates . These are useful quantities, but it is important to realize that  overestimates the stochastic growth rate (just as the arithmetic mean is always larger than the geometric mean, unless the variance is zero).

## 8.6 Demographic Models Have Strengths and Limitations

It would be nice if we could generalize the results of our demographic studies. For example, it would be great if we could model the demography of a population we are studying and apply those results to another population. Unfortunately, life is not that simple. From 1999 to 2003 Eelke Jongejans and Hans de Kroon (2005) studied three perennial herbs where they all co-occurred in five sites in the Netherlands. The populations responded differently to differences between sites and between years, and while the species had similar life histories, they had different patterns of response to environmental variation. Jongejans and de Kroon concluded that generally it is not valid to do "space for time" substitutions in demography—we cannot sample multiple sites and conclude that the variation among them is roughly the same as the variation among years in a single site, or vice versa. Nor can we use data from one species as a proxy for data in a similar or related species. Recent work has supported both conclusions. The range of variation can be astounding. Nathan Emery and his collaborators (2017) used a **common garden** to grow plants from 17 populations of a single species, *Actinotus helianthi* (flannelflower, Apiaceae), in different parts of its range in New South Wales, Australia. They found that germination success varied between 0.2% and 64.2%; when seeds were treated with smoke, germination varied between 59.3% and 97.5%. The percentage that survived to the conclusion of the experiment (day 153) varied from less than 1% to over

35%, and the percentage of individuals that produced seeds varied between 62% and 90%. A recent analysis by Judy Che-Castaldo and her collaborators (2018) of matrix models from 425 plant species in the COMPADRE database revealed that demographic rates are not well predicted by phylogenetic relatedness or by plant traits and confirmed that demographic rates vary considerably among populations within the same species.

These studies suggest that the details of demographic models are not generalizable, although the general approach and fundamental concepts apply very broadly. Constructed properly, demographic models can provide an invaluable set of analyses of observed demographic processes. But models for particular populations cannot tell us what other populations of the same or different species are likely to do. Indeed, Elizabeth Crone and her collaborators (2013) constructed matrix models for 82 populations of 20 species and asked how well the models described population dynamics during the time used to estimate the model parameters, and how well they predicted dynamics more than 5 years beyond that time. The result was that the models did quite well at describing the observed dynamics, but predicted population sizes were within the 95% confidence interval 5 or more years later for only 40% of the models. The poor forecasting ability of those models was a result of significant environmental variation, not of unmodeled density dependence or poor sample sizes. The researchers concluded that good forecasts may require models that include likely kinds of environmental changes as well as the demographic responses to those changes. Alternatively, some recent studies have found that forecasting quantities like plant cover or total population size may be done equally well using other statistical approaches (Tredennick et al. 2017; Greenville et al. 2018).

Courtesy of Elizabeth Crone

Elizabeth Crone

## Summary

- Because of their modular growth, individuals in plant populations can be defined based on whether they are genetically distinct (genets) or potentially physiologically distinct (ramets). The choice can have important effects on the questions one can ask.
- Survival and reproduction in plants usually depend more on a plant's size or developmental stage than on its age. When it can be determined, age can still be informative.
- Populations increase and decrease in size, and they also change in the numbers of individuals of a given age, size, or developmental stage—their structure.
- Studies of change in population size and composition depend on population models. The most widely used demographic models consider survival, growth, and reproduction as functions of the individual plant's size or developmental state.
- Demographic models can provide an estimate of the anticipated long-term rate of growth of the population at the time the estimate was made. That rate is a description of how the population is currently changing, not a prediction of how it will change in the long run, because future conditions are unknown.
- Ecologists get important information by asking how the long-term growth rate depends on particular components of growth (like the survival of seedlings).
- Plant populations vary in their growth rates due to environment fluctuations, heteogeneity among individuals, and sampling error. This random variation means that in the long term, populations grow more slowly than they would in a constant environment with the same average growth rate.

Courtesy of Sam Scheiner

9

# Evolution: Processes and Change

In other chapters we explore the various ways in which plants interact with their biotic and abiotic environment and consider the features of plants that permit them to cope with those different environments. Many of those features arose through the process of evolution by natural selection. Through this process, just a few simple consequences of biology and natural laws have resulted in the vast diversity of life. All of the species that make up ecological communities arose through evolution, and ecological processes provide the context for that evolution. The ecologist G. Evelyn Hutchinson emphasized this intimate relationship in his seminal book *The Ecological Theater and the Evolutionary Play* (1965). The intertwining of evolution and ecology is greatest in the process of natural selection. In this chapter we summarize the basic principles and processes of evolution and explore some of its outcomes.

How do we know that a particular plant trait was shaped by natural selection? Answering this question is not simple. There is no single method for doing so, and all methods require assembling multiple kinds of information. Historically, ecologists have sometimes been guilty of assuming that characteristics of organisms were optimal, and were always adaptations produced by natural selection, without carefully documenting whether that was, in fact, the case. These assumptions were termed "Panglossian" by Stephen Jay Gould and Richard Lewontin (1979), after the character Dr. Pangloss in the novel *Candide* by Voltaire. In that story, the naive Dr. Pangloss goes through life confidently declaring his assumption, "All is for the best in this, the best of all possible worlds."

Sometimes ecologists have constructed scenarios to explain what processes may have led to particular adaptations, and then simply accepted those scenarios without substantiation. For example, some plant species produce seeds with a sticky, mucilaginous coating. Some ecologists originally claimed that this coating was an adaptation for dispersal; seeds getting stuck on the feet of ducks was

Above: *Callistemon*, commonly known as "Howie's Fire Glow." Canberra, Australia

one particular scenario they proposed. Such scenarios are termed "just-so stories," after the title of the book in which Rudyard Kipling recounts fanciful tales about the origins of various animals. Today we understand that in many species, the coating has more to do with water retention by the seed than with dispersal.

The problem with just-so stories constructed by scientists is not that they are necessarily wrong; it is that they are based on scant evidence or on unfounded extrapolation from what is known. For example, while we cannot justifiably make the leap to sticky seeds having evolved by the process of natural selection acting to increase the dispersal success of all seeds, it has been documented with substantial evidence based on careful study that the sticky seeds of some species, such as *Phoradendron californicum* (desert mistletoe, Santalaceae), are in fact spread by birds (see Figure 6.15D).

Therefore, we must be cautious in either accepting or dismissing such adaptive speculations. Speculations are an indispensable first step in posing hypotheses to test whether a trait has come about by natural selection. As ecologists are usually knowledgeable about the organisms and systems they study, such speculations often prove eventually to be correct. But until evidence has been assembled that firmly supports these suppositions, they remain speculations. The examples of the results of natural selection presented in this chapter were chosen because they are based on more than just speculation. Most of them are based on multiple lines of evidence. As we examine each, we discuss the reasoning behind the conclusions concerning the shaping of the trait by natural selection.

Why is it so difficult to decide whether a trait was shaped by natural selection? While natural selection is a primary process in shaping plant form and function, other processes may also be responsible for shaping particular traits. Mammals have backbones not only because a backbone is a handy thing to have, but because they are vertebrates, and they have inherited this trait from vertebrate ancestors; all vertebrates share this trait for that reason. Likewise, angiosperm cells exhibit cellular respiration not because it is especially advantageous to angiosperms per se, but because the ancestors of angiosperms possessed this trait, which they have inherited and thus share with a much broader group of living organisms. These traits might (or might not) have been adaptations of the ancestors in which the traits originated, but in their modern descendants they are simply inherited traits.

Natural selection is only one of the factors that cause evolutionary change. Other processes may act in concert with natural selection, or they may act against it, or they may act independently of it. Mutation, for example, is a necessary source of the variation that natural selection requires. Migration of individuals from other environments can prevent adaptation to local conditions in response to natural selection. Genetic drift can also cause changes in gene frequencies. Nonrandom mating (see Chapter 6) causes changes in genotype frequency, and in turn those can affect the rates at which gene frequencies change. Each of these four processes can produce results that might at first be assumed to be the outcome of natural selection. In studying the process of evolutionary change, we are generally reconstructing a historical event, so our conclusions are almost always based on indirect inference, rather than on direct observation. All of these factors make the determination of adaptation by natural selection a challenging task.

## 9.1 Natural Selection Is a Primary Cause of Evolutionary Change

**Natural selection** is the process by which individuals with differences in their traits leave different numbers of descendants because of that variation. If those differences are **heritable** (have a genetic basis), then evolution occurs by natural selection. Thus, natural selection alone is not sufficient evolutionary change. **Adaptive traits** are ones that have evolved by natural selection. The suite of traits associated with CAM photosynthesis (see Chapter 2), for example, is adaptive in hot, dry climates because individuals that have those traits are able to leave more descendants than individuals that do not. But not all hot, dry climates contain CAM plants, and CAM can be adaptive in other environments. Arctic plants are short statured because by remaining close to the ground, they remain warmer, grow more, and ultimately leave more offspring than individuals that do not have those growth characteristics.

The principles of natural selection were first proposed as an explanation for evolutionary change by Charles Darwin in his book *On the Origin of Species by Means of Natural Selection* (1859); Alfred Russel Wallace (1858) also put forward the same theory at that time. Natural selection is one of the five central processes of evolution. The others are mutation, migration, genetic drift, and nonrandom mating.

One thing to keep in mind as you read about evolution by natural selection is that, while many people assume that it is something that occurred long ago, evolution by natural selection is going on all the time. It occurred long ago, it happened in the recent past, and it continues to operate on populations today. Scientists study patterns of, and find evidence for, evolution by natural selection at all of these time scales. And ongoing natural selection can affect ecological processes, such as population dynamics (see Chapter 8), as part of a feedback between the two sets of processes.

## Variation in phenotype is necessary for natural selection

An important starting point for any discussion of evolution and natural selection is variation. Variation is ubiquitous in nature. Nearly all natural phenomena vary at some level. This principle is expressed in the adage "No two snowflakes are exactly alike." Living beings are much more complex than snowflakes, with an even greater potential for different kinds and amounts of variation. Variation among individuals is essential for natural selection to result in evolutionary change.

Plants, like most living organisms, vary phenotypically. The term **phenotype** refers to all of the physical attributes of an organism. Phenotypic traits include outward appearance (such as height, leaf size and shape, flower color, or fruit number), life history characteristics (such as being an annual), macroscopic and microscopic anatomy (from tissues to organelles), physiology, and biochemistry (such as protein composition). Phenotypic variation can be extensive for traits such as seed size, which can vary 20-fold within a single population of a wildflower species (Figure 9.1), or biomass, which can vary over several orders of magnitude in trees. Other kinds of traits, such as flower size and solute concentrations in cells, tend to vary much less among individuals in a population. Still other traits, such as petal number or photosynthetic pathway, tend to be invariant within a species, although even these traits vary within some species (see Chapter 2). Furthermore, the particular pattern of growth and development exhibited by an individual plant depends on a complex interaction between its genes and its environment. All of these differences among individuals result in phenotypic variation.

Phenotypic differences among individuals can result from three types of variation: genetic, environmental, and developmental. The **genotype** of an individual is the information contained in its **genome**—the sequence of its DNA. This information is expressed through the processes of transcription of DNA to make RNA, and translation of the RNA into protein. Gene expression, and the development of the plant as a whole, is controlled by the information contained within the genome through various feedback mechanisms, including interactions with the external environment. A seed contains all of the information necessary for the growth and development of the adult plant. Development can be thought of as the unfolding of the information contained in the genome. Genetic variation—differences among the genomes of individuals—can result in phenotypic variation among individuals.

Individuals also experience different environments as they develop and grow, so the information expressed by the genome is dependent on the environmental context during development. The environment can vary in many ways, even over very small distances. Each grass shoot in a seemingly uniform meadow, for example, experiences somewhat different amounts of light, soil nutrients, and herbivory. The same genetic individual may have a very different phenotype when grown under different environmental conditions. In addition, individuals with different genotypes may respond differently to the same environmental conditions. This phenotypic variation in response to the environment is called **phenotypic plasticity** (Bradshaw 1965). Phenotypic plasticity may be reversible or not, and can occur at different stages of a plant's life. In plant ecology, the terms *acclimation* and *acclimatization* are also used to describe some aspects of plasticity. Even two individuals with identical genotypes growing in identical environments, however, do not necessarily look alike or function identically. Small, random differences in when and how genes are expressed, called **errors of development**, can lead to measurable differences in the adult plant.

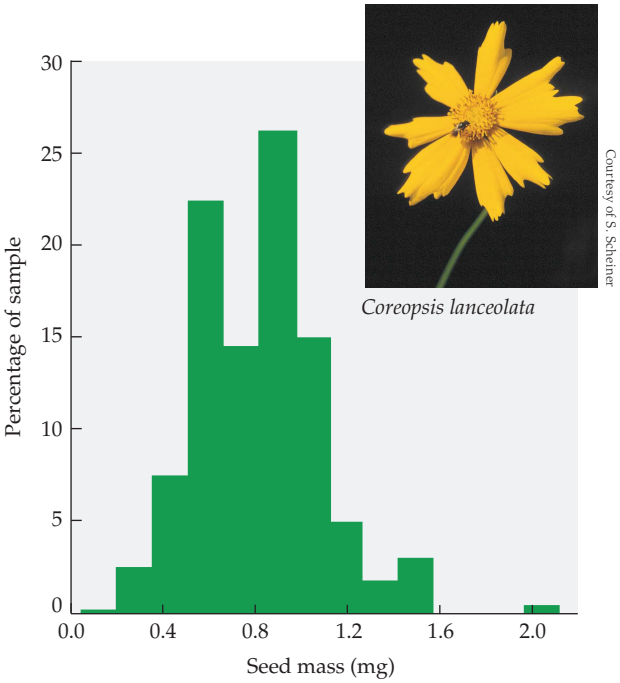



**Figure 9.1** Variation in individual seed mass in a population of *Coreopsis lanceolata* (Asteraceae) growing on an inland dune south of Lake Michigan. (Data from S. J. Banovetz and S. M. Scheiner. 1994. *Am Midl Nat* 131: 65–74.)

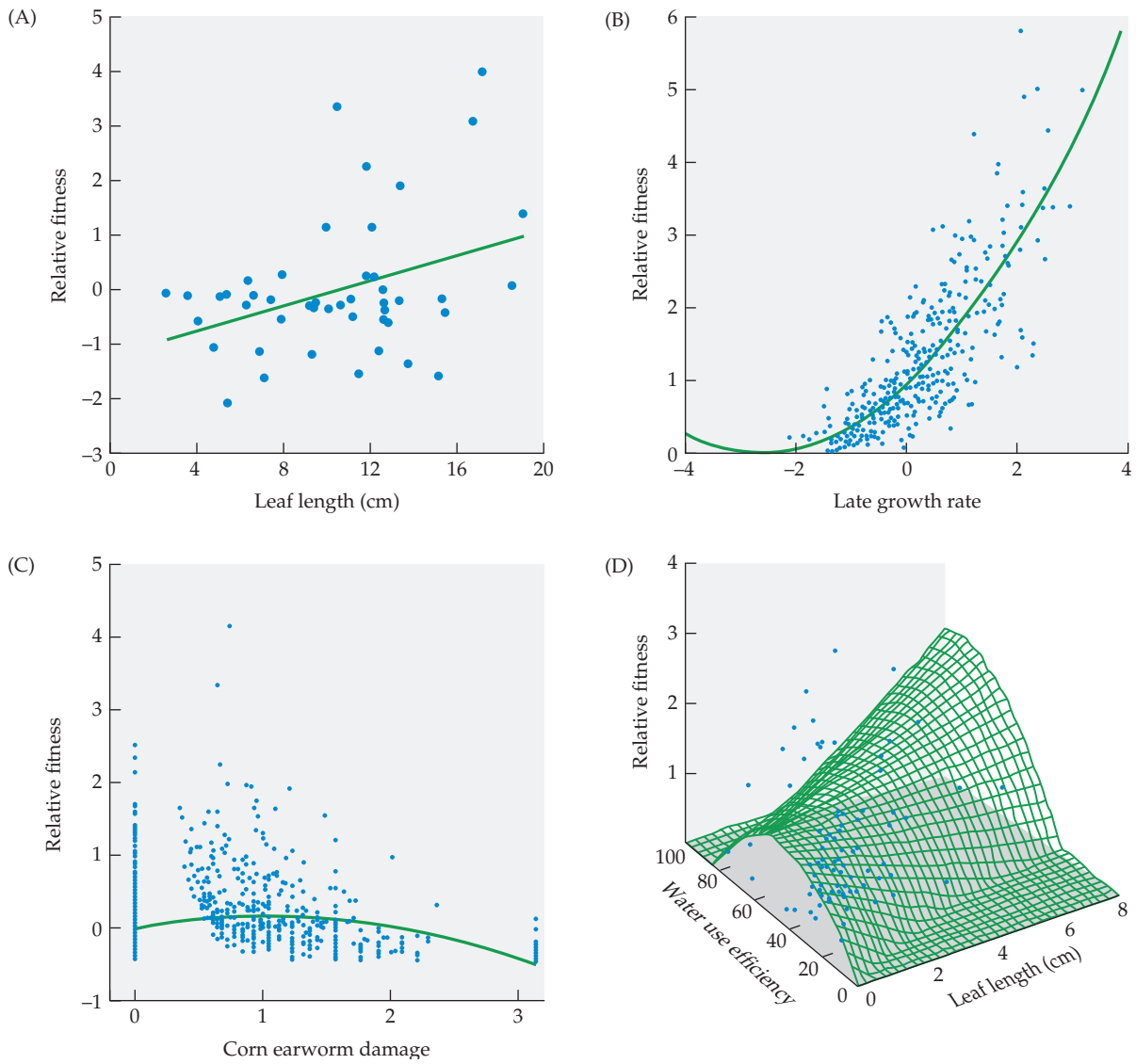
## Three conditions are necessary for evolution by natural selection

Natural selection requires three conditions in order to lead to evolutionary change: phenotypic variation among individuals in some trait, fitness differences related to those phenotypes (some individuals must leave more descendants than others as a result of those phenotypic differences), and **heritability** of the phenotypic variation (the phenotypic differences must have a genetic basis). If these three conditions are met for a trait within a population, then the frequency of that trait will change in that population from one generation to the next; in other words, evolution by natural selection will occur. Ecology plays a role in all three conditions.

Fitness differences (factors that lead to differences in the numbers of descendants) include differences in mating ability, **fecundity** (number of gametes produced), fertilizing ability, **fertility** (number of offspring produced), and **survivorship** (chance of surviving). In Chapter 8, we examined the consequences of these factors for the numerical dynamics of populations; in this chapter, we look at their evolutionary consequences.

What causes fitness differences? Differences among individuals in fitness may be due to chance and therefore may not be associated with trait differences in a consistent fashion. But often fitness differences are associated with differences in some trait among individual plants. **Phenotypic selection** occurs when individuals with different trait values have consistent differences in fitness. Phenotypic selection is the part of the evolutionary process most studied by ecologists.

A useful and convenient way to study the process of evolution by natural selection is to subdivide it into two parts: that which occurs within a single



generation—phenotypic selection—and that which occurs from one generation to the next—the genetic response (Endler 1986). Phenotypic selection consists of the combination of phenotypic variation and fitness differences. This part is what we often think of as natural selection. But for evolution to occur, the other part—the genetic response—is equally important. The **genetic response** is the change in the genetic makeup of the population that occurs from one generation to the next. The genetic response depends on the heritability of a trait. If a trait is heritable, then if phenotypic selection favors that trait in one generation, the next generation will have a greater proportion of individuals with that trait.

In a population of the grass *Danthonia spicata*, for example, individuals with longer leaves had higher survivorship and fecundity than individuals with shorter leaves (Figure 9.2A). Variation in leaf length is partially under genetic control. So the next generation in this population, on average, would be likely to have longer leaves. From one generation to the next, the change in the population may be small. But if the process of natural selection goes on for many generations, the population may become very different from the ancestral population. Much of the striking variation among species is due to such long-term evolutionary responses to natural selection.

We can characterize natural selection by the relationship between trait values and fitness. **Directional selection** occurs when individuals with the most extreme values (such as the smallest flowers, longest roots, or greatest stomatal conductance) for a trait have the highest fitness (Figure 9.2A, B). In this case, the population will continue to evolve in a single direction over time, if no other processes interfere. The primary change will be in the mean value of the trait in the population, as more and more individuals have extreme values over the course of many generations. **Stabilizing selection** occurs when individuals with intermediate trait values have the highest fitness (Figure 9.2C). Under stabilizing selection, the mean value of the trait will not change, but the variability in that trait will decline as fewer and fewer individuals have high or low values for the trait over the course of many generations. **Disruptive selection** has the opposite effect: it occurs when individuals with both high and low values of the trait have higher fitness than those with intermediate values. **Correlational selection** occurs when the pattern of selection on one trait depends on the value of another trait (Figure 9.2D). Correlational selection can result in highly coordinated suites of traits and complex adaptations.

◀ **Figure 9.2** Relationships between phenotypic variation and fitness differences result in phenotypic selection. If genetic variation for these traits also exists, then evolutionary change may occur. (A) Directional selection for greater leaf length in the perennial *Danthonia spicata* (poverty grass, Poaceae) in a natural population growing in a white pine–red oak forest in northern lower Michigan. Fitness was measured as the total number of spikelets produced over 5 years and was corrected for correlations with other traits. (B) Directional selection for a greater growth rate late in the season in the annual *Impatiens capensis* (jewelweed, Balsaminaceae) in a natural population growing in a deciduous forest in Wisconsin. The final dry weight of the plant, which is correlated with the number of seeds produced, was used as a proxy for fitness. The fitness function is curved, but still monotonically increasing. (C) Stabilizing selection for the amount of damage caused by the corn earworm to the annual *Ipomoea purpurea* (morning glory, Convolvulaceae) in an experimental population growing in an old field in North Carolina. Fitness was measured as the number of seeds produced by the end of the growing season. (D) Stabilizing selection for water use efficiency and directional selection for larger leaf size in the annual *Cakile endentula* (sea rocket, Brassicaceae) in an experimental population growing on a beach along southern Lake Michigan. Fitness was measured as the number of fruits produced by the end of the growing season. The optimal water use efficiency depends on leaf size, which is an example of correlational selection. (Part A data from S. M. Scheiner. 1989. *Evolution* 43: 548–562; B after T. Mitchell-Olds and J. Bergelson. 1990. *Genetics* 124: 407–415; C data from E. L. Simms. 1990. *Evolution* 44: 1177–1188; D after S. A. Dudley. 1996. *Evolution* 50: 92–102.)

## 9.2 Heritability Measures the Genetic Basis of Phenotypic Variation

Although at times it may seem as if the world consists of an infinite variety of species, that is not the case. It is easy to imagine organisms that could exist but do not, such as unicorns or dragons. Why not? One part of the answer is that the genetic basis of traits constrains evolution. For evolution to occur, there must be appropriate genetic variation. What limits this variation? To answer that question, we must first understand what we mean by genetic variation in an evolutionary context. Then we can explore the ways in which the environment interacts with genes to determine that variation.

### Heritability is a measure of resemblances among relatives

Heritability ($h^2$) is the amount of resemblance among relatives that is due to shared genes. Offspring tend to resemble their parents and their siblings because the phenotype of an individual is determined, in part, by its genotype, and an individual receives its genes from its parents and shares those genes with its siblings. Consider a trait such as height in an annual plant at the end of the growing season. We might do the following

experiment with *Brassica campestris* (Brassicaceae), a self-incompatible plant (see Chapter 6): We choose pairs of plants, pollinate one individual with pollen from the other in each pair, and cover the flowers to prevent any other plant from pollinating that individual. Thus, when we collect the seeds at the end of the growing season, we know exactly who the parents were. Before the parental plants die, we measure their heights. Then we plant the seeds, let them grow, and measure the heights of the plants at the end of the next growing season.

We then plot the heights of the offspring against the heights of their parents (Figure 9.3). In this case, we find that taller parents tend to produce taller offspring. We measure this tendency using a statistical technique called correlation. If offspring always exactly matched their parents, the correlation between parental height and offspring height would be 1.0. If there were no relationship, the correlation would be 0. In our example, the correlation is 0.41, and the slope of the line is 0.21; there is a resemblance, but some offspring are taller than their parents, while others are shorter. Negative correlations are also possible for some traits in some species, although they are very unusual.

One measure of the heritability of a trait is the slope of the line of a regression of offspring trait values on parental values. In the example above, because we used information from both parents, the slope is exactly equal to the heritability. If we had measured the trait in only one parent, we would have information about only half of the genes being contributed to the offspring, and the slope would have been one-half the heritability.

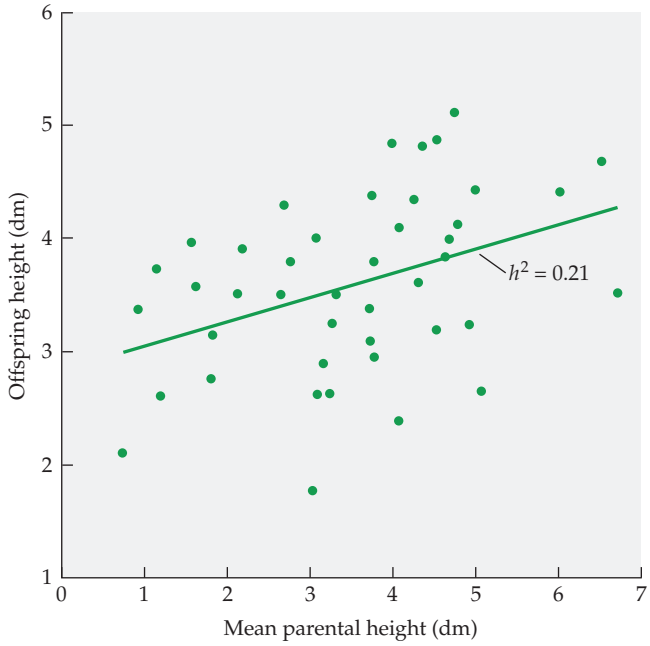



**Figure 9.3** Plot of offspring height against mean height of the parents in decimeters in the annual *Brassica campestris* (Brassicaceae) for plants grown in a greenhouse. The heritability of height in this species (the slope of the line) is 0.21. (Unpublished data courtesy of A. Evans.)

The other common way of measuring heritability is to measure the correlation among siblings. Suppose we took two seeds from each of many plants. We could germinate the seeds and grow the pairs of siblings, measure their heights, and construct a graph much like Figure 9.3, except that now the axes would be the heights of the two siblings, and each point would represent a sibling pair. Again, the slope would measure heritability, with the exact relationship depending on whether the plants shared both parents or just one parent. We can do such an analysis with cousins or with any individuals that are related, as long as we know their relationships. Nor are we restricted to using pairs of individuals. Various statistical techniques can be used to measure heritability in groups of related plants with different degrees of relatedness.

There is a critical distinction between the heritability of a trait and whether that trait has a genetic basis. Heritability requires that phenotypic differences among individuals be due, at least in part, to genetic differences among those individuals. In Box 9A we describe a case in which height is genetically determined. In that example, some individuals have genotype *AA*, some *Aa*, and some *aa*. Instead, imagine that all individuals in the population have the same genotype, *AA*. Assume, however, that height also depends on the amount of nitrogen in the soil. If the population is growing in a field that varies in soil nitrogen from spot to spot, then individuals will differ in height. However, none of those phenotypic differences will be due to genotypic differences. If we were to measure these plants, collect their seeds, and grow the offspring in that same field, the correlation between parental height and offspring height would be 0, and the heritability of height in that population would be 0. Yet there is still a gene in that population that determines height. In this case, the heritability of height is 0 because phenotypic variation in height is due to variation in an environmental factor, not to variation in the gene for height.

A lack of heritability can also be caused by effects of genetic dominance (Box 9B). Consider a recessive allele *b* that prevents chlorophyll from functioning in leaves (*BB* and *Bb* individuals both have functioning chlorophyll). If the species does not have an alternative carbon source (as some parasitic plants do), this lack would be lethal for individuals with two copies of the allele. Having functional chlorophyll is determined by a gene, but the heritability of that trait would be 0 because parental phenotypes would not predict the offspring phenotype (*BB* and *Bb* individuals have the same phenotype).

The previous example also demonstrates that the heritability of a trait depends on the frequencies of its

## BOX 9A A Simple Genetic System and the Resemblance of Relatives

This example is based on the work that Gregor Mendel did in the nineteenth century with garden peas (*Pisum sativa*, Fabaceae), although the details have been modified for illustrative purposes. Although it is based on a simple, one-locus system, nearly all traits of ecological interest are based on several to many loci. However, the same principles hold, no matter how many loci affect a trait.

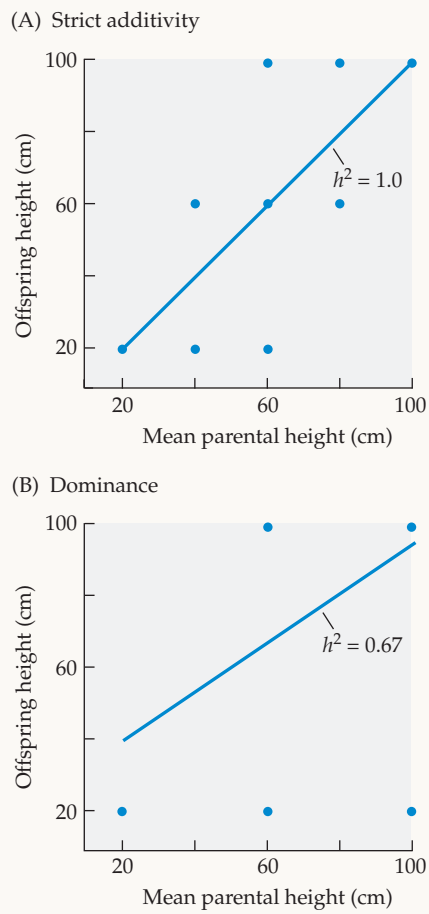
Consider a simple genetic system in which plant height is determined by a single diploid locus. We assume that individuals with genotype *AA* are tall (100 cm) and those with genotype *aa* are short (20 cm). We also assume that there are no environmental effects ($V_E = 0$ and $V_{G\times E} = 0$) or errors of development ($V_e = 0$).

### Case 1: Strict Additivity

If individuals with genotype *Aa* have a phenotype that is exactly intermediate between *AA* and *aa* individuals, then genetic variation is strictly additive. In this case, *Aa* individuals would be intermediate in height (60 cm tall). Because the effects of the alleles are strictly additive, we can predict the phenotypes of the offspring of a cross. If both parents are tall, the cross will be *AA* × *AA*, and all offspring will be tall. If both parents are short, the cross will be *aa* × *aa*, and all offspring will be short. If one parent is tall and the other short (*AA* × *aa*), all offspring will be 60 cm tall (*Aa*). That is also the height that we get by averaging the parental phenotypes; the mean offspring phenotype equals the mean value of the parents' phenotypes. If one parent is 100 cm tall and the other is 60 cm tall (*AA* × *Aa*), half the offspring will be 100 cm tall and half will be 60 cm tall. Again, the mean value of the parents' phenotypes, 80 cm, exactly equals the mean value of the offspring phenotypes. Note that, for this cross, no parent or offspring is actually at the mean height; the mean is a descriptor of the group, not a property of any particular individual. A graph of mean parental phenotype against mean offspring phenotype (part A of the figure) has a slope of 1.0. That is, the heritability of this trait is 1.0 because we can perfectly predict the mean offspring phenotype from our knowledge of the parental phenotypes.

### Case 2: Dominance

Now assume that *A* is dominant to *a*, such that *Aa* individuals are 100 cm tall. In this case, predicting offspring phenotypes becomes more difficult. If both parents are short, all offspring will be short. But if both parents are tall, their genotypes could be both *Aa*, or both *AA*, or one *AA* and the other *Aa*. In the latter two instances, all offspring will be 100 cm tall. But if both parents are *Aa*, then 1/4 of the offspring will be *aa* and will be short. The mean offspring phenotype will be 80 cm (3/4 × 100 + 1/4 × 20), even though the mean phenotype of the parents was 100 cm. If we assume that the two alleles exist at equal frequencies in our population, then a graph of mean parental phenotype against mean offspring phenotype will have a slope of 0.67 (part B of the figure). The heritability of the trait is less than 1.0 because some of the genetic variation is nonadditive due to the dominance relationship. In other words, some offspring differ phenotypically from their parents because of the effects of dominance; if they do not inherit the dominant allele from either of their parents, they do not resemble their parents. The exact heritability of a trait in a population depends on both the degree of dominance and allele frequencies in the population.




Plot of mean offspring phenotype against mean parental phenotype for two genetic systems. The slope of the regression line is the heritability of the trait. (A) Strict additivity; slope = 1. (B) Complete dominance; slope = 0.67. These heritabilities assume equal frequencies of the two alleles in the population.

alleles in the population. When the frequency of *A* is 1.0—all individuals have the *AA* genotype—the heritability of the trait is 0. Thus, heritability estimates for the same trait can differ among populations, or in the same population measured at different times. Heritability estimates are always specific to the population and environment in which they are measured.

## BOX 9B Using Genes to Track Pollen and Seeds and to Identify Species

In this chapter we mostly focus on the genetics of phenotypes—visible traits—because natural selection occurs on traits. Those traits are where most of the evolution and adaptations that are important for a plant's ecology happen. Those traits are almost always the result of several to many loci, and the measures of heritability and genetic variation discussed in the main text are the most appropriate way to study them using what is termed "quantitative genetics," or the genetics of phenotypes. But what about studying the genes themselves? Those studies are very useful for discovering the genetic basis of a trait, although currently not very useful for understanding the ecology of that trait. In the main text, we identified the genes that are responsible for growth responses to low light, but we do not need that information to determine how growth in high versus low light affects plant fitness (see Figure 9.8). There are two ways, however, that knowing a plant's genes can be very useful.

The first use is in tracking the movement of pollen and seeds. Unlike many animals, where we can track the movement of an individual by putting a band on a bird's leg or a radio collar on a lion's neck, for example, seeds and pollen are generally much too small and numerous to try to follow in that way. Instead, we can follow their movement by looking at their genes. Every individual, except for identical twins, has a unique DNA sequence. Even if two individuals are identical at 99.9999999% of their genome, because each genome contains tens or hundreds of millions of DNA base pairs, there will still be some that differ. We can use those differences to identify the source of a pollen grain or a seed. For simplicity, let's say that there are four individual plants. Plant 1 has a gene with the following DNA sequence: AATCGCTA. Plant 2 has the sequence AATAGCTA (note the change in the fourth base); plant 3 has the sequence TATCGCTA; plant 4 has the sequence GATCGCTA. For this stretch of DNA, each individual has a unique sequence. If we then go to a flower where pollen has been deposited and sequence the DNA in that pollen, we can determine exactly which plant that pollen came from, thereby tracking its movement. For a seed, we take advantage of the fact that its mitochondria and chloroplasts were from its female parent. (There are exceptions, but they are very rare.) Because mitochondria and chloroplasts have their own DNA, again we can track back to the plant where the seed developed. Also, in angiosperms the seed coat and its DNA is from tissue from the maternal parent (see Chapter 6), which can be compared with the DNA of the embryo. Any genes in the embryo that are not in the maternal tissue must have come from the paternal (pollen) parent.

How easy is it to find genes that differ among individuals? It used to be very hard. Now, however, with the cost of DNA sequencing down to a tiny fraction of a penny for each base pair, and dropping every year, it is very easy. Genomics is the study of very large sections of an individual's DNA. Genomics is much discussed in the press, and you may have heard about its importance in some of your courses. Its primary importance is in studying which genes and other parts of the genome are responsible for determining traits. For ecological problems, genomics mostly gives us a way to sequence enough of the DNA of an individual and possibly every individual in a population (if it is not too big) that we can find enough differences to allow us to determine the source of a pollen grain or a seed.

The second use is in identifying species. If we want to determine all of the species in a plot or a community (see Chapter 12), we need to carefully look at each individual and decide what species it is. In some cases, this is easy. In a northern forest that only has seven or eight species of tree growing in it, each species will have a unique leaf shape or bark type. But what about a tropical forest that might have 100 species of trees, many of which have identical-looking leaves? When we cannot use visible traits, we can use DNA sequences to identify species. Unlike the case where we are trying to uniquely identify individuals, now we want a part of the DNA where everyone within a species has exactly the same sequence, but where that sequence differs among species. Some parts of the genome evolve more slowly that others because of strong selection against change. Again, through the use of genomics we can sequence enough DNA of each species that we can find places that have these sorts of differences among species. This use of genomics is sometimes referred to as "DNA bar coding" because the DNA sequence acts like a bar code on a package in the supermarket. DNA bar coding cannot be used to identify new species; you have to have known reference samples to establish which stretches of DNA have the right sorts of variation among, but not within, species. But once you have those, they can be a big help in identifying the species of a plant from just a small fragment. Recent advances in sequencing technology have created devices that can be attached to a smartphone for quick IDs in the field. While still in the development phase, these devices have great promise for plant community surveys in the near future.

### Phenotypic variation can be partitioned into genetic and nongenetic components

Another way of thinking about heritability is to consider the various sources of phenotypic variation described above. If we measure height in a population of annual plants, it will vary due to differences in genotype, environment, and errors of development. The fraction of that variation that is caused by genetic differences is

heritability. Mathematically, we can express this idea as follows. First, consider the total phenotypic variance, which measures how much the values of some phenotypic measure differ from the average value for the population, across all of the individuals in the population. We symbolize phenotypic variance as $V_P$. By a combination of experimental and statistical techniques, we can determine how much of that variance is due to different causes. Breaking up variance into its components is called variance partitioning. In the simplest case, we might want to partition the phenotypic variance only into that part due to genetic variance ($V_G$) and that part due to all other causes ($V_e$):

$$V_P = V_G + V_e$$

Then, heritability would be $h^2 = V_G/V_P$, the fraction of phenotypic variance that is due to genetic differences among individuals. This concept of heritability is identical to that of heritability defined as the resemblance among relatives, just a different way of measuring it.

Genetic variance can be further partitioned, however. In a diploid organism, **dominance** occurs when the expression of an allele depends on the properties of the other allele at the same locus (see Box 9A). In both diploid and haploid organisms, **epistasis** occurs when gene expression depends on the properties of alleles at other loci. For example, flower color pigments are produced through a chain of biosynthetic steps, each of which is controlled by a different enzyme that is coded for by a different gene. The amount of genetic variation that is manifested in a population can be a result of differences in the direct expression of alleles at a locus (additive variance, $V_A$), or differences in the expression of different combinations of alleles at each locus (dominance variance, $V_D$), or differences in the expression of combinations of alleles at different loci (epistatic variance, $V_I$). Again, by various experimental and statistical procedures, we can partition this variance into its causes:

$$V_G = V_A + V_D + V_I$$

If heritability is calculated as just the fraction of phenotypic variance that is due to additive genetic variance ($h^2 = V_A/V_P$), then we are speaking of **narrow-sense heritability**. If it is calculated as the total genetic variance ($V_G$), then we are speaking of **broad-sense heritability**. This distinction is important because the response of a trait to natural selection depends on its narrow-sense heritability. Under both directional selection and stabilizing selection, all additive genetic variance for a trait under selection will eventually disappear, unless other processes introduce new variation. Phenotypic variation may remain, however, due to plasticity and errors of development.

Heritability values tell us whether there is genetic variance for a trait in a population and, if so, whether there is just a little variation or a lot of variation. In terms of evolution, the amount of genetic variance may impose a constraint on evolution. If there is no genetic variance ($h^2 = 0$), the constraint is strong. No matter how much natural selection there is on a trait, there will be no evolutionary change. If there is a little bit of genetic variation, the constraint is weak; there will be some change, but it will be small, and evolution will proceed slowly. If there is a lot of genetic variation, there is almost no constraint on the potential for evolutionary change in the population.

### The environment can interact with the genome to determine the phenotype

How does the environment influence heritability measures? This is where the ecology of a population probably has its greatest influence on phenotypic variation. Our original definition of heritability assumes that differences among individuals with different genotypes do not depend on their environment. That is, we assumed that if an individual is 10% taller than another when growing in one environment, it will still be 10% taller in a different environment. But what happens when this is not true? Suppose, for example, that when a certain plant species is grown under shady conditions, all individuals are small and about the same size, but when the plant is grown in a sunny spot, some of those individuals are much larger than the others due to genetic differences. In other words, the genetic differences are apparent in some environments, but not in others.

These kinds of differences in genetic expression as a function of the environment are referred to as **genotype-environment interactions**. Genotype-environment interactions are the genetic component of phenotypic plasticity. If we also consider the strictly environmental effects on the phenotype, we can further partition the total phenotypic variance of a population as

$$V_P = V_E + V_{G\times E} + V_G + V_e$$

That is, the total phenotypic variance is equal to the two components of plasticity, $V_E$ and $V_{G\times E}$, plus the genetic variance, $V_G$, and variance due to errors of development, $V_e$ (Figure 9.4).

The presence of variation resulting from genotype-environment interactions can have large effects on heritability. Heritability is not simply a result of the genetic differences among individuals: Those genetic differences must result in phenotypic differences. Some kinds of genetic differences among individuals never result in phenotypic differences; for example, some types of variation in noncoding regions of the DNA do not do so. In other cases, whether genetic differences result in phenotypic differences depends on the environment. When variation in genotype-environment interactions is present, the amount of genetic variance

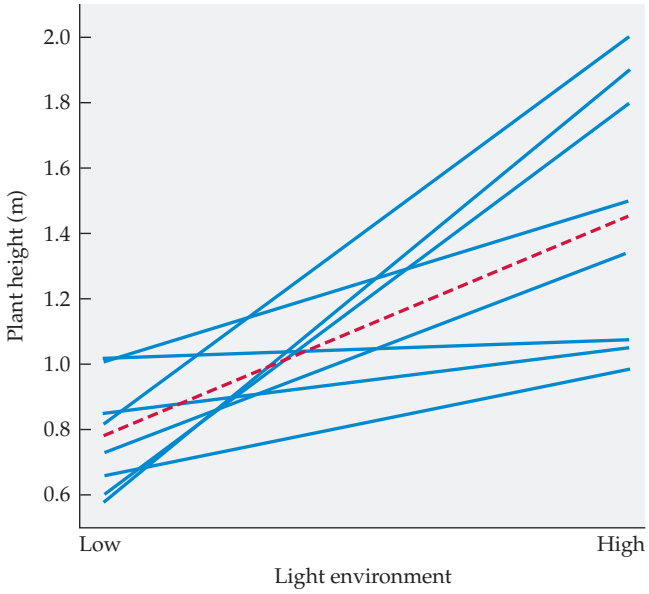



**Figure 9.4** An example of phenotypic plasticity and genotype-environment interaction for plants growing in high-light and low-light environments. In this example, cuttings were taken from individual genets and grown in each environment. Each solid line connects the mean heights for each genet in the two environments. The dashed line connects the overall means in each environment and represents $V_E$. The variation in average heights among genets represents $V_G$. The extent to which the solid lines are not parallel represents $V_{G\times E}$.

seen in each environment may differ. In the hypothetical example given above, the plants grown in the shade were all of similar size; in other words, phenotypic differences were minimized in that environment. If the heritability of height were measured only in the shade, we would conclude that it was very low, because the amount of phenotypic variance would be low. On the other hand, if heritability were measured only in the sunny environment, it would be larger. Thus, evolution would be constrained in the shady environment because of a lack of heritable variation.

### Genotypes are often nonrandomly distributed among environments

The final term that must be included in accounting for the total phenotypic variation of a trait in a population is **gene-environment covariance**, which is abbreviated Cov(G,E). A nonrandom relationship between any two factors is called covariance; such a relationship can be either positive or negative. Covariance is closely related to correlation; one is a mathematical transformation of the other. In this case, we are interested in covariance between genetic and environmental effects on a trait. Such covariation is often found in nature because individuals are usually not randomly distributed across environments. This is especially true of plants, as seeds often end up close to the maternal parent plant (see Chapter 6).

Positive covariance between genetic and environmental effects occurs when genetic traits are positively associated with responses to the environment. For instance, more vigorous competitors might dominate small patches of rich soil. In this case, plants that are already genetically capable of growing more rapidly will also be growing under better conditions, while slow growers will be relegated to poorer conditions. Thus, genetic differences in growth rate are exaggerated by environmental effects.

Nonrandom distribution of genotypes can act instead to minimize genetic differences. Negative covariance between genetic and environmental influences occurs when genetic and environmental influences have opposite effects. Consider a population of shrubs in which larger individuals produce more seeds. If you took the seeds of several individual shrubs and grew them all under optimal conditions in a garden, you would find a positive correlation between parental size and offspring size. In a natural population, however, most of the seeds will germinate in the shade of the parent plant. Because large plants produce many seeds, those seeds will be growing under crowded conditions. Plants that are genetically capable of growing larger (the offspring of the large shrubs) will be growing under poorer conditions, so genetic and environmental influences will be acting on size in opposite directions. Therefore, a negative covariance might exist.

Adding the term for gene-environment covariance to the equation for phenotypic variance given above, we end up with our final equation describing phenotypic variation in a natural population:

$$V_P = V_E + V_{G\times E} + V_G + Cov(G,E) + V_e$$

Since $V_P$ is the denominator in calculating heritability, changes in any of its components will affect the heritability of a trait. Thus, the ecology of a plant can affect all three components of natural selection: phenotypic variation through effects on development, fitness differences through effects on survival and fecundity, and heritability through genotype-environment interactions and gene-environment covariance.

## 9.3 Patterns of Adaptation Are the Result of Natural Selection

Natural selection can result in three different patterns of adaptation. First, individuals may become specialized to perform best in different environments. Within a population, for example, individuals of one phenotype might have the highest rate of survivorship during periods

of drought or in the driest spots, while individuals of another phenotype might survive best during wetter periods or in microhabitats with moister soil. Second, individuals may become phenotypically plastic, changing their form in different environments so as to match the trait value with the highest fitness in each set of environmental conditions. In response to changes in water availability, for example, individuals might produce leaves of different sizes and shapes. This pattern would allow them to have high survivorship during both wet and dry periods. Finally, all individuals in a population may converge on a single, intermediate phenotype that does at least passably well in all environments, but not as well as plants specialized for the different environments. This last pattern is sometimes referred to as a jack-of-all-trades strategy, as in the saying "A jack-of-all-trades is a master of none."

Which of these patterns of adaptation will prevail depends on a complex combination of factors, including how variable the environment is, how that variation is distributed in time and space, and how much genetic variation exists in the population. If the environment varies over short time periods, then phenotypic plasticity is often favored. Usually we think of phenotypic plasticity with regard to traits that become fixed during development, such as leaf shape, but it may also exist for traits that are reversible, such as some defensive chemical compounds (see Chapter 11). In aquatic plants from many unrelated families, leaves on the same plant may be produced above or below the water. Submerged leaves are characteristically feathery and highly dissected, while emergent leaves (those above the water surface) on the same plant have very different shapes, often less dissected or entire (Figure 9.5). In water, $CO_2$ diffusion is much slower than in air, and the highly dissected underwater leaves have increased surface areas and a smaller boundary layer (see Chapter 3), allowing better $CO_2$ uptake rates. When environmental changes occur over much longer time spans, however, phenotypic plasticity is much less likely to be favored. Instead, specialization is often favored, giving rise to patterns such as the contrasting leaf shapes among ecotypes of *Geranium* and *Achillea* discussed later in this chapter.

## Heavy-metal tolerance is an example of genetic differentiation

Let's examine one of the best-documented cases of natural selection and local adaptation in plants. Many other examples of adaptation are presented throughout Parts I, II, and III of this textbook, although we do not examine in detail the evidence that those traits are the results of natural selection. The example here is also important historically: It was the first demonstration of fine-scale genetic differentiation in response to selection by a known factor in the environment. The example involves a **cline**—a gradient in allele frequencies or another population characteristic—resulting from genetic differentiation and adaptation. Clines can occur over very short distances as well as at larger geographic scales.

In the 1960s, A. D. Bradshaw and his students (e.g., Jain and Bradshaw 1966; McNeilly and Antonovics 1968; Antonovics and Bradshaw 1970) began to study local adaptation of grasses to differences in soil conditions due to mine waste contamination. In Great Britain, zinc and copper have been mined for centuries. The soil left over from the digging and ore extraction, known as tailings, was simply dumped outside the mines. Although most of the ore had been removed, these tailings still contained high concentrations of the metals, too low to be worth extracting but still high enough to be toxic to most plants growing on the soil. Some plants, however, managed to grow on the tailings—in particular, the grasses *Anthoxanthum odoratum* and *Agrostis tenuis*. Although some of the mines had been abandoned for only a century—probably fewer than 40 generations for these perennial species—the researchers found that populations of these grasses on the mine tailings were tolerant of the heavy metals, while populations growing in ordinary adjacent pastures were not. The tolerance for heavy metals was achieved by the evolution of biochemical mechanisms that prevent the uptake of the toxic ions. These differences in the genetic composition of populations occurred over very short distances—within a meter of the mine boundary (Figure 9.6).

The abrupt nature of the genetic boundary illustrates several aspects of the evolutionary process. First, the alleles for heavy-metal tolerance did not spread outward from the population on the mine tailings. Why? Experiments were performed in a greenhouse in which plants from both the mine tailings and the adjacent pasture were grown in soil with and without the heavy-metal contaminants. As expected, plants from the pasture population died when grown with the heavy metals, usually as seedlings. However, when grown in the absence of the heavy metals, the pasture plants grew much more quickly than those from the mine tailings population. Something about the mechanism of heavy-metal tolerance reduced growth under uncontaminated conditions. This interaction is an example of a **trade-off**, a reduction in fitness in one feature of an organism due to an increase in fitness in another feature. In this case, heavy-metal tolerance and growth in uncontaminated soils are traded off against each other. Thus, different alleles are favored on and off the mine tailings, an example of disruptive selection. As a result, what began as a single population has diverged genetically into two populations due to natural selection.

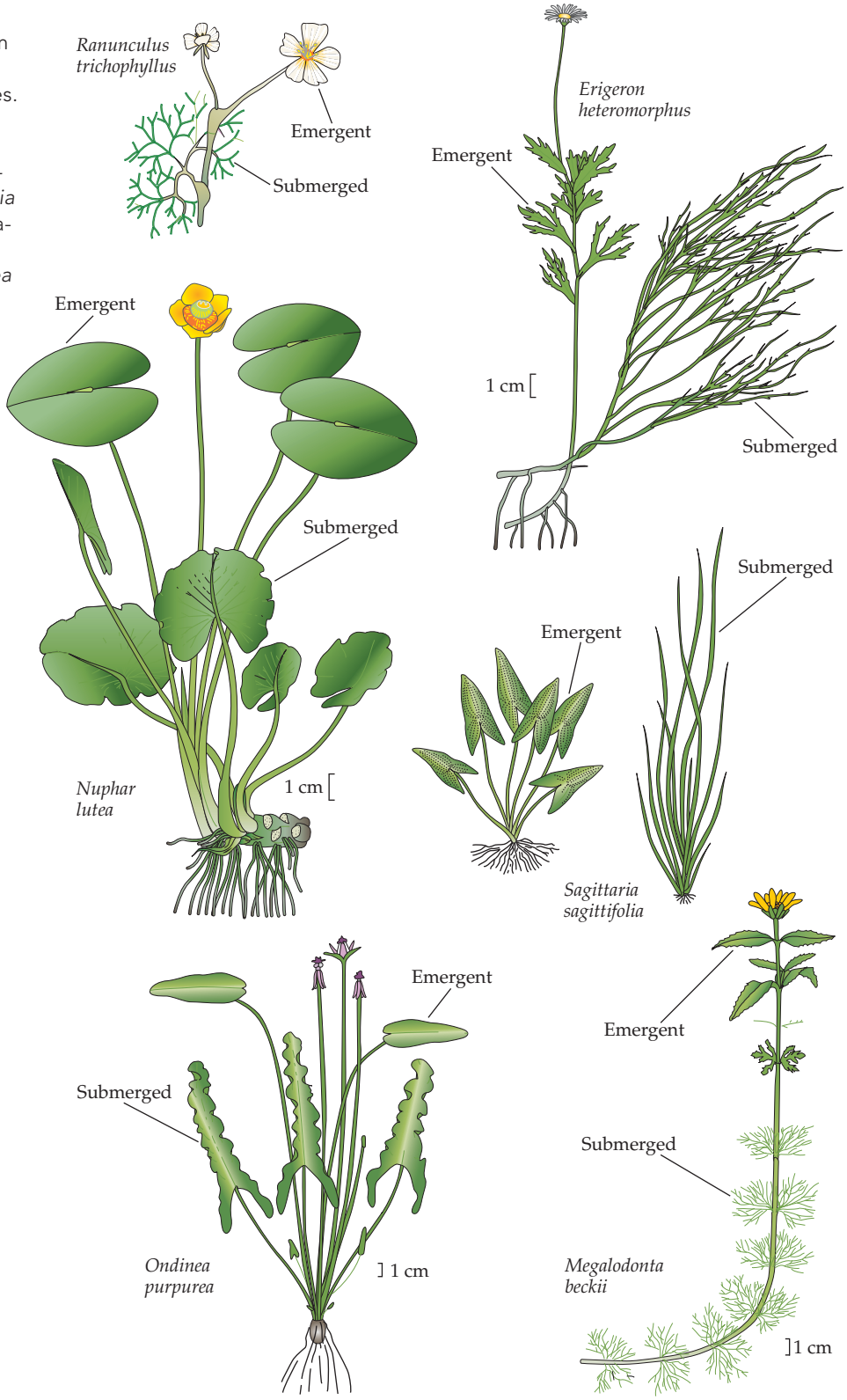



**Figure 9.5** Phenotypic plasticity results in adaptive differences in form between emergent and submerged leaves of several aquatic plant species. Clockwise from top left: *Ranunculus trichophyllus* (threadleaf crowfoot, Ranunculaceae), *Erigeron heteromorphus* (fleabane, Asteraceae), *Sagittaria sagittifolia* (Hawaii arrowhead, Alismataceae), *Megalodonta beckii* (Beck's water marigold, Asteraceae), *Ondinea purpurea* (purple water lily, Nymphaeaceae), and *Nuphar lutea* (yellow pond lily, Nymphaeaceae).

Other features of the populations have also evolved as a consequence of this divergence. All grasses are wind-pollinated (see Figure 6.6B), so pollen from the adjacent pastures can easily land on the stigmas of plants on the mine tailings, and vice versa. The abruptness of the genetic boundary was found to be related to the prevailing wind direction, with more of the "wrong" genotype being found on the downwind side. If an individual is

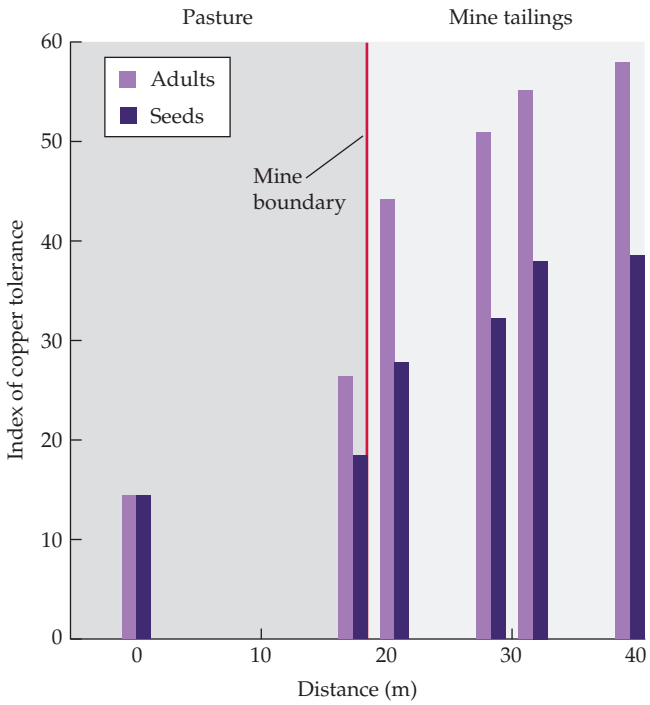



**Figure 9.6** Ability of adults and offspring grown from seeds of *Agrostis tenuis* collected from mine tailings and from adjacent uncontaminated pasture to grow in the presence of copper. A century after the mine was abandoned, the genetic composition of the grass population changes dramatically at the mine boundary, where alleles for enzymes that prevent the uptake of toxic ions are suddenly favored. (After T. McNeilly. 1968. *Heredity* 23: 99–108.)

pollinated by a plant from the other population, then its offspring will be less well adapted to the local conditions, as shown by the lower tolerance of the plants grown from seed in Figure 9.6. The problem of "wrong" pollinations results in selection for mechanisms that reduce cross-pollination. In these populations, two mechanisms were involved in reducing outcrossing (Antonovics 1968; McNeilly and Antonovics 1968). First, individuals on the mine tailings tended to flower earlier than those in the pasture, with the difference being greatest for plants just at the boundary of the mine tailings. Second, plants near the boundary tended to self-pollinate more than plants farther away from the boundary. Both changes reduced the chance of an individual receiving the "wrong" pollen from plants adapted to the alternate conditions (toxic or uncontaminated soils). These changes in mating patterns were a secondary consequence of the primary selection for heavy-metal tolerance. They acted to reinforce the genetic differentiation of the populations. Over a long enough time, such reinforcement can lead to total genetic isolation of the populations, and ultimately to speciation.

These studies clearly demonstrate, for several reasons, that the differences between populations were due to natural selection. First, the investigators demonstrated the presence of all three necessary components of natural selection: phenotypic variation, fitness differences, and heritability. Second, they found that secondary predictions were also met. The change in flowering time, for example, was a consequence of natural selection for heavy-metal tolerance. Third, they found the pattern of adaptation repeated both among populations of the same species at different mines and among different species. Such repeated patterns of differentiation would be very unlikely to have occurred by chance alone. For all of these reasons, we conclude that the differences in heavy-metal tolerance between populations on and off mine tailings are due to evolution by natural selection.

## Adaptation to different light conditions is an example of adaptive plasticity

For the most part, plants are not mobile. This observation, while obvious, has profound implications for the evolution and adaptation of plants. If the environment becomes unfavorable in one spot, a mobile animal can move to a more favorable location. But what can a plant do?

Because rooted plants cannot move, they must be able to tolerate environmental variation. Plasticity might help a plant perform better under different environmental conditions. But this may not always be the case; in fact, plasticity may lead to changes that are neither advantageous nor disadvantageous, or even to changes that are harmful. How do we test whether plasticity is adaptive—that is, whether it contributes to plant fitness—in any particular case? Can we determine the conditions under which it is most advantageous? Many characteristics of plants vary with the environment. While some of this variation is adaptive, some changes in plant appearance and function are merely unavoidable consequences of plant physiology, such as having yellow leaves when deprived of sufficient nitrogen. A meta-analysis of phenotypic plasticity in plants found that only half of the measured traits were plastic, and only one-third of the traits (two-thirds of the plastic traits) showed adaptive plasticity (Palacio-López et al. 2015). Demonstrating that phenotypic plasticity is adaptive requires, first, that there be more than one possible response to different environmental conditions and, second, that there be a consistent relationship between plasticity and fitness. The following study demonstrates how one can test the hypothesis that plasticity is adaptive.

Johanna Schmitt and her colleagues studied populations of *Impatiens capensis* (jewelweed, Balsaminaceae) growing under different light conditions (Schmitt 1993; Dudley and Schmitt 1995, 1996). This species is an annual that grows in deciduous forests of the northeastern and north-central United States. It grows

in conditions from full sun to forest understories, often along streams, and can reach very high densities.

A typical response to crowding in plants is stem elongation. This strategy allows a plant to overtop its close neighbors, thus gaining more light (Figure 9.7). In uncrowded situations, however, plants remain short because elongation has costs as well as advantages. More resources must be put into support structures rather than flowers and seeds, and the elongated stem is thinner, placing the plant in greater danger of falling over. Is it possible that a plant can tell whether it is growing near other plants and is thus at risk of eventually becoming shaded by its neighbors? If so, a plant finding itself in this situation will benefit by elongating its stem as it grows. If it can detect the fact that it is growing in an uncrowded environment, it may be better off not elongating its stem very much. It turns out that the elongation response is controlled by the ratio of red to far-red light that a plant receives, which is detected by chemicals called phytochromes (see Chapter 2).

Is the elongation response truly adaptive? Schmitt and colleagues addressed this question by manipulating plant form and growing conditions. Sets of seedlings from the same maternal plant were initially grown in a greenhouse under two conditions: (1) with filters that blocked the red part of the spectrum and (2) with full-spectrum controls that reduced light by the same amount without changing the red:far-red ratio. These treatments created two types of plants, an elongated form and a short, stout (suppressed) form. These experimental plants were then transplanted into a forest at either high or low density. Fitness was measured as the total number of seed capsules present at the end of the summer. As predicted by the adaptive plasticity hypothesis, the elongated form had higher fitness under crowded conditions, while the suppressed form had higher fitness at low densities (Figure 9.8).

The pattern of plasticity also differed in an adaptive way among populations growing in different natural environments. The researchers compared three populations growing within 1 km of one another, one in a forest clearing and the other two under the forest canopy. They grew plants from all three populations

Courtesy of Johanna Schmitt

**Figure 9.7** *Impatiens capensis* (Balsaminaceae) grown in the sun (high red:far-red ratio) and in the shade (low red:far-red ratio). Short, stout plants are favored in low-density (sunny) conditions, while tall, thin plants are favored in crowded (shady) conditions.

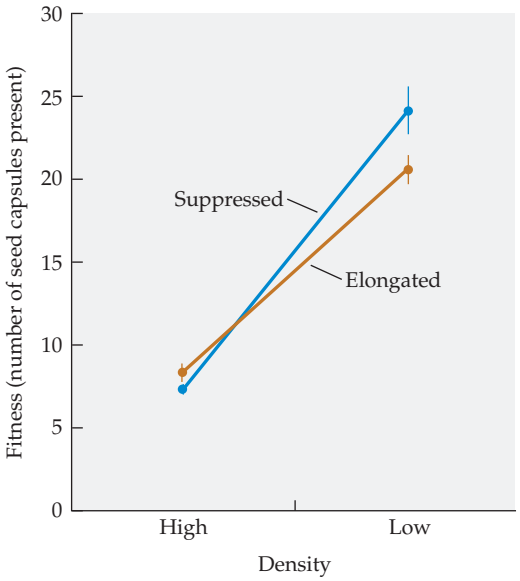



**Figure 9.8** A test of adaptive plasticity of stem elongation in *Impatiens capensis*. Elongated plants were produced in a greenhouse with filters that blocked the red part of the spectrum; suppressed plants were produced with full-spectrum controls that reduced light by the same amount without changing the red:far-red ratio. The two sets of plants were then grown in a forest at high or low density. The bars indicate one standard error. (After S. A. Dudley and J. Schmitt. 1996. *Am Nat* 147: 445–465.)

together in a greenhouse. The clearing population had a stronger elongation response than the other two populations. This result is predicted on adaptive grounds: In the forest, a plant that grows taller than its neighbors will still be in the shade. On the other hand, in a clearing, a plant that overtops its neighbors will be in bright light. Thus, being plastic is more adaptive for the population in the clearing. As in the heavy-metal tolerance study, the researchers showed the existence of the three necessary components of natural selection: Plants varied in how plastic they were, differences in plasticity led to fitness differences, and plasticity was shown to be heritable. Finally, they showed that differences in plasticity among populations were in the direction expected if plasticity were evolving by natural selection.

## Environmental effects can extend over generations

So far we have examined ways that the environment affects natural selection through direct interactions with individuals. The environment also can have more indirect effects that play out over multiple generations. One pathway is through **parental effects** on the phenotype of their offspring, which are sometimes referred to as maternal effects because they are most often associated with interactions between mothers and their seeds or eggs. The simplest form of such effects is through maternal provisioning of seeds, for example by how much endosperm each seed has (see Chapter 6). The parent plant can also determine the environmental conditions under which the seed is likely to germinate. For example, Toshiyuki Imaizumi and colleagues (2017) showed that in *Arabidopsis thaliana* (Brassicaceae) the day length experienced by the maternal plant strongly influenced seed germination even more than the day length experienced by the seed itself.

Parental influences can extend beyond the seed stage through heritable changes in chromosomes. Without changing the sequence of DNA bases (the A, T, G, and C), methyl ($CH_3^+$) groups can be added to chromosomes. Chromosomal DNA is wound around protein structures called histones, and changes in their chemical makeup can change the structure of that winding. The environment can influence such changes in the chemical makeup of the chromosome or its associated molecules, thereby affecting the ways that genes are expressed and changing the phenotype of the individual. Changes that do not involve alterations of the DNA base-pair sequence are collectively known as **epigenetic effects**. Such effects are very important for determining the process of plant development. Some such changes have been shown to be heritable. For example, methylation patterns in *A. thaliana* can persist at least three generations (Whittle et al. 2009). So the environment can indirectly affect the phenotype of individuals in subsequent generations, which is known as **transgenerational plasticity**.

We do not yet know how often transgenerational plasticity occurs or how often it is adaptive. Even less is known about whether transgenerational plasticity mostly happens through direct parental effects on the offspring's phenotype or environment or through heritable epigenetic effects. One of the most extensive sets of studies is that of Elizabeth Lacey and colleagues (Case et al. 1996; Lacey 1996; Lacey and Herr 2000; Lacey et al. 2003), who looked at the effect of parental and grandparental temperature on growth, survival, and reproduction in the perennial herb *Plantago lancelata* (Plantaginaceae). Along with direct environmental effects on the seeds during development and through the female parental plant, they also found several indirect grandparental and male parental effects: the temperature of the grandparental plant affected leaf area and flowering time, and the temperature of the male parental plant affected seed weight, germination rate, and flowering time. Such indirect environmental effects are consistent with epigenetic changes to the DNA. Importantly, they showed that these changes also affected fitness as measured by offspring seed set and that the transgenerational plastic changes were in the direction that increased fitness. However, the direct effects of the environment during development were substantially larger than the maternal, paternal, or grandparental influences. More research is needed on this system and others before we know how often such indirect multigenerational effects are important as compared to direct, within-generation environmental effects.

## Phenotypic plasticity is important for understanding other ecological concepts

The term *plasticity* is used in two different contexts by ecologists, and it means two different things. We have already discussed one context—trait variation among individuals due to environmental influences. The other context is whether a species is capable of tolerating a wide range of conditions. Ecologists might at first believe that this species is highly plastic. This broad tolerance could be the result of individuals that are all genetically alike but are phenotypically plastic. However, the differently appearing individuals could be genetically different, with each type specialized to each environment. So, a species may consist of a combination of generalist, plastic individuals or specialist genotypes (or a combination of the two). The same terms—*generalist* and *specialist*—are also used to refer to different species if they are found in many different types of environments or just one or a few types of environments. In plant ecology, the idea that the total variation in a

species can result from some combination of plastic and nonplastic genotypes traces to an influential paper by Bradshaw (1965). However, only in the past 30 years have these distinctions begun to be more widely recognized (see reviews by Schlichting 1986; Sultan 1987; West-Eberhard 1989; Scheiner 1993; Via et al. 1995; Pigliucci 2001; West-Eberhard 2003; DeWitt and Scheiner 2004; Richards et al. 2006).

An important related concept is that of the **niche** of a species (Grinnell 1917), which describes the range of environmental conditions in which a species can live and reproduce. The niche of a species can also be used to describe the role of a species in a community, including its interactions with other species. The niche of a species is determined by the sum of the niches of all the individuals belonging to that species (Figure 9.9); each individual may have a wide or a narrow niche itself. We speak of two aspects of the niche of a species, the **fundamental niche** and the **realized niche** (Hutchinson 1957). The fundamental niche is the range of conditions that a species is physiologically capable of growing in; the realized niche is the range it is actually found in. Factors such as competition, herbivory, and lack of pollinators or seed dispersers all act to reduce a species' realized niche from the potential of its fundamental niche; these factors will be covered in more detail in Part III. A comprehensive review of niche concepts can be found in *Ecological Niches* by Jonathan Chase and Mathew Leibold (2003).

Parents can also influence their offspring through the process of **habitat construction**, the process by which an organism alters its own (or another species') environment (Sultan 2015). For example, the needles (leaves) of pine trees are acidic. Over many, many years as dead needles build up on the forest floor, the soil will become more acidic, favoring some species over others. Over shorter time periods, during the process of succession (see Chapter 13) as trees grow up, they reduce the amount of light on the forest floor, in some cases preventing their own seedlings from growing and instead favoring other species.

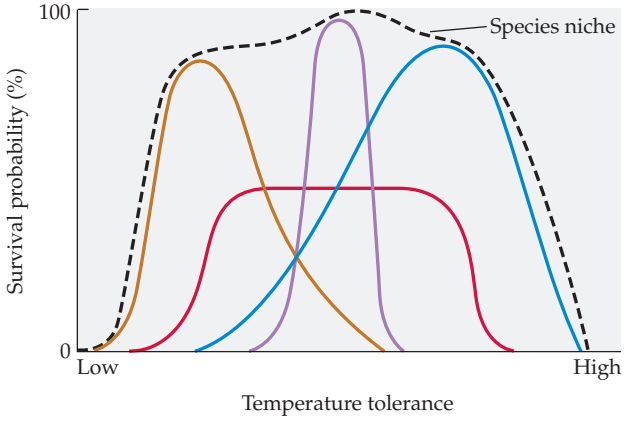



**Figure 9.9** The niche of a species is the sum of the niches of its members. Here we consider one aspect of a niche, temperature tolerance. In this graph, each solid curve is the survival probability of a single individual as a function of temperature. The temperature range of the curve is one measure of the plasticity of that individual. The dashed curve is the niche of the species.

## 9.4 Natural Selection Can Occur at Levels Other Than the Individual

Natural selection usually occurs at the level of the individual. Individuals differ in fitness, and genetic responses to natural selection are measured by looking at differences among the individuals in a population from generation to generation. However, none of the three components of natural selection requires that individuals be the only focus of the process. Plants, in particular, have growth forms that can result in other units being the focus of natural selection.

The first question we must address is, What is an individual? In many plant species, an individual is obvious. There is a central trunk or stem that grows, reproduces, and dies. But many other plants are **clonal**, existing as a set of genetically identical, possibly interconnected, but semiautonomous to fully autonomous units (see Chapter 6). Many grasses are good examples. From a single grass seed grows a plant that first sends up leaves and puts down roots. Later, it produces specialized underground stems called **rhizomes**, which cause the plant either to grow larger or to spread over a wider area. New leaves and roots produced at nodes along the rhizome are called **tillers**. Water and nutrients pass along the rhizomes to different parts of the plant. However, once the new tillers are large enough, the connections between them can be cut and each tiller can survive as an independent plant. Even if the connections remain, the transfer of water and nutrients among tillers may be small.

Deciding exactly what constitutes an individual, then, is not simple. The British ecologist John Harper proposed the following distinction (Figure 9.10): A **genet** is a genetic individual, the product of a single seed. A **ramet** is a potentially physiologically independent unit of a genet, such as the tillers of a grass plant. A single genet may consist of many separately functioning ramets. While individual ramets may come and go, a genet can exist for a long time. Quaking aspen (*Populus tremuloides*, Salicaceae) is a tree that can spread by sending up new trunks from its roots. While individual trunks

Pete Davis/CC BY-SA 3.0

John Harper

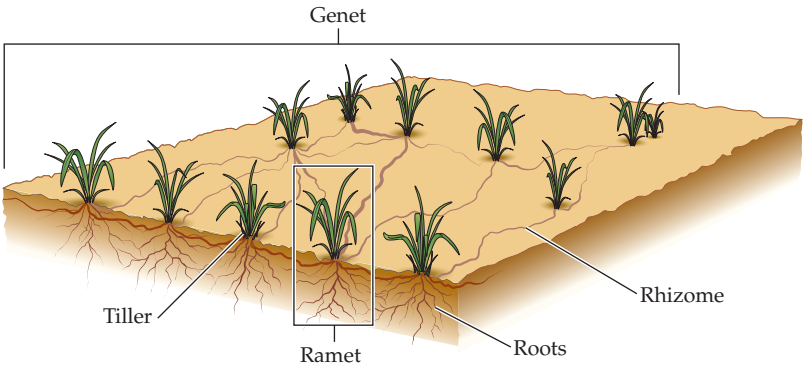



**Figure 9.10** Diagram of a grass plant that grows by vegetative spread. The entire plant is a single genet; each tiller constitutes a single ramet.

usually live only 50 or 60 years, one quaking aspen genet was estimated to be more than 10,000 years old (Mitton and Grant 1996; Figure 9.11). Other common examples of plants that grow clonally, or vegetatively, are kudzu (*Pueraria*, Fabaceae), strawberries (*Fragaria*, Rosaceae), and cattails (*Typha*, Typhaceae).

Natural selection can occur whenever its three components—phenotypic variation, fitness differences, and heritability—are present for a trait. While most selection occurs at the level of the individual, selection at other levels has been observed (Lewontin 1970). First consider selection within an individual plant. **Mutations** in gametes (egg or sperm cells) produce heritable variation. However, in plants, **somatic mutations**—mutations in ordinary cells making up the body of the plant, rather than in gametes—can also lead to variation that is subject to natural selection. If a somatic mutation occurs in a meristematic cell (see Figure 6.1) that gives rise to a new ramet, it can result in genetic variation among the ramets within a genet. If those genetic differences lead to phenotypic differences that affect fitness, then selection can act on the ramets. In addition, such a somatic mutation can become heritable, as the new ramets eventually flower and produce gametes, potentially carrying that mutation into the next generation. Selection can also occur above the level of the individual, among populations or among species. When trying to determine whether a particular feature of a plant species is a result of evolution by natural selection, we should consider the possibility that the feature is a result of selection at a level other than the individual. As a result, evolutionary processes can be complex: selection can occur at more than one level at a time, and these different selective processes might operate in the same direction or in opposite directions. Because evolution occurs over very long stretches of time, even very slow processes, or ones that occur very rarely, can be important.

Courtesy of Jeff Mitton

**Figure 9.11** This quaking aspen genet, growing in the Rocky Mountains of Colorado, is estimated to be more than 10,000 years old.

## 9.5 Other Processes Can Cause Evolutionary Change

One way of classifying processes that contribute to evolution is to group them into those that add genetic variation to a population and those that eliminate it. Natural selection acts to eliminate genetic variation as alleles that lead to increased fitness become fixed in a population and other alleles are lost. An important constraint on evolution is the lack of sufficient genetic variation for natural selection to act on. For evolution by natural selection to continue, genetic variation must reenter a population.

## Mutation, migration, and sexual reproduction are processes that increase genetic variation

The ultimate source of all genetic variation is mutation: change in DNA sequences. Mutation includes not only changes in single base pairs (e.g., an AT pair changed to a GC pair), but also deletions, insertions, rearrangements of parts of chromosomes, and duplications of part or all of a chromosome. Some mutations, such as a substitution in the third position of a codon, may have no effect on the phenotype. Other mutations, such as deletions, may destroy the function of a gene. Duplications of whole genes may have no immediate effect on the phenotype but may allow the evolution of a new function in one of the two copies. In order to produce a permanent change in a population, a mutation must occur in gametes or gamete-producing cells. In plants, these include much of the plant body, since flowers typically arise in many different sites on a plant. This is one important difference between plant and animal evolution: plants do not have specialized reproductive-cell lineages that are separated off during early development, but nearly all animals do. The consequence is that some somatic mutations in plants can potentially be passed on to sexually produced offspring, but mutations that humans have in our ears or livers, while they may affect our lives, cannot be passed on to children. Although mutations are generally random events, it is crucial to remember that this does not mean that evolution is always random (as is sometimes mistakenly believed). Random mutations are just the source of variation for other evolutionary processes, such as natural selection.

Mutations are relatively rare events. Changes in single base pairs occur at a frequency of $10^{-8}$ to $10^{-10}$ per base pair per generation. For an average-sized gene, mutations occur at a frequency of $10^{-5}$ to $10^{-7}$ per gene per generation. On average, in any given population of most species, one individual in ten has a new mutation somewhere in its genome each generation.

Another source of the genetic variation in a population is **migration**, the movement of individuals from one population to another, which can introduce new alleles into a population. Migration in plants occurs primarily by the movement of spores, pollen, and seeds (see Box 9B). For seed plants, the movement of pollen is particularly important. Pollen can be carried by the wind for tens of kilometers, or by insects, bats, or birds for varying distances. Seed movement tends to be more restricted, although studies of wind-dispersed seeds show that dispersal over fairly long distances is common. A study of five Canadian tree species, for example, found substantial densities of seeds nearly a kilometer from their source (Figure 9.12). Dispersal of seeds over very long distances also can occur through movement by animals. Pollen and seed movement is discussed in more detail in Chapter 6.

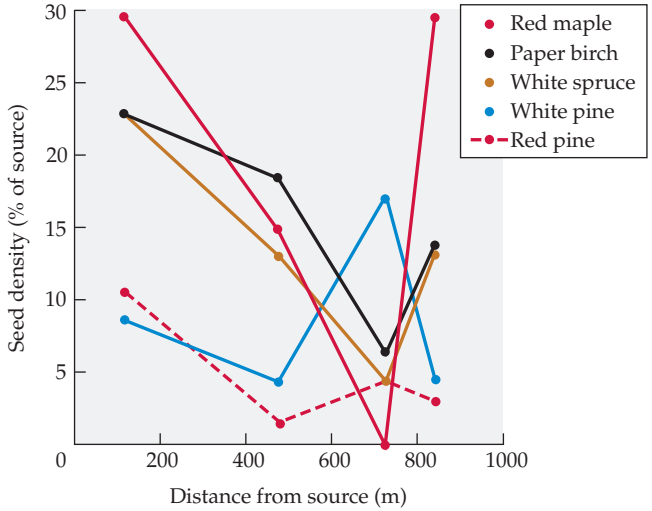



**Figure 9.12** Densities of seeds at varying distances from their source, expressed as a percentage of seeds found directly under the parental tree of each of five species. (After D. F. Greene and E. A. Johnson. 1995. *Can J Bot* 73: 1036–1045.)

Because gene expression can depend on the presence of alleles at other loci, new combinations of alleles at different loci can create new variation. For this reason, the process of sexual reproduction, which includes crossing-over, meiosis, and recombination, is an important source of new genetic variation.

## Genetic drift is a process that decreases genetic variation

Several processes can act to eliminate genetic variation from a population. Probably the most important of these, after natural selection, is genetic drift. The term **genetic drift** refers to changes in gene frequencies due to random sampling effects. We can illustrate genetic drift with a simple example: Imagine a population that consists of yellow-flowered and white-flowered plants. In this population, flower color has a genetic basis, and that color does not affect either survival or reproduction. If every plant has the same chance of surviving and reproducing, most of us intuitively think, flower color will not change over time. But it can. In one generation, by chance alone, yellow-flowered plants might have more offspring than white-flowered plants. In the next generation, the frequencies of the two flower colors in the population will have changed, but not because of natural selection. Rather, the change in trait frequency is due to genetic drift. But will things even out over time, as most people's intuition suggests? The answer is, Very rarely. Any amount of variation in the number of offspring among individuals will likely result in some combination of more offspring of one color than the other. Although the frequencies of the two colors may go up and down from one generation to the next, given enough time the genes responsible for one color or the other can be lost from the population purely by chance (Figure 9.13).

Genetic drift can occur at many life history stages, including meiosis, pollination, seed germination, and maturity. In most populations, changes due to genetic drift will be small. Generally, genetic drift is important only in populations smaller than 100 individuals; in populations greater than 1000 individuals, its effects are usually much smaller than those of other evolutionary processes. Thus, population size is a critical parameter in determining how evolution will proceed (see Figure 9.13).

*Ipomoea purpurea* (morning glory or bindweed, Convolvulaceae) provides an example of the effects of population size on genetic drift. This species is an annual weed of agricultural fields in the eastern United States.

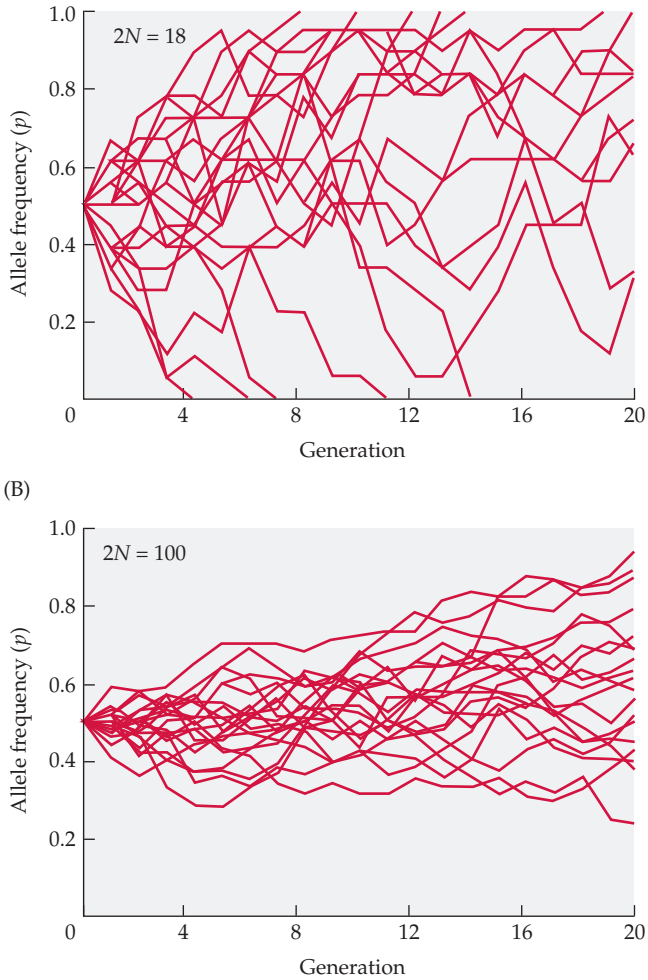



**Figure 9.13** Changes over 20 generations in gene frequencies due to random sampling. (A) An allele frequency of $p = 0.5$ implies that there are 9 copies of the *A* allele and 9 copies of the *a* allele in this small population. (B) An allele frequency of 0.5 implies that there are 50 copies of each allele in this much larger population, which results in smaller oscillations in allele frequencies. As population size decreases, the effects of genetic drift increase. (After D. Hartl and A. Clark. 1989. *Principles of Population Genetics*, 2nd ed. Oxford University Press/Sinauer Associates: Sunderland, MA.)

Its flowers vary in color; some individuals have blue flowers and some have pink flowers. Flower color is a simple one-locus trait, and blue is dominant to pink. The chief pollinators of this plant are bees, and gene flow among populations is primarily by pollen. Seeds are passively dispersed and stay close to the parent plant, although long-distance seed dispersal probably occurs in conjunction with agricultural or horticultural activities. Population densities vary among fields, and distances between plants range from 1 to 6 m.

A study of flower color was initiated by Michael Clegg and his associates (Ennos and Clegg 1982; Brown and Clegg 1984; Epperson and Clegg 1986) to determine the evolutionary processes responsible for variation within and among populations. The study looked at patterns both within fields and among fields across the states of Georgia, South Carolina, and North Carolina. The percentage of individuals with pink flowers varied greatly from field to field, from 5% in some fields to 55% in others. This variation occurred over a number of different distances, from as short as 50 m (the shortest distance between fields) to as long as 560 km (the farthest distance between fields). Fields that were located close to one another were no more similar in their color morph frequencies than distant fields. This pattern most likely means that long-distance seed dispersal occurs through (accidental) human activities.

In contrast, distance did affect the resemblance among neighboring individuals within fields (Figure 9.14). Within a field, the researchers found clusters of plants; some

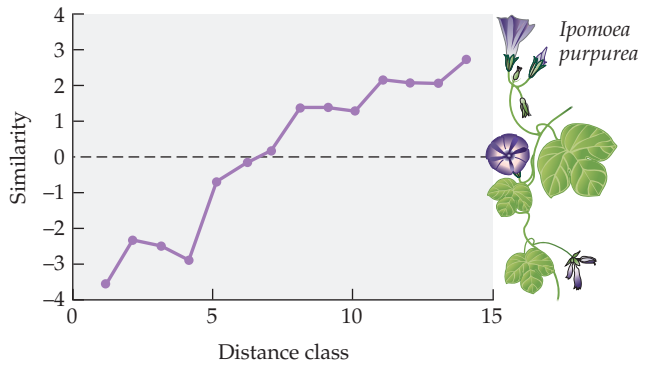



**Figure 9.14** Spatial autocorrelation in clusters of *Ipomoea purpurea* (morning glory, Convolvulaceae) across a field near Athens, Georgia. Similarity is measured as the likelihood that a pink-flowered plant will be found next to a blue-flowered plant. At short distances, the similarity is negative, indicating that pink-flowered plants tend to be found next to other pink-flowered plants, and the same for blue-flowered plants. In contrast, at longer distances, clumps dominated by blue-flowered plants alternate with those dominated by pink-flowered plants. The point at which the line crosses zero indicates the size of the average cluster, approximately 50 m in this field. (After B. K. Epperson and M. T. Clegg. 1986. *Am Nat* 128: 840–858.)

clusters had only plants with pink flowers, other clusters had only plants with blue flowers, and still other clusters had some plants with pink and some with blue flowers. On average, each cluster contained 10 to 15 plants. The pollinators—bees—tend to travel from one plant to the nearest plant, so alleles tend to move very short distances from one generation to the next. The researchers concluded that the pattern of variation in flower color within fields was the result of a combination of genetic drift and migration by pollen movement. Random sampling caused allele frequencies to become different between clusters. Pollen and seed movement caused nearby clusters to be similar but was not sufficient to keep more distant clusters from becoming different. **Spatial autocorrelation** (resemblance as a function of distance) can result from genetic drift, although it can also be caused by other processes.

### These evolutionary processes have important conservation implications

Population size is of particular concern for species conservation. In a widespread species, even if each population is small, genetic variation can still be maintained across the species as a whole as long as there is some migration between populations. If the entire species is reduced to a single population or a small number of populations, then maintaining genetic variation can become difficult. Genetic variation can be lost quickly through genetic drift and natural selection because unfavored genotypes do not leave many offspring. New variation through migration will not be available because the few populations will all contain the same alleles. New variation through mutation will be rare because the rate of new mutations in a population depends on the number of individuals. In such small populations, if the environment changes in such a way that new alleles would increase fitness, those new alleles may not be present, and the population may be unable to adapt to the new conditions. Such concerns are even greater today because of human-caused changes in the environment (see Chapter 19). Global warming may cause large changes in local environments in just a few decades. Small populations, unable to adapt, will be faced with extinction. And regardless of evolutionary effects, small populations are also at increased risk of extinction for demographic reasons (see Chapter 8).

## 9.6 Evolutionary Processes Can Affect Variation among Populations

So far, we have discussed evolutionary processes that affect variation within a single population. We can also consider how evolutionary processes affect variation among populations. Natural selection can increase or decrease variation among populations. If natural selection favors different trait values in different locations, different populations of a species may become adapted to different conditions. Eventually, the species may split into two species. Conversely, if natural selection favors the same trait values in different locations, the populations will tend to remain similar.

Mutation and genetic drift both tend to increase variation among populations. It is very unlikely that exactly the same mutations will occur in different populations. Because genetic drift is random, it is unlikely that different populations will experience exactly the same random changes. These tendencies for populations to diverge will be greatest in small populations.

Migration tends to decrease genetic variation among populations. Migration brings alleles into a population from nearby populations. These alleles may represent genetic variation previously lost from the population, or they may represent a mutation that has appeared in another population. If rates of migration between two populations are high enough, the populations will have the same alleles at almost the same frequencies.

All of these processes—natural selection, mutation, genetic drift, and migration—can act together to determine the evolution of a set of populations. Imagine a set of populations that initially are genetically identical. If the populations are relatively small, they may start to diverge from one another by genetic drift and mutation. One result is that new alleles, and new combinations of alleles, will appear in each population. These changes can result in several possible evolutionary outcomes. First, natural selection may favor different combinations of alleles in each population, resulting in the populations becoming different from one another. Sufficiently high rates of migration among populations, however, can sometimes prevent a population from evolving adaptations to the local environment. Therefore, another possible evolutionary outcome is that all of the populations will evolve to the average conditions. A third outcome may result if the populations differ in how well they each adapt to local conditions. The most successful population may send out many more migrants than the other populations. Eventually, all of the populations will resemble that successful one.

## 9.7 Ecotypes Are Different Forms of a Species That Are Adapted to Different Environments

Plant ecologists and taxonomists have recognized since before the time of Linnaeus, well over 250 years ago, that plants of the same species growing in different places may have different appearances. Early in the twentieth century, researchers in Europe and in the United States

began careful experimental work to better understand the causes of some of these differences. In Sweden, Göte Turesson (1922) collected plants belonging to several dozen species from habitats all over Europe and grew them in a common garden at a single site in Sweden. He found that many of the differences that he noticed when the plants were growing in their natural habitats were retained when they were all grown in the same environment. Turesson coined the term **ecotypes** to describe populations of a species from different habitats or locations that possess genetically based differences in appearance and function. Ecologists usually use the term *ecotype* to refer to such differences that appear to be adaptive.

At about the same time that Turesson was doing his experiments in Sweden, Jens Clausen, David Keck, and William Hiesey (1940) were doing similar work in California. These researchers established three common garden sites at low, intermediate, and high altitudes, from the Carnegie Institution at Stanford University, south of San Francisco, where they worked, up into the Sierra Nevada (the spectacular mountains that John Muir, the famous conservationist, called "the range of light"). Plants from a large number of species were collected along a **transect**—a line along which samples are taken—running from the coast of California inland, up to the highest part of the Sierra Nevada, and down the arid eastern side of the mountains (Figure 9.15). The plants were propagated vegetatively, and genetically identical copies of many different individual plants were grown in each of the three gardens. Clausen, Keck, and Hiesey studied two groups of species most intensively: several closely related species of *Achillea* (yarrow, Asteraceae), and *Potentilla glandulosa* (cinquefoil, Rosaceae).

Like Turesson, these researchers found that many of the differences in morphology and phenology among plants from different sites were still present when those plants were compared in the common gardens (Figure 9.16). *Achillea* plants from lower-altitude sites were larger than plants from higher altitudes, with longer leaves and taller flower stalks. Plants from higher

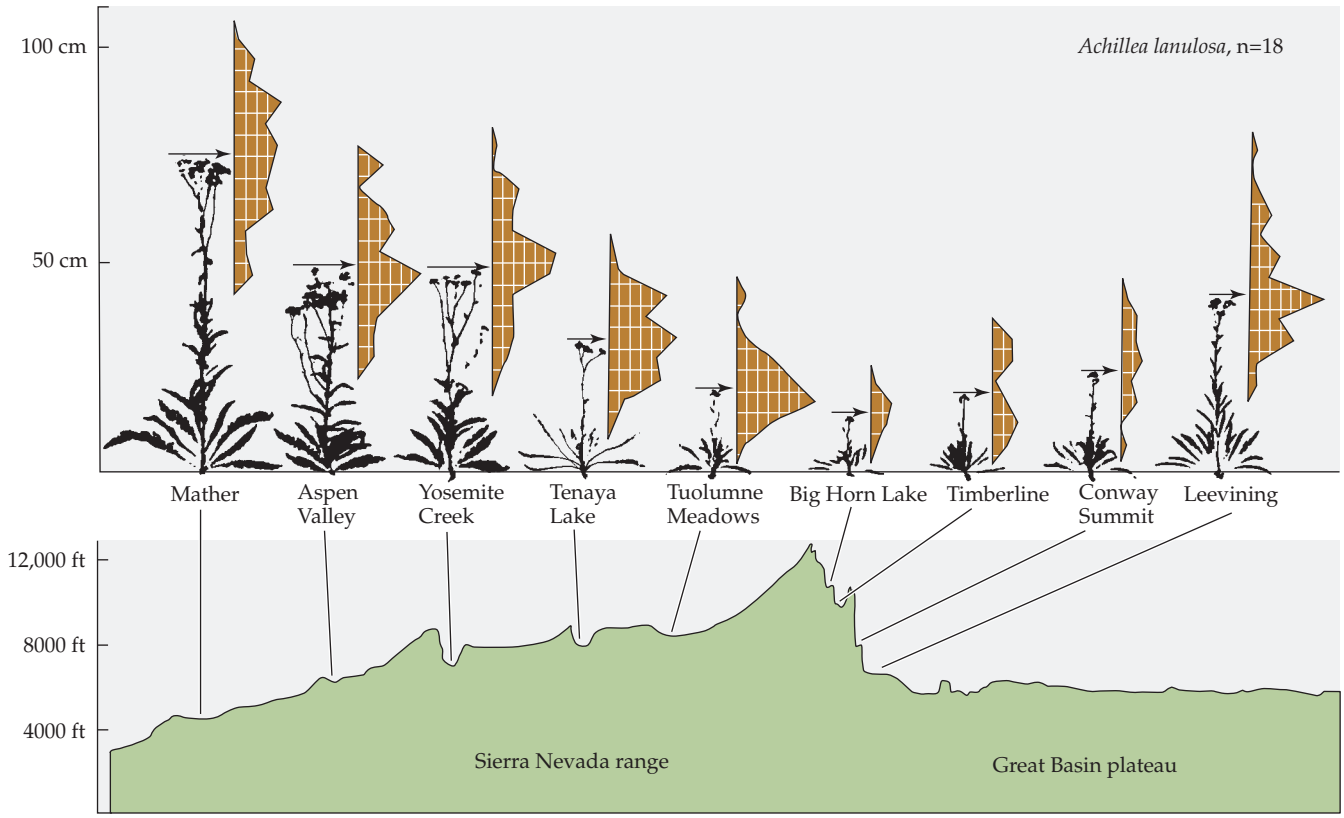
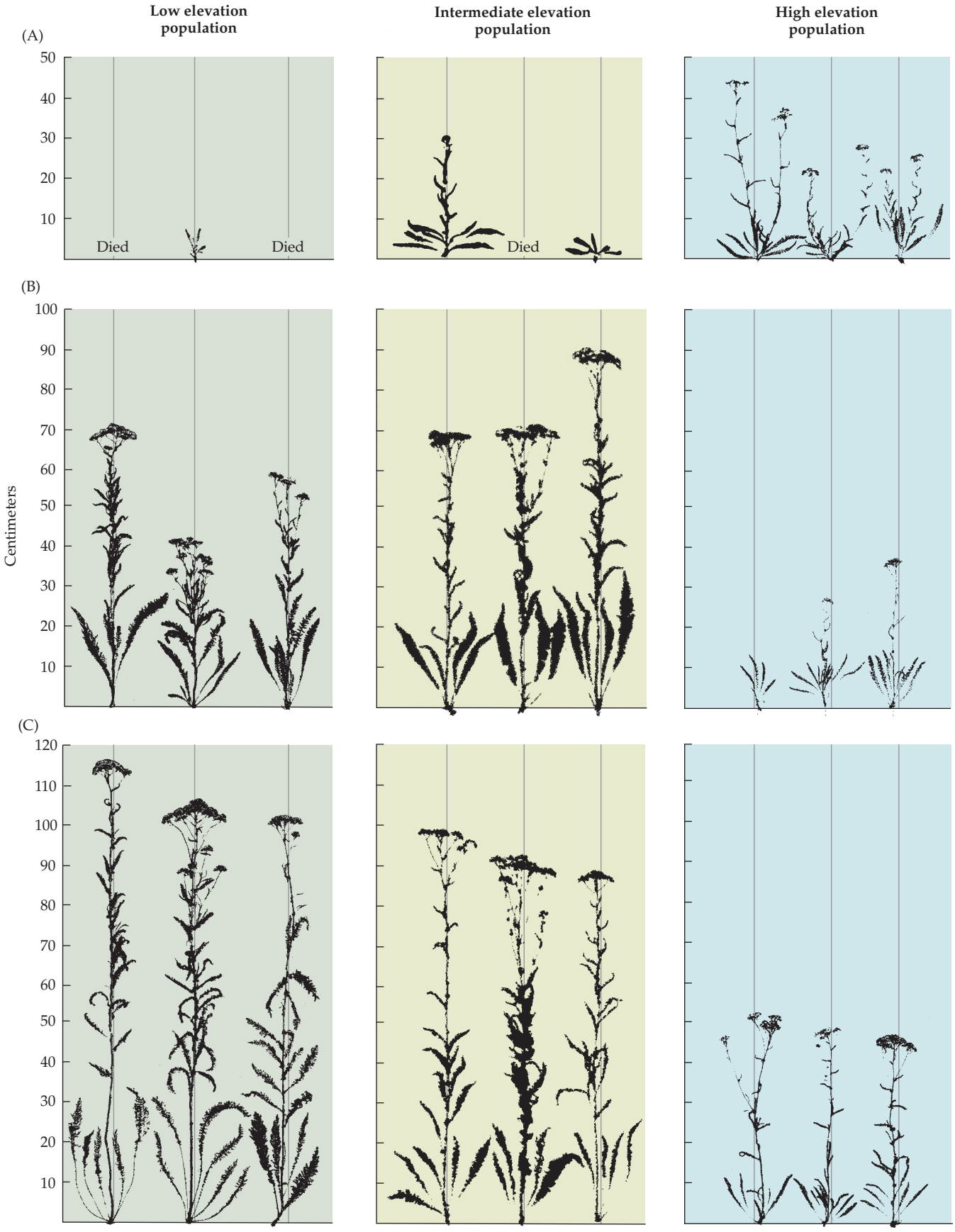



**Figure 9.15** Part of Clausen, Keck, and Hiesey's transect across California, from the foothills of the Sierra Nevada to the Great Basin desert, showing (below) the altitude of the sites from which plants were collected and (above) the appearance of *Achillea* plants when grown in a common garden at Stanford University near sea level. The small graphs show the variation in height among individual plants from each site, with the specimens pictured representing plants of average height and the arrows indicating mean plant height. Note that horizontal map distances are compressed. (From J. D. Clausen et al. 1948. *Experimental Studies on the Nature of Species. III. Environmental Responses of Climatic Races of* Achillea. Carnegie Institute of Washington, Washington, DC.)


Low elevation population
Intermediate elevation population
High elevation population
(A)
(B)
(C)
Centimeters
Died
Died
Died
50
40
30
20
10
100
90
80
70
60
50
40
30
20
10
120
110
100
90
80
70
60
50
40
30
20
10

◀ **Figure 9.16** Examples of plants (vertical panels from left to right) grown from clones of low-, intermediate-, and high-elevation populations of *Achillea* when grown in common transplant gardens at high elevation (A: Timberline, 2100 m) intermediate elevation (B: Mather, 1400 m) and low elevation (C: Stanford, 20 m). (From J. D. Clausen et al. 1948. *Experimental Studies on the Nature of Species. III. Environmental Responses of Climatic Races of* Achillea. Carnegie Institute of Washington: Washington, DC.)

altitudes generally entered dormancy sooner and began growing later than plants from lower, warmer sites. These alpine plants also tended to flower later than those from lower-altitude sites. Plants generally had the highest survival rate and best performance in the garden with the conditions that most closely resembled those in which they grew naturally (Clausen et al. 1948; Núñez-Farfán and Schlichting 2005). The researchers also noted striking differences in the appearance of the leaves of plants from different sites along the transect. Leaves from the high-altitude populations not only were smaller, but tended to have a dense gray pubescence (a mat of short hairs on the leaf surface) and were much more compact in shape. Leaves from lower-altitude plants were smooth and green and were highly **dissected** (had a blade divided into small but connected parts, making the leaf look very feathery).

Many years after these original studies, Jessica Gurevitch and colleagues (Gurevitch 1988; Gurevitch and Schuepp 1990; Gurevitch 1992) returned to the original sites to carry out further studies on the same populations that Clausen, Keck, and Hiesey had worked on. As anyone who has visited or lived in California might expect, many of the original sites where plants had been collected are now under asphalt. However, the Carnegie Institution still maintains the sites at Mather and Timberline, where two of the common gardens once were located, and plants from these intermediate- and high-altitude populations were collected for study. The populations differed genetically in their photosynthetic characteristics. There were also genetically based differences between the populations in both leaf size and leaf shape, particularly in the degree to which these complex leaves were dissected. As the earlier researchers had noticed, the Mather population had much more feathery, highly dissected leaves, while leaves from the Timberline plants were more compact (Figure 9.17).

Could there be an adaptive reason for these differences in leaf shape? Energy budget studies demonstrated that the highly dissected leaves of the Mather plants had a thinner boundary layer with greater heat conductance (see Chapter 3), making it more likely that they would remain close to air temperature. The compact leaves of the high-altitude plants had a thicker boundary layer and less heat conductance, and energy budget calculations suggested that these leaves could warm considerably above air temperatures. Thus, in the warm, dry environment at the lower altitude, leaves would remain relatively cool, while high-altitude plants might be able to warm up above the chilly air temperatures common in their environment to maximize photosynthesis and growth.

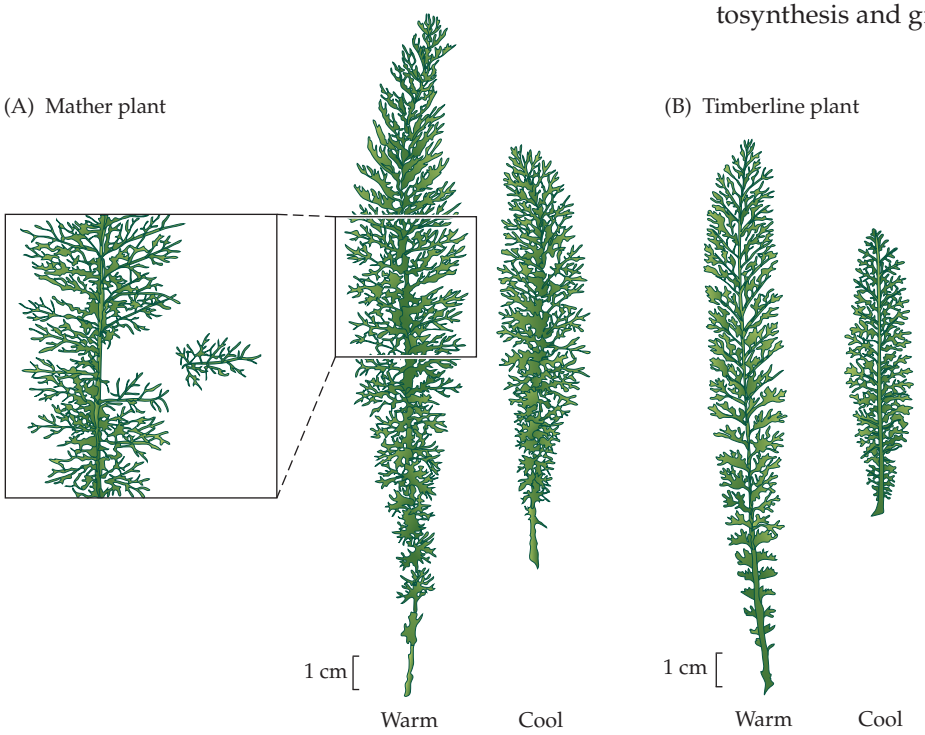



**Figure 9.17** Leaves of plants from (A) the Mather and (B) the Timberline populations of *Achillea lanulosa* when grown together in common greenhouse environments under warm and cool conditions. A part of the midsection of a Mather leaf is enlarged to show the complex shape. The longer Mather leaf is approximately 11.5 cm long. (After J. Gurevitch. 1992. *Genetics* 130: 385–394.)

Another study taking a similar approach was carried out by Martin Lewis (1969, 1972) on the widespread European species *Geranium sanguineum* (Geraniaceae). This species grows in a variety of habitats. As Turesson and other researchers had noticed in other species, plants from the driest sites had the most highly dissected leaves with the narrowest lobes, while those from moister sites had larger, less dissected leaves (Figure 9.18). Lewis found genetically based differences in leaf shape and size among plants collected from a large number of different sites in Europe. There was a cline in the degree of leaf dissection, with the most highly dissected leaves from the driest environments and the more compact leaves from moister sites closer to the ocean. A different geographic cline in leaf size was correlated with the openness of the habitat, with the largest-leaved plants from the least open, most tree-shaded habitats. Based on physiological and energy budget studies, Lewis hypothesized that differences in leaf energy budgets due to these differences in leaf size and shape would lead to differences in leaf temperatures. The more dissected leaves from the drier habitats would remain closer to air temperature, thus avoiding overheating when drought forced stomatal closure. Other physiological advantages were also predicted for leaves of each ecotype in its "home" environment.

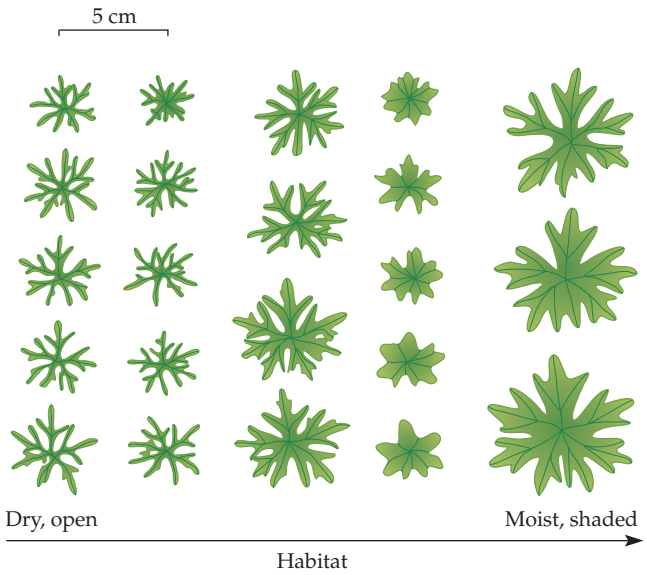



**Figure 9.18** Differences in leaf shape among *Geranium sanguineum* collected from sites in a variety of habitats from northern to central and eastern Europe, then grown in a common garden. The sites ranged from dry, open sites (alvar, xeric limestone, left) to intermediate sites (steppe, mixed steppe–woodland, and coastal cliff tops) to moist, shaded sites (woodland, right). (After M. C. Lewis. 1972. *Sci Prog* 60: 25–51.)

Originally, it was believed that there was a strict dichotomy between clines and ecotypes. Clines were thought to represent gradual differences in genetically based traits that are responsible for gradual morphological and physiological differences over geographic space. Ecotypes, in contrast, were thought to represent sharp differences among populations that are distinct genetically and in various phenological traits. Most plant ecologists no longer see this distinction as being very useful. Genetic variation responsible for morphological, physiological, and phenological differences may occur over many different scales, sometimes with sharp distinctions among populations and sometimes with gradual change. The abruptness of the change depends on the nature of the evolutionary processes acting on these groups of plants: selection, genetic drift, and migration all shape the rate of genetic change over space. We may choose to call distinct populations "ecotypes" for convenience, or we may wish to emphasize the gradual nature of change and call the gradient a "cline." As natural habitats grow ever smaller and more fragmented, there will be more and more cases of gradual genetic and phenotypic change in plant and animal populations, which will become isolated, thus resembling ecotypes, even though they originated as members of a continuously varying cline.

## 9.8 Natural Selection Can Cause the Origin of New Species

One outcome of evolution by natural selection is **speciation**, the production of new species. Speciation may occur when two or more populations of the same species become isolated from each other and adapt to different environmental conditions over a long time. Eventually, the populations' responses to differences in natural selection in their disparate environments can result in populations that have become so different that they are reproductively isolated, meaning that they can no longer interbreed. Most biologists consider **reproductive isolation** to signify that two populations have become different species—which can also occur through processes other than natural selection. Usually, such differentiation happens in populations that are geographically distant from each other (**allopatric speciation**) (Figure 9.19), but it can occur in adjacent populations (**parapatric speciation**), or even within a single population (**sympatric speciation**). These three modes of speciation are common to both animals and plants; however, plants undergo some unique speciation processes, as we will see shortly.

The **biological species concept** defines a species as a group of actually or potentially interbreeding organisms that are reproductively isolated from other such

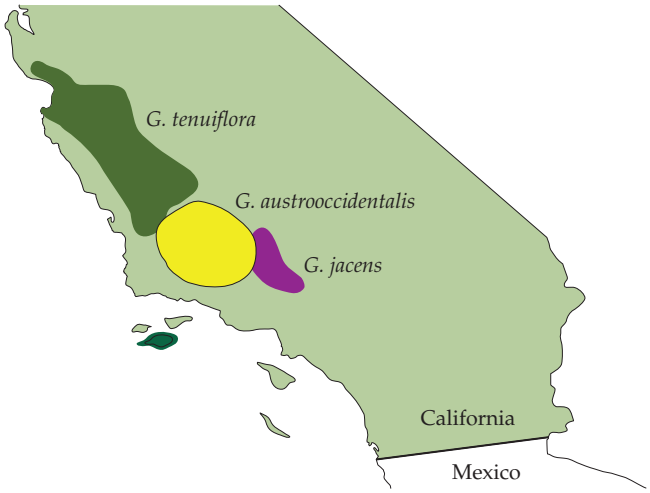



**Figure 9.19** Geographic ranges of three species in the genus *Gilia* (Polemoniaceae). The most likely explanation for this pattern of adjacent ranges is allopatric speciation in an originally widespread species. (After V. Grant. 1971. *Plant Speciation*. Columbia University Press: New York; V. Grant and A. Grant. 1960. *Aliso* 4: 435–481.)

groups (Mayr 1942). A species thus forms an evolutionary unit within which the genetic changes in populations through time are linked. In most cases, species defined in this way are identical to species defined by taxonomists. **Taxonomic species** are generally defined by shared morphological characters or DNA sequences. One problem with using morphological criteria alone to define species is the existence of **cryptic species**, organisms that appear to belong to the same species and yet are reproductively isolated from each other. In plants, cryptic species occur among polyploids (see below). More common are individuals that appear to be different, especially if they come from distant locations, yet are still able to interbreed. If those individuals are connected by others that are morphologically and geographically intermediate, they would probably be classified as members of the same species.

The reproductive barriers between species are not always impermeable. **Hybridization** occurs when members of different species in the same genus mate and produce viable offspring (Figure 9.20). (*Hybridization* has a different meaning to agronomists, who use the same term to refer to crosses between different strains of a single species.) The extent of hybridization varies greatly among genera. In most genera, hybridization is a very rare event, if it occurs at all. On the other hand, some genera are notorious for hybridizing frequently, including *Quercus* (oaks, Fagaceae), *Crataegus* (hawthorns, Rosaceae), and *Atriplex* (saltbush, Amaranthaceae). In some cases, species identities become very difficult to discern as the boundaries between species become indistinct, and we speak of the group of hybridizing species as a **hybrid swarm**. Hybridization appears to be much more common in plants than in animals. This difference may be due in part to the more passive nature of mating in plants, which must rely on wind, water, or animals to transport pollen from one flower to another. Hybrid swarms tend to be more common in wind-pollinated genera.

Another feature of evolution common to plants is **polyploidy**: the duplication of the entire set of chromosomes, resulting in two or more copies of the genome in each cell. Polyploidy can come about in two ways. First, consider a diploid ($2n$) individual that, by some error, produces pollen and ovules without meiosis (the reduction division), resulting in diploid gametes with $2n$ chromosomes instead of haploid gametes with $1n$ chromosomes. If that individual self-pollinates, its offspring will have four copies of each chromosome instead of the usual two; these offspring will be **tetraploid**, with $4n$ chromosomes. Such doubling does not require self-pollination, however; if two separate individuals both produce unreduced, diploid gametes and mate with each other, a tetraploid offspring will result. This process is referred to as **allopolyploidy** if the gametes come from individuals of different species in conjunction with hybridization. It is referred to as **autopolyploidy** if both gametes come from individuals of the same species.

Polyploidy does not have to stop at $4n$; further doublings can take place. Mating with $2n$ individuals can result in $6n$ offspring. In one species of saltbush, *Atriplex canescens*, individuals ranging from $2n$ to $20n$ are known. Ferns, which are an ancient lineage, often have large numbers of chromosomes. The *Polypodium vulgare* (Polypodiaceae) group has individuals with $2n$ = 74, 148, and 222. The high base number (74) is undoubtedly due to much older polyploidy events. The highest chromosome number known in any plant occurs in the fern *Ophioglossum reticulatum* (Ophioglossaceae), in which $2n$ = 1260. The modular developmental system of plants, at least in some taxa, is apparently able to handle these ploidy changes without difficulty. A typical consequence of polyploidy is an individual that is larger or more vigorous than its parents, but otherwise similar in appearance. The larger size comes about in part because cell size is proportional to nucleus size; more chromosomes result in a larger nucleus and, thus, a larger cell. There can also be changes in life history traits, including phenological traits (see Chapter 7).

Individuals with odd ploidy numbers (e.g., $3n$ or $5n$) are also possible. Such individuals cannot produce gametes through typical meiotic processes because the chromosomes cannot pair properly. Through a process known as agamospermy, some plants can produce seeds without fertilization or meiosis (see Chapter 6); dandelions (*Taraxacum officinale*, Asteraceae) and grapefruit (*Citrus*

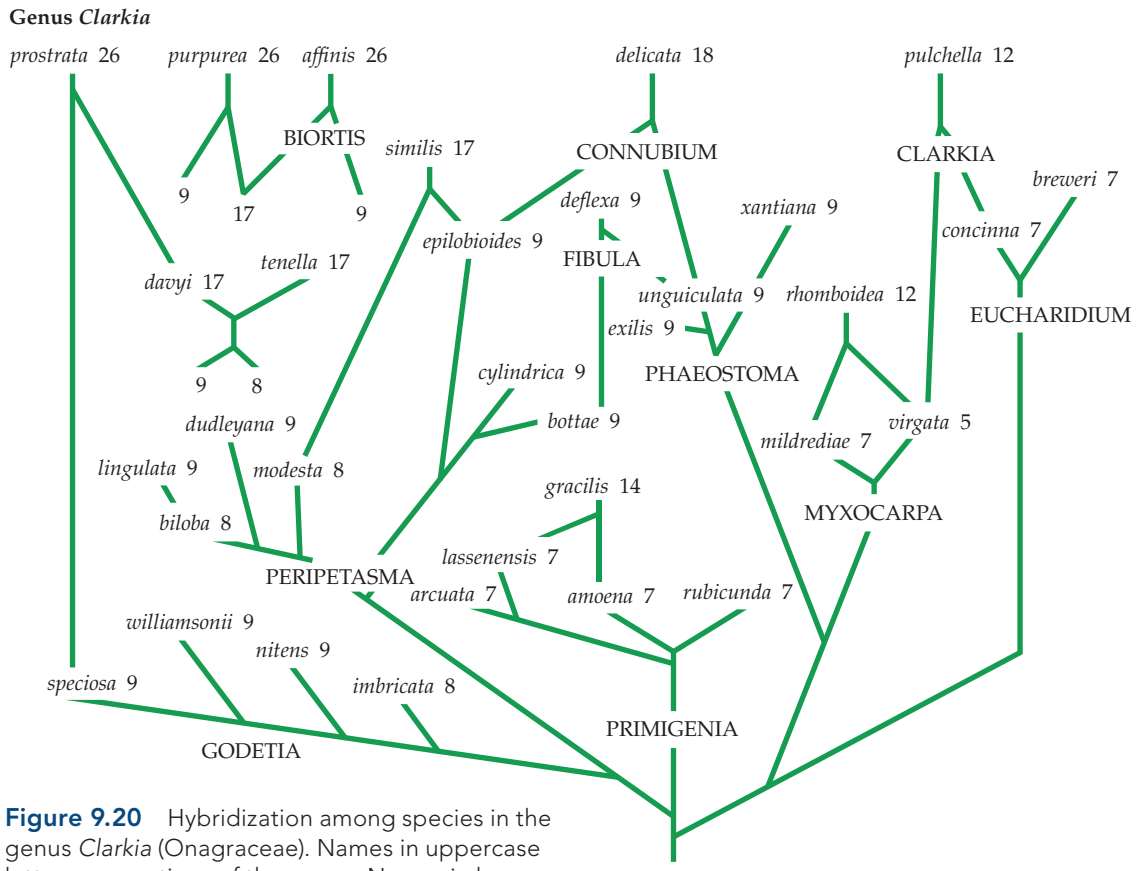



**Figure 9.20** Hybridization among species in the genus *Clarkia* (Onagraceae). Names in uppercase letters are sections of the genus. Names in lowercase letters are species, given with their haploid chromosome number. Polyploid forms have arisen from hybridization between species. (After H. Lewis and M. Lewis. 1955. *Univ Calif Publ Bot* 20: 241–392.)

*paradisi*, Rutaceae) are two such plants. (Sometimes pollination is required to initiate seed development, even though fertilization does not occur.) Seeds produced by agamospermy contain an exact duplicate of the parent's genome. Agamospermy is not restricted to species with odd ploidy numbers. Species with odd ploidy numbers can also evolve to have even ploidy numbers through chromosome doubling or hybridization.

**Aneuploidy**, the gain and loss of individual chromosomes, is another common way in which plant lineages change their chromosome numbers. Aneuploidy is especially common in lineages that are already polyploid. In the genus *Hesperis* (Brassicaceae), for example, species are known in which $n = 7$, 14, 13, and 12. In this instance, the most likely evolutionary pathway was first a doubling of the number of chromosomes from 7 to 14, then a subsequent loss of first one, then a second chromosome.

One result of polyploidy is instant reproductive isolation. A $4n$ offspring cannot mate with its $2n$ progenitors, except by a new, rare hybridization event. Thus, under the biological species concept, such a polyploid individual is a member of a new species. In order for that species to become established, however, the polyploid individual must reproduce, and the population must grow. It is only because most plants produce both male and female gametes and are self-compatible that such individuals are likely to reproduce. *Atriplex* provides an interesting exception; as mentioned above, this genus contains many hybrids and polyploids, yet it is dioecious (male gametes and female gametes are produced on different individuals). Despite the reproductive isolation caused by polyploidy, taxonomists tend to classify individuals with different ploidy numbers as members of the same species if their overall appearance is similar—an instance of taxonomic species differing from biological species.

## 9.9 Adaptation and Speciation Can Happen through Hybridization

Hybridization can do more than create reproductive isolation. It can also lead to new genetic combinations that are adapted to different environmental conditions than those of either of the parental species. A constraint on speciation through hybridization is that the new hybrid must compete with the parental species. Because the hybrid initially will be rare, most matings will be with the parental species. If the hybrid is polyploid, these matings will result in the production of no offspring or sterile offspring (Levin 2002). If the hybrid is diploid, the offspring may be fertile but are likely to be genetically swamped by the parental species and fail to lead to a

new species. Ecological or spatial isolation of the hybrid can provide an escape from this trap.

Loren Rieseberg and colleagues (2003) studied hybridization in wild sunflowers, *Helianthus* (Asteraceae; see Figure 7.4). *H. annuus* and *H. petiolarus* are both widespread in North America, the former inhabiting clay-based soils that are wet in spring and dry out later in summer, while the latter is found in sandier soils with less vegetative cover. The two species hybridized several times between 60,000 and 200,000 years ago, producing *H. anomalus*, *H. deserticola*, and *H. paradoxus*, all diploids. In contrast to the parental species, the hybrids have much smaller geographic distributions and occupy much more extreme environments: *H. anomalus* occurs on sand dunes in Utah and northern Arizona, *H. deserticola* is found on dry, sandy soils in the Nevada desert, and *H. paradoxus* is limited to a few brackish salt marshes in western Texas and New Mexico.

How is it that the hybrids are able to inhabit more extreme environments then the parental species? You might expect that the hybrids would be intermediate in phenotype (see Box 9A). In this instance, the genes of the two parental species had diverged sufficiently that new dominance and epistatic interactions in the hybrids created the extreme phenotypes.

Two approaches were used to demonstrate that the new hybrid phenotypes were adaptive. First, the traits of the hybrids were compared with those of other species that live in the same or similar habitats. *H. anomalus* has large seeds, rapid root growth, and succulent leaves, like other species found on sand dunes. *H. deserticola* flowers rapidly and has narrow leaves, like other desert plants (see Chapter 3). *H. paradoxus*, like other halophytes, can reduce the effects of sodium and other mineral ions through active exclusion, internal sequestration, and increased leaf succulence (Welch and Rieseberg 2002; Lexer et al. 2003). Second, the parental and hybrid species were grown in a common garden in the habitat of each of the hybrids. As expected, each hybrid was favored in its own habitat. The researchers examined selection on genes associated with the hybrid traits and found that those genes were selected for in the expected habitats.

One limitation to understanding the process of adaptation is the length of time since the events of interest occurred. So how did Rieseberg and associates know that the extreme phenotypes originated with hybridization, rather than evolving through natural selection afterward? They re-created the hybrids by crossing the parental species and found that their synthetic hybrids had phenotypes similar to the natural hybrids. While not proving that the adaptive gene combinations of the hybrids arose during the hybridization event, these results provide strong indirect support for that hypothesis. Thus, Rieseberg's work shows the potential of hybridization to bring about new genetic combinations leading to adaptation to new ecological conditions and speciation. Yet to be determined is how common this process is in the evolution of new species.

## Summary

- Natural selection is the central evolutionary process determining the forms of living organisms.
- It comes about through the interaction of three components: phenotypic variation, fitness differences linked to that variation, and a genetic basis for that variation. All three components are necessary for evolution by natural selection to occur.
- The ecology of a plant plays a role in all three components, influencing the phenotype by affecting the course of growth and development, determining fitness differences, and influencing the expression of genes.
- Natural selection acts on variation among phenotypes; almost all plant traits vary among individuals within populations and species.
- Heritability of variation means that at least some of the difference among individuals in a population is due to genetic differences.
- Heritability is measured by studying resemblances among relatives.
- Heritability values depend on gene frequencies within populations and genotype-environment interactions, and they can be affected by the nonrandom distribution of genotypes, and thus, heritability estimates are specific to the population and environment in which they are measured.
- Natural selection has shaped the characteristics and nature of the plants that we see around us, even though that process occurred in the distant past or is occurring now at rates too slow to measure directly.
- Mutation and migration are important for supplying genetic variation.
- Genetic drift can act in two ways, decreasing genetic variation in a population and causing evolution in the absence of natural selection.

# PART III Population Interactions and Communities

Courtesy of G. Fox


**10** Competition and Other Plant Interactions 261

**11** Herbivory and Other Trophic Interactions 297

**12** Community Diversity and Structure 333

**13** Community Dynamics and Succession 371

**14** Local Abundance, Diversity, and Rarity 397

Courtesy of G. Fox

10

# Competition and Other Plant Interactions

Scientists have been studying plant competition for well over 150 years, and it has been considered to be a key factor affecting almost every aspect of plant existence. Competition was an essential part of Charles Darwin's argument for evolution by natural selection (Figure 10.1). In *The Origin of Species* (1859), Darwin wrote:

> *Every being, which during its natural lifetime produces several eggs or seeds, must suffer destruction during some period of its life, and during some season or occasional year, otherwise, on the principle of geometrical increase, its numbers would quickly become so inordinately great that no country could support the product. Hence, as more individuals are produced than can possibly survive, there must in every case be a struggle for existence, either one individual with another of the same species, or with the individuals of distinct species, or with the physical conditions of life.… As the mistletoe is disseminated by birds, its existence depends on birds; and it may metaphorically be said to struggle with other fruit-bearing plants, in order to tempt birds to devour and thus disseminate its seeds rather than those of other plants.*

In addition to competition, the outcome of the "struggle for existence" can depend on other factors, including the adaptations of plants to their abiotic environments (see Chapter 2, Chapter 3, and Chapter 4), interactions with herbivores (see Chapter 11), and chance (see Chapter 8), and competition can interact with all of those factors. Darwin was primarily concerned with the role of competition in setting the stage for natural selection. But competition also has many ecological effects. In fact, as natural selection shapes adaptations, it also determines many of the ways in which plants compete.

Above: *Eqisetum* (horsetail, Equisetaceae) seedlings in a clearing near Cape Mears, Oregon, U.S.A.

Courtesy of S. Scheiner

**Figure 10.1** Charles Darwin studied plant competition in a plot beside his home in the village of Downe in the London Borough of Bromley. He used methods that would be very familiar to plant ecologists today, simply putting small wooden stakes into the ground next to each plant that he wanted to track.

**Competition** can be defined as a reduction in fitness of competing individuals due to shared use of a limiting resource. At the population level, this translates to a reduction in population growth rates: when only intraspecific competition is involved, the growth rate of that population is reduced as compared with a population of the same size but not experiencing competition. When two or more populations are competing for limiting resources, the growth rates of all competing populations are reduced. Competition is often inferred as the reason that populations do not increase exponentially. In agriculture, competition is generally defined in terms of a reduction in crop yield; while this is analogous to the fitness and population growth definitions, it differs from them because yield is not related to fitness or population growth in any simple way. Competition can affect individual plants at all life stages, and its effects can have a major impact on populations, communities, and landscapes, as well as on species' distributions and abundances at even larger scales (see Chapter 19). Competition can occur between individuals of the same species, or of different species.

Ecologists have asked many questions about competition: What are the mechanisms by which plants interact with one another? What determines the outcome of competition among different individuals? Is competition characteristically more severe in some environments than in others? How important is competition, relative to other processes, in shaping community structure? How does competition affect the survival of individual plants, and their sizes? What resources do plants compete for, and are competitive interactions different for different resources?

While the study of competition has a long history, more recently scientists have begun to recognize that plants' interactions with one another form a spectrum from negative to neutral to positive interactions. Sometimes the same individuals interact negatively at one time and positively at another time. Some scientists have argued that **allelopathy**—negative interactions modulated by chemicals excreted by or leached from a plant into its environment that result in negative effects on neighboring plants—may be important, although the strength of the evidence for this has been much debated. At the other end of the spectrum, **facilitation**—positive interaction among plants—is increasingly attracting the attention of researchers, and we consider some of the results of that research here as well.

Ecologists who study animal populations have argued a great deal about the importance of competition in determining population structure, abundance, and the composition of communities. Plant ecologists, on the other hand, have generally accepted that the effects of competition are obvious and pervasive. Weeds frustrate gardeners, and every year farmers spend hundreds of millions of dollars on herbicides to reduce the effects of competition on crop plant productivity. However, because its effects are so complex, plant ecologists have debated just about everything else about plant competition, from how to define and measure it to when and where it is important.

We begin this chapter by examining what is known about how plants compete with one another and how competition affects the survival and growth of individual plants within a population. After reviewing methods for studying species interactions, we look at the effects of competition and facilitation on populations, population distributions, and community composition. It is these aspects of competition that have been the subject of the greatest controversy. We examine some of these debates and take a look at the evidence ecologists have gathered on the role and importance of competition.

## 10.1 Individuals Compete for Limited Resources

How do individual plants compete, and why does one individual succeed in competition while another fails? Ecologists understand a lot about the mechanisms of competition between individuals. Intraspecific competition and interspecific competition have often been treated separately in textbooks and other discussions. We do not separate them, because many of the ways that competition affects individuals, populations, and

communities are not qualitatively different if the competitors are of the same or another species. Questions about competitive outcomes are important because competitors can reduce a plant's biomass and growth rate and decrease its ability to survive and reproduce. The number of seeds produced by a plant is highly correlated with that plant's mass, so successful competitors that accumulate more mass often have more resources to put into reproduction.

Plant growth is highly plastic (see Chapter 6), and the mass, height, number of leaves, and reproductive output of an individual plant can vary over orders of magnitude depending on growth conditions. When plants are grown without close neighbors, they are generally much larger and produce more offspring than similar individuals surrounded closely by others, and often they have a very different morphology (see Figure 9.7).

### What are the mechanisms of resource competition?

Young plants usually have the highest mortality rates, so competition may often have its greatest effects in early life history stages. As a consequence, there are more studies of competition among seedlings than among other stages. Ecologists also study seedlings because these studies are easier to carry out and can more quickly lead to results. But competition also occurs at other life history stages and affects more than early survival and growth. Mature plants can be overtopped and overgrown by surrounding plants, leading to reductions in growth, reproduction, and ultimately, survival. This is probably a common mechanism during succession from old fields to forests, for example, in which shorter herbaceous plants are replaced by shrubs and eventually by trees (see Chapter 13). The root systems of different plants compete for water, nitrogen, and other nutrients. Plants may also compete for pollinators or seed dispersers (see Chapter 6).

We said above that individuals compete for limiting resources. A resource is a substance (or object) that is required for growth, maintenance, or reproduction. It must also be something that can be consumed, in the sense that if one individual acquires some of that resource, it is no longer available to others. For example, nitrogen, water, and phosphorus are all needed for plants to persist, and they are certainly consumable—so they are resources. On the other hand, temperature and soil pH are not resources: a given plant requires them to be within some range, but neither temperature nor pH can be consumed by other plants. Pollinators (see Chapter 6) are not usually consumed by plants, but their time and attention certainly can be! Sunlight is tricky—plants certainly do not deplete the sun's output of energy, but the amount of light on a flat surface (the surface of the Earth on which plants are rooted) is in limited supply and is depleted by plants. For competition to occur, the resource must also be limiting to population growth. Acquiring more water, for example, cannot affect population growth for most aquatic plants because the supply greatly exceeds the need.

What are the resources for which plants compete, and how does the nature of those resources affect the nature or outcome of competitive interactions? Plants compete for light, for water and mineral nutrients from the soil, for space to grow and acquire resources, and for access to mates. The needs of plants of different species for light, water, and basic mineral nutrients are relatively similar (see Chapter 4), in contrast to the resources needed by animals, which may differ far more widely. Unlike resource competition among mobile animals, most of the competitive interactions experienced by a plant occur among near neighbors. Resource competition is local, but the scale of what we mean by *local* depends on the resources in question. The shade of a neighbor reduces the plant's ability to photosynthesize, and the roots of the plants immediately surrounding it may absorb water and nitrogen that it needs to function. Plants even a short distance away may have no effect on it at all. Consequently, the density experienced by a plant is largely the density of the plants in the immediate patch around that individual, and the average density of plants in the surrounding field or forest may be essentially irrelevant to the degree of crowding it actually experiences (Figure 10.2). The effects of neighbors generally decrease sharply with distance. One exception is competition for the attention of animals that carry pollen and disperse seeds (as Darwin noted); plants may be competing for their visits with other plants some distance away.

To understand how plants compete for a resource, one must know how the resource is supplied and also how plants acquire it. The amount of water available, for example, depends among other things on the recent amount of precipitation, the depth of the soil, the soil texture, weather, and the abundance and activities of other plants, and therefore it is affected by season, vegetation cover, plant physiognomy, and other factors. Water becomes available in pulses, and the duration of the interval of water stress between those pulses varies with climate, season, and topography as well as biotic factors. Nitrate nitrogen is largely dissolved in the soil water, and it moves with the soil water into the roots or is leached and becomes unavailable as water drains or runs off from the part of the soil in which roots are present. Phosphorus, in contrast, is fairly immobile in the soil and is available only to roots in the immediate vicinity of the nutrient ions. Soil resources also depend on the presence and activity of mycorrhizae, bacteria, and other soil organisms on or in the vicinity of plant roots.

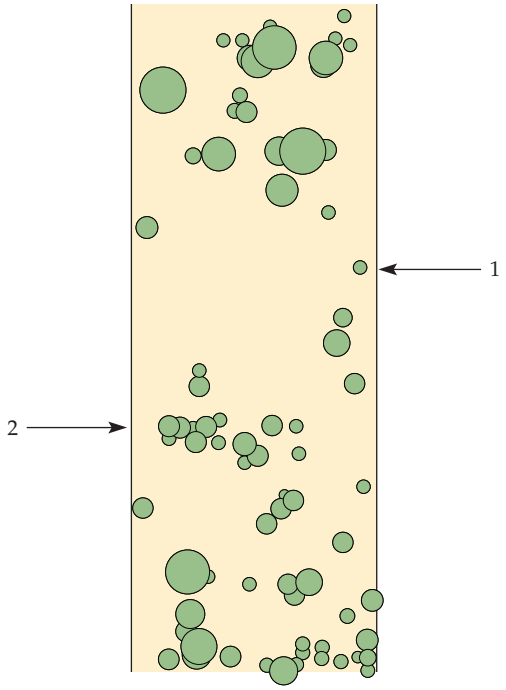



**Figure 10.2** A transect through a natural population of *Myosotis micrantha*. Each circle represents an individual plant; the size of the circle shows the relative plant size. Plants in the same population may experience relatively low densities of neighbors, as in the part of the transect shown at 1, or high densities of neighbors, as in the part of the transect shown at 2. The *x* and *y* axes are not on the same scale: the distance along the vertical dimension from region 1 to region 2 is about 35 cm; the width of the transect is 20 cm. (Data from C. Wilson and J. Gurevitch. 1995. *J Veg Sci* 6: 847–852)

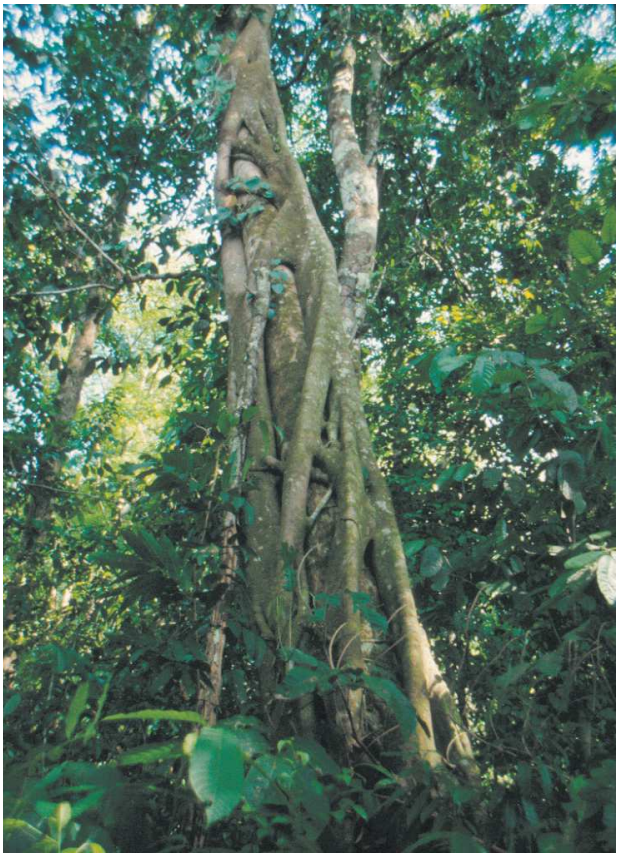
Plants can respond to the patchiness of available resources by adding new growth in the areas where resources are available. Forest trees respond to the higher levels of light created by gaps in the canopy by branch growth into the gaps. Young saplings in a newly created gap grow rapidly in height. Some individuals and some species are much better at responding to gaps and other patches of high resource availability than others. Similarly, dense masses of roots may develop in nutrient-rich patches (Robinson et al. 1999). Such root proliferation can maximize the amounts of nutrients available to the plant and may be an important component of competitive success. Root proliferation may allow a plant to monopolize nitrogen supplies when plants are competing on poor soils and where nitrogen is patchy (which is more likely when it occurs in less mobile forms than nitrate). Roots of different species are known to differ in nutrient uptake rates, at least under laboratory conditions, and these differences may affect their competitive abilities.

Light is in some ways the most peculiar resource for which plants compete. One plant can shade others, though above the canopy the amount of light is undiminished. Plant canopies can very effectively mop up photons, reducing light at ground level to less than 1% of the incident light in many communities. It can be very dark not only for saplings growing in a dense forest understory, but also for seedlings emerging in a grassy meadow and shaded by taller herbaceous plants. An important characteristic of light competition is that it is largely directional: for a given plant, a tall neighbor to its south has a much different effect than a neighbor of the same size to its north. Much of the energy supplied by sunlight comes from a point source (although shortwave reflection and scattering [see Chapter 2] are important in some habitats and under some weather conditions), while soil resources are more likely to be available in three dimensions. The space available to most plants—the plane at ground surface—is essentially two-dimensional. Competition for light is therefore quite different from competition for soil resources.

A dramatic example of competition for light is the overgrowth of trees by vines. In the tropics and subtropics, strangler figs (*Ficus*, Moraceae) grow on trees, encircling their hosts; initially the hosts are protected from high winds by the srength of the enveloping structure, but eventually the hosts are killed by the strangler figs (Figure 10.3A). In temperate forests in the eastern United States, trees along forest edges or gaps are frequently overgrown by either native or invasive vines (Figure 10.3B). Sometimes vines cloak and weaken otherwise healthy trees, but sometimes they overwhelm only trees previously weakened by other factors, such as insects or disease. It is not fully understood whether trees have mechanisms by which they can defend themselves from vines (such as chemical means or shedding of bark or limbs). When trees are overburdened by the weight of vines, they become much more likely to be toppled in windstorms and killed. This vulnerability probably results both from the weakening of the root system as a consequence of loss of photosynthate as the vines increasingly shade the leaves and from the extra weight of the vines.

One general approach to thinking about mechanisms of plant competition is to consider the effects plants have on resources. The extent of resource depletion can be used as a measure of the competitive effect of one species on others (MacArthur 1972; Armstrong and McGehee 1980). The quantity $\boldsymbol{R^*}$ is defined as the average amount or concentration of a resource remaining in the environment after a population of a single species grown alone has used all that it can take up. Robert MacArthur (1972) first hypothesized that the outcome of competition among animal populations should be determined by the "$R^*$ rule": over the long term, in a constant environment, the species with the lowest $R^*$ is predicted to competitively displace all other species. David Tilman

(A)



© Simon Haigh/Shutterstock.com

(B)

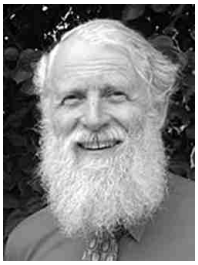
Courtesy of J. H. Miller, USDA Forest Service

**Figure 10.3** (A) A strangler fig (*Ficus* sp., Moraceae), a tropical plant that begins life as several vines that encircle a tree, kill the tree that they grow on, and subsequently merge to become a tree. (B) *Pueraria lobata* (kudzu, Fabaceae), originally from east Asia, is invasive in the southeastern United States, where it can overgrow trees and shrubs.

(1980, 1985) later expanded the application of the $R^*$ rule in the resource-ratio hypothesis and more explicitly focused on its implications for plant competition. Appealing though this idea seems, the $R^*$ rule has been subjected to only a few experimental tests (reviewed by Grover 1997 and Miller et al. 2005). Later work was still dominated by studies that do not directly test the predictions of the theory. Most of the experimental studies were conducted in lab and other controlled environment studies. We will return to Tilman's perspectives and predictions on plant competition when we discuss trade-offs and strategies later in this chapter.

Joseph Craine and collaborators (2005) argued that what really determines competition for soil nutrients, and thus limits plant growth in nature, is not $R^*$, the concentration of nutrients in the soil at equilibrium, but the rate at which nutrients are able to reach the roots by diffusion. In models based on competition by diffusion limitation, these researchers found that the mechanism of nutrient competition depended on the rate of diffusion of nutrients to the root surface, and preemption of that nutrient supply is what leads to competitive dominance—not the $R^*$ rule, reduction of nutrient concentrations in the soil solution. Preemption occurs by a plant increasing the zone of its own resource capture at the expense of the neighbor with which it is competing.

Beyond direct competition for resources, another means by which plants can compete is by "getting there first"—arriving in a newly available microhabitat before other **propagules** or diaspores (seeds, vegetative reproductive units, or other dispersal units) and being able to hold that space against later arrivals. Peter Grubb (1977) hypothesized that differences among species in the conditions and circumstances required for germination and establishment, which he called the species' **regeneration niche**, might be an important factor in species coexistence. He argued that this is one of the means by which species with superior dispersal capabilities dominate some habitats to the exclusion of other species. We discuss disturbance, colonization, and succession further in Chapter 13.



Courtesy of Peter Grubb

Peter Grubb

One valuable insight is that we can summarize plants' competitive abilities as having two components: the **competitive effect** of a plant on its neighbors, and the **competitive response** of a plant to its neighbors. This distinction, first made by Deborah Goldberg and Patricia Werner (1983), can help us to understand the ways in which plants interact. Surprisingly, the rankings of competitive effect and competitive response may be uncorrelated—a species may be an excellent response competitor, but a poor effect competitor (Goldberg 1990).

Courtesy of Deborah Goldberg

Deborah Goldberg

## Resource competition often depends on plant size

In competition, bigger is usually better. Larger plants generally have a competitive advantage over smaller ones; they can affect their smaller neighbors a great deal but be affected by them very little. Taller plants shade shorter neighbors, and plants with more extensive root systems can garner more water and nitrogen from the soil. Larger plants generally take up more resources, produce disproportionately more flowers, attract more pollinators, and set more seed than smaller neighbors.

But the story is not so simple: size does not always play a role in plant competition, and the role of size in determining competitive outcomes may take some unraveling to figure out. Plants of extremely different sizes or ages may not be competing at all, or they may be competing in different ways than individuals that are closer in size or in age. A mature forest tree usually does not compete directly with an herbaceous understory plant growing beneath it. Understory plants are adapted to grow in the shade, and the roots of the tree and the herb are often found at very different depths in the soil and draw on different supplies of soil resources. So can we assume that plants of very different growth forms never compete? Not so. Some herbaceous perennials in the forest understory have been observed to flower only when a canopy tree falls, creating a gap. A small sapling growing beneath a canopy tree may have no measurable competitive effect on the mature tree, but the shade of the tree may completely prevent the sapling from growing larger. Bring down the big tree, and the sapling may rapidly increase in height and biomass. But foresters have also found that removal of understory plants with herbicides can sometimes lead to great increases in the growth rates of mature lumber trees, presumably because there are a great many understory plants, and together they intercept water or reduce soil nutrients available to the trees. Competition between smaller and larger plants is generally highly asymmetric (the larger plant has much greater effects on the smaller and is itself affected less), but this is also not always the case.

Another complication in understanding the effects of plant size on competitive interactions is that larger plants may become disproportionately susceptible to factors that reduce their effectiveness in competition. For example, a plant with a greater leaf surface area may lose more water through transpiration than a similar plant with a smaller leaf area. Consequently, the larger plant may be affected more by a dry period than its smaller neighbors, increasing its likelihood of dying or reducing its future effectiveness in competition.

## Plant competition frequently occurs between seedlings

Mature plants of most species produce a great many seeds, but few of those seeds survive to become mature individuals. (In a stable population, on average, only one seed survives to reproductive maturity over the life-span of a mature plant, but of course some individuals are more successful than others in producing offspring.) Seedlings and seeds are the life history stage most vulnerable to many environmental hazards, from drought and other abiotic effects to predation and competition. Competition among seedlings is often intense, with many seedlings dying or failing to mature as a consequence. Seeds and seedlings have been of great interest to ecologists and evolutionary biologists as the life history stages during which many ecological patterns are established and during which intense natural selection might be operating.

In a classic experiment, C. M. Donald (1951) showed that in **monoculture** (populations or plantings of a single species), each of several species of British annual pasture plants grown in Australia and sown over a wide range of densities had a remarkably consistent total final dry weight (or "yield") in a given area (Figure 10.4). This is called **constant final yield** (Weiner and Freckleton 2010). Whether seeds were planted sparsely or very densely (above a certain minimum density), the total aboveground dry matter at final harvest was constant for each species. The total yield increased when more resources were supplied, but the same relationship held: the average plant size became smaller as density increased. In a reanalysis of these data, Tatuo Kira and his colleagues (1953) showed that as plant density increased, the mean weight of individual plants decreased in a linear fashion when both were expressed on a log scale (Figure 10.5). However, in most natural settings, the mixture of plant species generally means that yield will not be constant. In addition, competition-caused mortality will ultimately limit any quantity like yield.

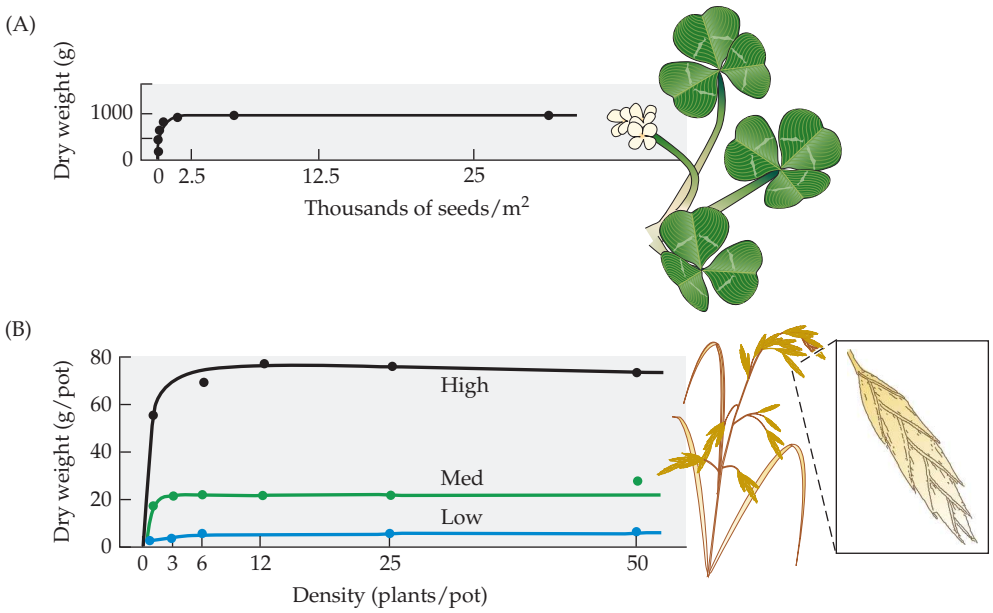



**Figure 10.4** Relationships between yield of dry matter (g) and plant density for two pasture plants. (A) *Trifolium subterraneum* (sub clover, Fabaceae), measured after flowering (density is expressed as thousands of seeds sown/m$^2$). (B) *Bromus unioloides* (Poaceae) at low, medium, and high levels of fertilization with nitrogen (density is expressed as plants/pot). Note that the *x* axis is on a geometric scale in both (A) and (B). (After C. M. Donald. 1951. *Aust J Agric Res* 2: 355–376, with permission from CSIRO Publishing.)

In an influential paper on intraspecific competition in plants, Andrew Watkinson (1980) described the effect of intraspecific density on plant performance as

$$w = w_m(1 + aN)^{-b} \tag{10.1}$$

where $w$ is mean plant weight, $N$ is plant density, $w_m$ is the mean dry weight of an isolated plant at a given time, and $a$ and $b$ are model parameters fit for particular data sets. The density at which yield is first affected by density determines the value of $a$, and $b$ is related to the shape of the density-yield relationship. These relationships have been useful in describing the relationship between density and yield in monocultures of many different species under a range of environmental conditions (Watkinson 1985). This model has been used to describe plant populations in many studies, but it does not provide insight into the underlying mechanisms or population-level consequences.

Mean plant size can be a misleading measurement, however. If a measure (like size) follows a normal distribution, the mean is "meaningful" and tells you a lot about the values in the population. When a measure is far from being normally distributed (that is, when the values don't fit a bell curve), the mean is, well, pretty meaningless as an indication of the individual measurements in that population.

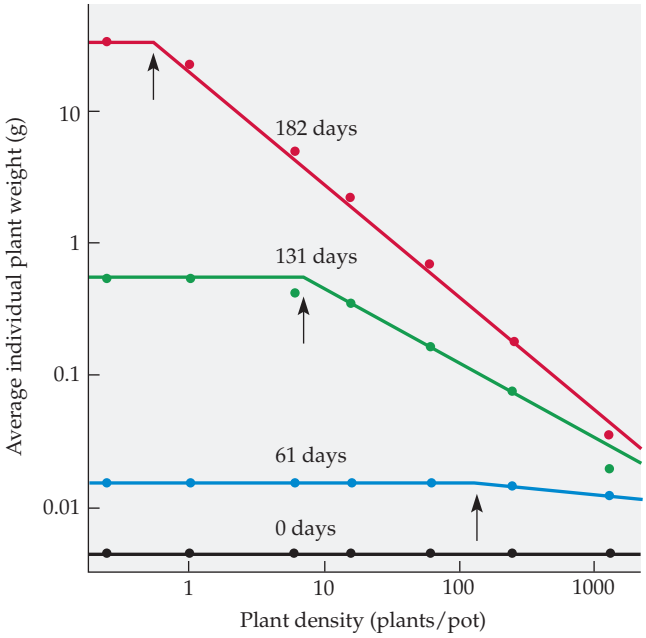



**Figure 10.5** Average individual plant weight (g) for *Trifolium subterraneum* (sub clover, Fabaceae) planted over a range of densities (expressed as number of plants/pot) and harvested at 0, 61, 131, and 182 days after seeds were planted. Note that both axes are on a $\log_{10}$ scale. The arrows show the densities at which plants began to reduce one another's growth at different ages. At early harvests, only very dense plantings showed reductions in average plant weight, but by the last harvest, all but the lowest-density planting demonstrated weight reductions related to density. Decline in weight with density is linear on a log-log scale. (After T. Kira et al. 1953. *J Inst Polytech* Osaka City Univ D 4: 1–16; Data from C. M. Donald. 1951. *Aust J Agric Res* 2: 355–376.)

Size relationships among individuals in even-aged, dense monocultures have been well studied in greenhouses and in a few studies in natural populations. Individual plant sizes are generally extremely uneven in such stands. Typically, a few large individuals dominate the available area, while most individuals remain very small. These highly unequal size distributions are called **size hierarchies** because the sizes form a hierarchy from largest to smallest plants. Size hierarchies have important implications for plant fitness (which is highly unequal among individuals), for population demography (the average survival and fecundity of individuals is fairly meaningless in such a population; demographic analyses taking size into account are needed, see Chapter 8), and genetic drift (since the contribution of individual plants to the next generation is highly unequal, see Chapter 9).

One cause of size hierarchies is hypothesized to be **asymmetric competition** (Weiner 1990; Schwinning and Fox 1995), in which the largest individuals have disproportionate negative effects on their smaller neighbors. It has been suggested that small initial differences in access to light may lead to greater and greater inequality in size over time as competitive effects magnify the initial small differences (Figure 10.6). In an influential review, Susanne Schwinning and Jacob Weiner (1998) examined hypotheses and evidence for the causes of asymmetric competition. It was clear that competition is not always asymmetric. A striking example of this is Aaron Ellison's (1987) study of dense natural populations of *Salicornia europaea*, a succulent salt marsh annual. Experimental thinning of these populations did not affect the degree of size inequality, suggesting that asymmetric competition was not responsible for the size hierarchy. Schwinning and Weiner suggested that asymmetric competition may be unlikely when plants grow taller but not wider when competing (as is true for *Salicornia*) and may also be less likely for clonal plants (de Kroon et al. 1992). More generally, they suggested that competitive asymmetry requires that the resource be "preemptable" and that this is generally more likely to be the case for light than for soil resources.

Size hierarchies can be caused by patchy resources, especially when plants are much smaller than the resource patches. Chester Wilson and Jessica Gurevitch (1995) examined the spatial relationship of plant sizes in a dense natural stand of *Myosotis micrantha* (forget-me-not, Boraginaceae), a small winter annual (Figure 10.7). These plants had extremely unequal sizes; there were large numbers of very small plants, with a steep drop in numbers of plants in the larger categories, and very few of the largest individuals. The researchers hypothesized that if asymmetric competition was the cause, then large individuals should be surrounded by small, suppressed neighbors. They found instead that the opposite was true. Large plants had large immediate neighbors, and small plants were associated with small neighbors. Individual plant mass was also highly correlated with the combined mass of neighbors, so the population formed a mosaic of patches of large plants and patches of small plants. Plants without close neighbors were much larger than plants with neighbors, however, so competition probably did affect plant size. The researchers concluded that asymmetric

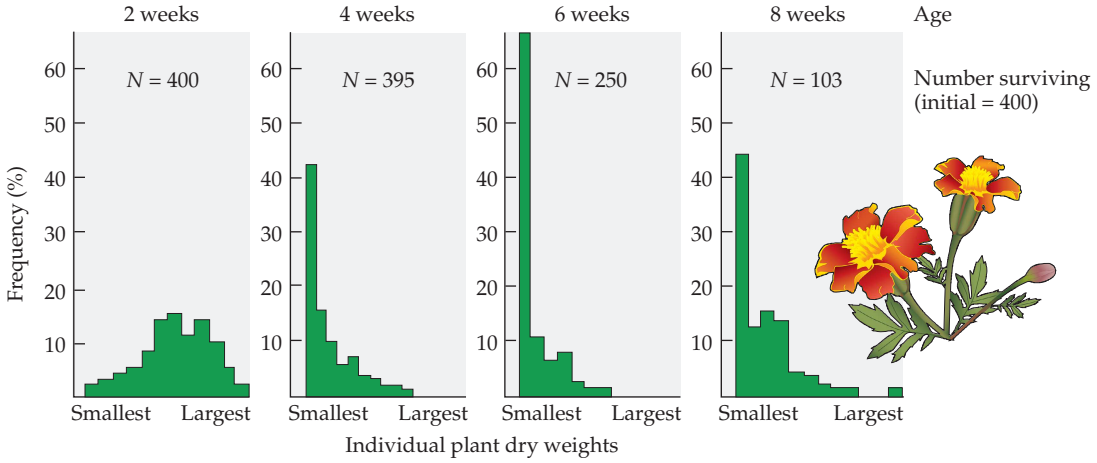



**Figure 10.6** Frequencies of dry weights of individual seedlings of *Tagetes patula* (marigold, Asteraceae), an annual plant, grown in a greenhouse experiment, at 2, 4, 6, and 8 weeks. The number of surviving plants is shown at the top of each graph. At 2 weeks, the distribution of dry weights is close to a normal (bell) curve, but the distribution becomes increasingly unequal (hierarchical) as the population grows older, with many small plants and a small number of very large individuals. Over time, death removes the smallest individuals from the population (self-thinning), so by 8 weeks the population is somewhat less unequal than at 6 weeks. (After E. D. Ford. 1975. *J Ecol* 63: 311–333)

Courtesy of J. Gurevitch

**Figure 10.7** *Myosotis micrantha* (forget-me-not, Boraginaceae), an annual plant, growing in a natural population in southern New York State, U.S.A.

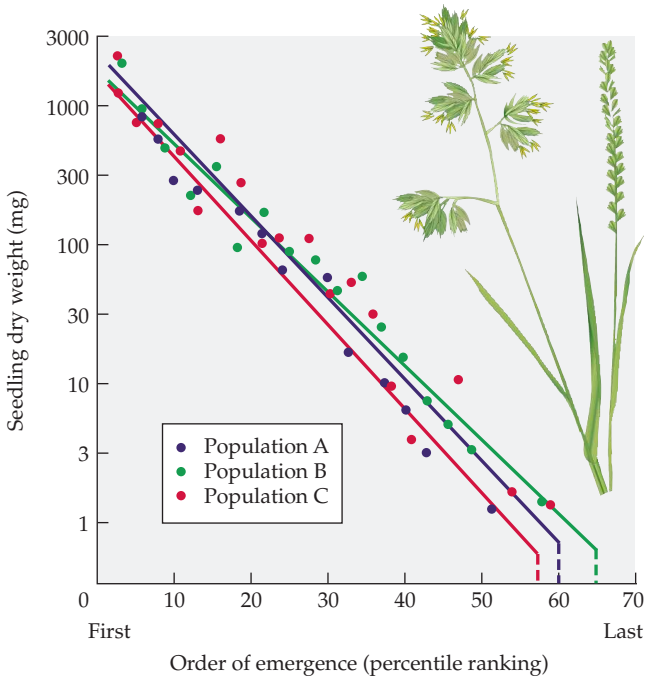



**Figure 10.8** The effect of relative order of emergence (percentile ranking) on seedling dry weight (mg). Each line shows the relationship (the regression) for one of three different populations of *Dactylis glomerata* (orchard grass, Poaceae) in a greenhouse experiment. The consequences of emerging sooner or later than one's neighbors can be enormous: seedlings with the lowest percentile ranking—those that germinated and appeared aboveground first—were more than 1000 times larger than those emerging last. Seedling dry weight (mg) is graphed on a log scale. (After M. A. Ross and J. L. Harper. 1972. *J Ecol* 60: 77–88.)

competition was unlikely to have caused the extreme size hierarchy found in this natural population. Rather, the size hierarchy was probably caused by variation in plant density or patchy resource distribution.

In addition to the factors just discussed, differences in seed size and in the order of emergence (speed of germination) can also contribute to the development of size hierarchies, particularly in annual plants. Two or more of these factors can interact to generate and maintain size hierarchies as well (Miller et al. 1994). A number of studies have quantified the effects of relative order of emergence on plant size. Among a group of seedlings germinating together, a small head start may confer a large advantage. For example, M. Ross and John Harper (1972) showed a strong relationship between the order of emergence and plant size when seeds of the grass *Dactylis glomerata* (orchard grass, Poaceae) were planted at high densities (Figure 10.8). They hypothesized that this result was due to the disproportionate share of resources taken by earlier-emerging individuals. However, earlier germination has disadvantages as well; if it did not, natural selection would lead to an arms race of ever-earlier-germinating seeds (Miller 1987). For example, earlier germination in spring might result in greater likelihood of mortality from late frosts. Thus germination time is an example of a trait that often is subject to stabilizing selection (see Chapter 9).

## Seedling competition can lead to self-thinning

Many experiments on plant competition have been conducted in which seeds of herbaceous plants (usually annuals) are planted in monospecific stands.The seedlings grow until they begin to crowd one another. As the plants grow larger, crowding becomes severe, and eventually some individuals die. This type of density-dependent mortality is known as **self-thinning** (after the gardening and forestry practice of thinning: removing smaller or weaker individuals in overly dense plantings). It is important to realize that the term *self-thinning* does not imply a voluntary, altruistic self-sacrifice of the weaker individuals for the general good (which is certainly not what is happening!). Thus, while plants are highly plastic, this plasticity ultimately has limits—otherwise, mortality would not occur.

Kyoji Yoda and colleagues (1963) proposed what they called the –3/2 thinning law to describe the general relationship between the average mass per individual and the density of survivors:

$$w = cN^{-3/2} \tag{10.2}$$

where $w$ is the average dry weight per plant, $c$ is a constant that differs among species, and $N$ is density. When plotted on a log-log scale, the relationship between mean dry weight and density is linear, with a slope of –3/2 (Figure 10.9). Yoda and his colleagues tentatively explained this relationship as a consequence of the weight of a plant being directly related to its volume (a cubic

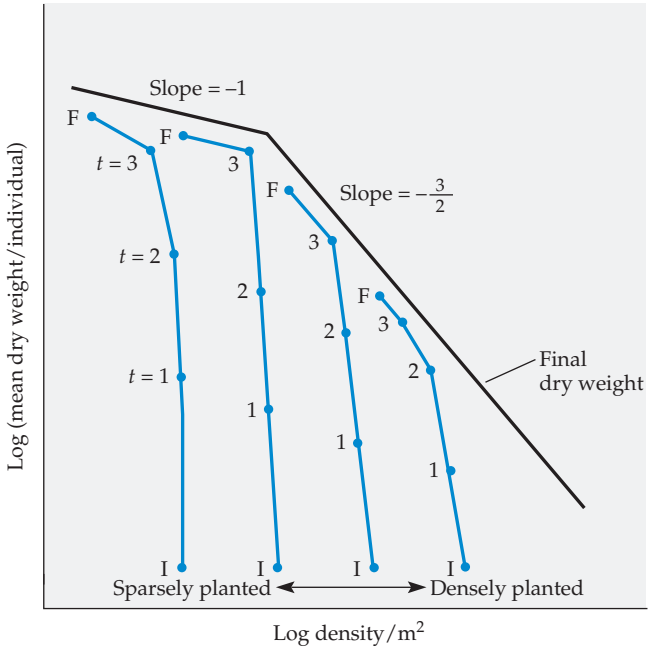



**Figure 10.9** Effects of planting at different densities on the mean dry weights of individuals as seedlings age. Each of the blue lines represents a different initial planting density. The lowest point on each line is the initial dry weight at germination (I), and the highest point is the final dry weight (F), with weights shown at time intervals $t$ = 1, 2, and 3. (After S. Kays and J. L. Harper. 1974. *J Ecol* 62: 97–105.)

measure), and the density of plants being determined by area (a squared term). More densely planted stands begin to experience mortality sooner, and at smaller individual plant sizes, than more sparsely planted stands. Factors that increase the size of individuals, such as greater soil fertility or greater initial seed size, also increase the speed and amount of mortality. The larger the plants are, the fewer can be packed into a given area. So, while it may seem counterintuitive, factors that favor plant growth may also contribute to higher mortality. Indeed, adding fertilizer or water to a patch of plants (whether it occurs naturally or in a garden) may stimulate growth but is also likely to increase mortality.

There is no doubt that self-thinning occurs in crowded stands: some plants die. The **–3/2 thinning law** was an attempt at a simple and general ecological law, but there is considerable evidence that the process is not so regular that it can be described by a single numerical relationship for all populations. J. White (1985) found that the –3/2 thinning law appeared to hold for some forest tree plantations and for certain other plants in even-aged monocultures, but many studies of single populations have been plagued by serious statistical problems (Weller 1987; Lonsdale 1990; Weller 1991). The explanation about plant volume versus surface area requires strong assumptions about plants holding constant shapes as they grow—assumptions known not to hold (Norberg 1988). Jonathan Silvertown and Jon Lovett-Doust (1993) argued that the relationship holds well if the –3/2 slope is considered to be an upper limit. Self-thinning slopes also vary among species, and they may depend partially on plant shape and development.

## 10.2 There Are Several Approaches to Experiments for Studying Competition

Ecologists have used a variety of approaches in controlled environments, gardens, and natural populations to measure the effects of competition and, in some cases, to test hypotheses regarding its effects on individuals, populations, and communities. They have also used a variety of methods to quantify the outcome of these experiments. Understanding the ways the experiments have been conducted and data have been presented is important because these factors can have large effects on the results, sometimes changing studies' conclusions (Grace 1995; Weigelt and Jolliffe 2003).

### How we quantify competition can affect experimental results

One of the most common approaches to quantifying the intensity of competition is to use an index, usually a ratio that standardizes responses across species and environments so that they can be compared on the same scale. A very widely used index is the **relative competition index** or relative competition intensity (RCI):

$$\text{RCI} = \frac{P_{\text{monoculture}} - P_{\text{mixture}}}{P_{\text{monoculture}}} \qquad (10.3)$$

where $P_{\text{monoculture}}$ and $P_{\text{mixture}}$ are plant performance in monoculture (only one species) or in mixture (with two or more species), and performance is usually measured as dry mass or growth rate.

Despite its popularity for measuring competition intensity, RCI has a number of limitations. Ratios expressed on an arithmetic scale have poor statistical properties and are asymmetric (changing the numerator affects the ratio differently than changing the denominator by the same amount). This makes it difficult to interpret! Indices are generally dimensionless; because they have no units, it is not clear how to interpret them in terms of the underlying competitive process or in terms of the population-level consequences. For example, in Equation 10.3 you can see that however $P$ is measured, it is in the same units in the numerator and the denominator, so they cancel out; what does an RCI of a particular value mean?

Another widely used competition index is the absolute difference between performance in a monoculture and mixture, the **absolute competition index** (ACI):

$$\text{ACI} = P_{\text{monoculture}} - P_{\text{mixture}} \tag{10.4}$$

Still another index is the **log response ratio** (LRR):

$$\text{LRR} = \ln\left(\frac{P_{\text{mixture}}}{P_{\text{monoculture}}}\right) \tag{10.5}$$

Which measure should be used? Many different indices have been used in studying plant competition. Alexandra Weigelt and Peter Jolliffe (2003) discuss 50 different indices and the considerations to use in selection and evaluation. Rather than discussing 50 indices, we note that it is questionable whether any one index can accurately represent the essential features of a competitive interaction. Any index simplifies reality and thus has inherent limitations (as do indices expressing any kind of relationship, from species diversity to stock market performance). Any attempt to compress complex data into a single index suffers from a loss of information. In addition, competition indices unrealistically assume linear responses to neighbor density, many different outcomes can result in the same index, and these indices tend to be strongly influenced by initial plant size (Grace et al. 1992). The use of so many different indices makes comparisons among studies difficult. Scientific progress relies in part on standardization, and further understanding of plant competition may be impeded by this proliferation of approaches and measures. Ideally an index of competition should be directly linked to measures of fitness (see Chapter 9) or population growth rate ($\lambda$, see Chapter 8). Overwhelmingly, published studies instead use quantities like biomass, yield, cover, or relative growth rate. While these quantities are components of fitness or population growth rate, and are often closely correlated with them, that is not necessarily true in any specific case. More importantly, perhaps, data on the effects of competition at one life stage do not necessarily provide information about the effects over a plant's lifetime, which is what we really need in order to draw conclusions about the effect of competition on populations. This is not to say that indices are useless, just that one needs to be aware of the limitations of any one index so that they are used in conjunction with other types of information.

In making comparisons of competition intensity across environments or species, an alternative to the use of indices is to examine graphs of performance and test for statistical interactions, comparing performance (e.g., biomass) directly without converting to indices. Better still would be to follow the growth (or other measure of performance) of the competing plants over time, following their trajectories graphically and statistically to determine how the interaction affects them over time.

### Competition experiments were originally conducted in greenhouse or garden environments

Many plant competition experiments, particularly those conducted in the mid to late-twentieth century, have been conducted in greenhouses (called glasshouses in some countries). Greenhouse experiments offer the advantages of relatively controlled conditions and precise manipulation of the factors of interest. Their limitations include the uncertainty in extrapolating results to natural communities, as well as artifacts of greenhouse conditions, including wind and humidity conditions very different from those in nature. An important potential artifact is the large effect that being grown in pots has on plant-water relations and root structure. Although the same kinds of experiments carried out in greenhouses and growth chambers could be carried out in gardens or natural communities with pots set in or above the ground, this has not as often been done, because the goals of greenhouse and field experiments often differ. In fact, there is no strict demarcation between the level of control in greenhouse, garden, and field experiments. For instance, technical advances have extended our ability to manipulate many environmental factors in the field, from soil temperatures to atmospheric $CO_2$.

One issue that sometimes arises is that experimental manipulations may have effects on a scale different from the questions being studied and may therefore have consequences that may be misleading. For example, if we water an experimental plot in an arid setting, we may well cause more plant growth, but (because we are watering only one plot in a large region) doing so may also attract more herbivores. Similarly, adding $CO_2$ tends to increase the ratio of carbon to nitrogen in new plant tissue, which may make the plants less subject to herbivory. Experimental manipulations frequently have some unintended consequences (much as all drugs have some side effects, ranging from very mild to severe), and these must be considered carefully in designing and interpreting experiments (Fox et al. 1995).

Three basic experimental designs have been commonly used in greenhouse competition experiments: substitutive (replacement), additive, and additive series (Gibson et al. 1999). **Substitutive designs** test the relative strength of intraspecific versus interspecific competition by altering the frequencies of two hypothesized competitors while keeping the total density constant (de Wit 1960; Harper 1977). These experiments were once the most important tool for studying plant competition but have been criticized on a number of grounds. They are subject to theoretical and statistical limitations, suffer

from restrictive assumptions, and offer limited ability to extrapolate their results (Connolly 1987). As these problems have become better known, substitutive designs are no longer frequently used.

Simple **additive designs** manipulate the total density of neighbors, usually over a range of densities, while keeping the density of the target species constant (usually a single individual). These experiments have been faulted for confounding density with species proportions and for too often basing their conclusions on a single measure of final yield. They offer various advantages, however, for addressing a number of questions (Gibson et al. 1999). For example, Deborah Goldberg and Keith Landa (1991) used an additive design to compare competitive effects and responses on a per-individual and per-unit-biomass basis. It is useful in such designs to include a "no neighbor" treatment.

**Additive series designs** and "complete additive" designs, also called **response surface experiments**, in which both densities and frequencies are varied, have a number of advantages (Inouye 2001). These designs offer more information than simple additive designs, but they are large and complex. The information gained may not always be worth the additional effort and cost involved. Multispecies competition experiments, likewise, are complex and are done only occasionally.

A few other experimental approaches have been used on occasion in greenhouse competition studies. Experiments in which root and shoot competition are decoupled (Figure 10.10) can be useful in determining the ways in which competitors affect one another more precisely. Fan designs were created to efficiently study the effects of density. Hexagonal designs test the effects of different frequencies of neighbors; hexagonal fans test both density and frequency simultaneously. These designs all consist of plants arranged in precise spatial patterns to provide data on density and spacing effects and on the effects of the proportions of inter- and intraspecific competition. Experiments of this type have the advantage of testing multiple factors simultaneously, but they have been criticized for having statistical drawbacks, including lack of independence, statistical biases, and other limitations (Gibson et al. 1999). Creative replication and variation on such designs might potentially overcome some of these limitations.

Jessica Gurevitch and a group of her students (1990) compared the intensity of intra- and interspecific competition by quantifying the performances of individual plants grown alone (no competition), grown with varying numbers of intraspecific neighbors, and grown with varying numbers of interspecific competitors. Plants were grown in a range of pot sizes to compare resource availability and competitive effects due to plasticity in plant size. The effects of intra- and interspecific competition were assessed by the reduction in growth for plants grown with neighbors in comparison to plants grown without neighbors. Growth with and without neighbors was compared directly rather than by using indices. The researchers argued that it was most informative to evaluate plant responses to competitors by comparing performance in either intra- or interspecific competition with performance in the absence of competition. In the more conventional approach, performance without competitors is not measured, and performances in intraspecific and interspecific competition are assessed using an index such as RCI, judging one against the other in a single ratio. According to Gurevitch and coauthors, the effects of intraspecific and interspecific competition are thus confounded in such experiments.

An unusual garden experiment was carried out by Deborah Goldberg and her colleagues (2001) on the effects of density on desert annual communities in Israel. They collected seed banks from sand dunes in sites

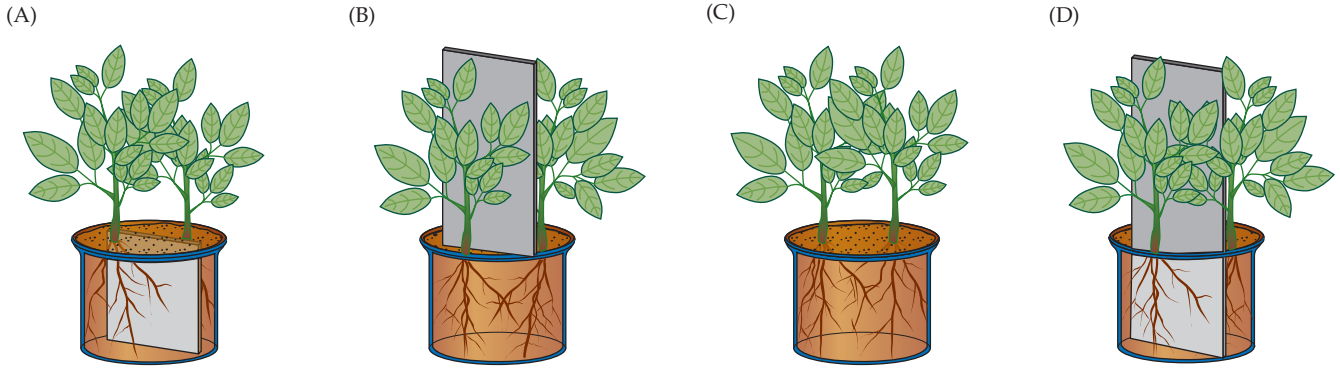



**Figure 10.10** One type of experiment designed to separate root and shoot competition while keeping total soil volume constant. (A) Shoot competition alone (note barrier to root interaction). (B) Root competition alone (note barrier to shoot interaction). (C) Both roots and shoots are allowed to compete. (D) Root and shoot barriers are combined to eliminate both aboveground and belowground interaction.

ranging from mediterranean to desert environments that differed in rainfall, productivity, and diversity. They were able to include entire annual plant communities in their seed bank samples by sieving the sand to collect all the seeds. Then samples of the entire plant community, in the form of seeds, were planted at different densities in a seminatural garden setting. The researchers found strong effects of density on seedling emergence, survival to the end of the growing season, and final plant size. However, its effects at each of these life history stages differed. The largest effects of density were its negative effects on seedling emergence (germination and very early survival). Density had a weak but positive effect on subsequent survival, with the lowest survival in the lowest-density plots (we will explore some possible reasons for this result when we discuss facilitation later in this chapter). It had a negative effect on average final plant size, although this effect was not as strong or as consistent as the negative effect on emergence. This research highlights the importance of studying competition at more than one stage of the life history.

Most competition experiments in natural communities have been concerned with simple questions regarding whether neighbors affect biomass, growth, or some other fitness component in a single location at a particular time. We do not even know whether the most abundant species in a community are typically those that are competitively superior, because we have limited data that address that question, despite its central importance in understanding the role of competition in community structure.

The most common experimental design in natural communities involves the removal of all or some neighbors of a target individual (Goldberg and Scheiner 2001). Removals of different species or different functional groups (e.g., woody versus nonwoody neighbors) may be contrasted. While there is no reason why neighbor densities could not be increased rather than reduced, this has less commonly been done in plant competition experiments under field conditions.

Other experiments manipulate target species abundance over natural gradients of productivity or neighbor density, usually by transplanting individuals into native vegetation. Generally, only the growth of mature plants is measured in such experiments, although a few studies have examined population responses. Both seed additions and adult removals were used by Norma Fowler (1995), for example, to examine the density dependence of demographic responses in two perennial grasses, *Bouteloua rigidiseta* and *Aristida longiseta*, in Texas. She found that the density dependence of demographic parameters for both grasses was weak—competition was less important than other factors for population regulation in these plants. In other experimental approaches, target species may be transplanted into communities that differ in species composition rather than in productivity or density. We will return to field competition experiments later in this chapter when we consider evidence for the effects of competition on community composition and species coexistence in nature.

A long-standing approach to studying the effects of competition in forests has been to manipulate aboveground and belowground competition separately, by cutting whole trees or large branches to create gaps in the canopy and by trenching to remove root competition (Coomes and Grubb 2000). Trenching involves slicing through the soil to a standard depth (often about 50 cm, but sometimes deeper) around small plots to sever root connections to trees and shrubs outside the plot. Usually trenching is used on cleared plots where seedling growth and survival are compared with the same measures in untrenched plots. Sometimes soil is removed from the trench, and barriers (plastic, sheet metal, etc.) are put in place to prevent root regrowth. For example, Ignacio Barberis and Edmund Tanner (2005) combined gap creation with trenching in a tropical semievergreen forest in Panama to assess the effects of aboveground and belowground competition on experimentally planted tree seedlings. They found that gaps greatly increased growth for the seedlings of the four tree species they studied but that trenching resulted in strong increases in growth only in gaps—that is, the tree seedlings did not respond to trenching when they were shaded by an overhead canopy (and one of the four species did not respond to trenching at all).

There have been fewer general studies of competition in recent years than in the past. Many researchers have begun to focus instead on questions about the circumstances under which competition plays an important role, and about how competition and other interactions—including facilitation and soil-plant feedbacks—may interact. We turn to these questions in the next two sections.

## 10.3 Interactions among Species Range from Competition to Facilitation

Plants can interact in a wide range of ways besides competing for resources. There are a number of theories about what makes plants inferior or superior competitors under different conditions. The outcome of competition may be either more or less important than other factors in determining species coexistence. One particular type of negative interaction, allelopathy—"chemical warfare" between plants—has been the subject of a great deal of study and controversy. Many different kinds of positive interactions between plants, known collectively as facilitation, may also influence both individual fitness

and community structure. Finally, we look at the evidence for allelopathy and facilitation and the systems in which they have been studied.

### Different theories attempt to explain how competitive trade-offs lead to strategies

Are there competitive strategies (suites of traits that co-occur and can be predicted to be successful under particular circumstances)? We have already encountered the concepts of trade-offs and strategies in the context of life history traits in Chapter 7. Do certain sets of traits generally make for better competitors, or are there trade-offs that make some traits advantageous in certain kinds of competitive situations but a handicap in others? This is yet another area in which the interpretation of patterns and the predictions made by different ecologists have a long history of diverging sharply.

Early in the twentieth century, John Weaver and Frederic Clements (1929) observed that in prairies, taller grasses had a competitive advantage, and at least some current researchers have also concluded that plant height, at least in herbaceous perennial communities, is generally associated with competitive dominance (Keddy 2001). But traits that confer an advantage to individuals or species competing for light may not necessarily be advantageous in competing for other resources, such as soil nutrients.

Extending ideas first developed by Robert MacArthur (1972), David Tilman (1982) and others have laid out the following view of plant competitive strategies. They assumed that there are different plant traits that confer a competitive advantage on species possessing them under different conditions or environments. The traits also have various costs. In any environment, they hypothesized, the best competitors under the existing conditions will be dominant. For example, tall plants that shade neighbors are superior competitors for light and should be competitively superior in fertile, high-productivity communities. In communities on infertile soils with low productivity, species that are superior competitors for soil nitrogen should be competitively dominant (Tilman 1988). Superior competitive ability in infertile soils depends on the ability to reduce soil nutrients to a level below that at which competitors can exist, and to persist at that low nutrient level, according to this theory.

In contrast, J. Philip Grime (1977) and others have argued that certain traits, particularly those that confer rapid growth rates under favorable conditions, are always associated with competitive superiority. In favorable, high-productivity environments, these superior competitors will always dominate. In unfavorable environments with low productivity, competitive ability will not be advantageous; instead, characteristics that confer stress tolerance, such as long-lived leaves with high mean residence times for nutrients (see Chapter 4 and Chapter 14), will determine species dominance and persistence (discussed in greater depth in Chapter 13). In disturbed habitats, other traits, such as dispersal ability, will be favored. Tilman's and Grime's theories make sharply different predictions, particularly concerning the traits that are most consistent with success in nutrient-poor, low-productivity environments: competitive ability to reduce soil nutrients to levels below which competitors cannot survive, and to persist at those low levels (Tilman), or characteristics such as high leaf nutrient retention rates that result in tolerance of unfavorable conditions, rather than success in competitive interactions (Grime). We discuss some of the data on contrasting adaptations to habitats with low and high soil fertility and their consequences at the ecosystem level in Chapter 4 and Chapter 14.

Paul Keddy and colleagues (1998) carried out a large garden experiment using Canadian wetland species. They tested the relationship between plant traits hypothesized to correlate with relative competitive *effects* (of each species on the other species with which it was competing) and competitive *responses* (to competition from the other species). The competitive effect of a species increased as its **relative growth rate** (**RGR**) increased, with RGR calculated as

$$\mathrm{RGR} = [\ln(W_2) - \ln(W_1)]/\Delta_t \qquad (10.6)$$

where the *W*s are plant sizes measured at two times, and $\Delta_t$ is the elapsed time. Species' competitive responses were not related to either growth rate or competitive effect and were similar across a range of very different communities; that is, the response of a species was consistent even when its neighbors were very different. These results suggest that faster-growing, taller species have greater competitive effects on neighbors. Nevertheless, a single experiment, however well done, cannot resolve a controversy. The question remains about how general these results are: we do not know whether this conclusion holds for other herbaceous temperate systems or whether it holds in different kinds of plant communities.

In a test of similar questions in KwaZulu-Natal, South Africa, Richard Fynn and colleagues (2005) compared strategies and competitive abilities of five perennial grassland species. These grass species are ordinarily found at sites that differ greatly in soil fertility (as well as in other factors, such as grazing and fire). In a comparison of two of these species hypothesized to be competitive dominants in soils of differing fertility, the researchers found a reversal of competitive superiority in soils with low and high phosphorus levels. *Panicum maximum* (guinea grass, Poaceae), a tall, broad-leaved grass found on fertile soils, was competitively dominant in experimental plots with high phosphorus levels when grown with *Hyparrhenia hirta* (thatching grass, Poaceae),

a tall narrow-leaved species from infertile soils (Figure 10.11). However, *H. hirta* was the superior competitor when the two species were grown together in experimental plots with very low phosphorus. (Both of these grasses are, incidentally, highly invasive on other continents and islands where they are non-native.)

The study by Fynn and colleagues (2005) offers evidence to support Tilman's prediction that different species should be competitively superior in different environments, and it contrasts with the result for the study by Keddy and associates (1998). However, we should interpret these results cautiously. The South African grass study also included other soil conditions and other species pairs, and the results were far less consistent, so the generality may be limited. From these and other studies, it is still not clear whether there are consistent relationships between competitive ability, soil fertility, and plant strategies. Nor is it clear whether the patterns observed hold for different kinds of plant communities, or just the communities of herbaceous perennials in which these relationships have been most commonly tested.

The relative allocations by plants to roots, photosynthetic organs, stems, and reproduction are traits that can affect plant competitive abilities. Shoot architecture has a large effect on the shading of neighbors. Root structure and distribution in the soil (including lateral extent, depth, and the degree to which roots fill space) affect water and nutrient uptake relative to neighbors' abilities to access these resources. Relative allocations to nonwoody versus woody structures and to perennating structures capable of regeneration versus current photosynthetic or reproductive tissues also affect competitive interactions. It is difficult to make generalizations about how different allocation strategies affect the outcome of competition, however, because their consequences differ for different combinations of species and in different environments.

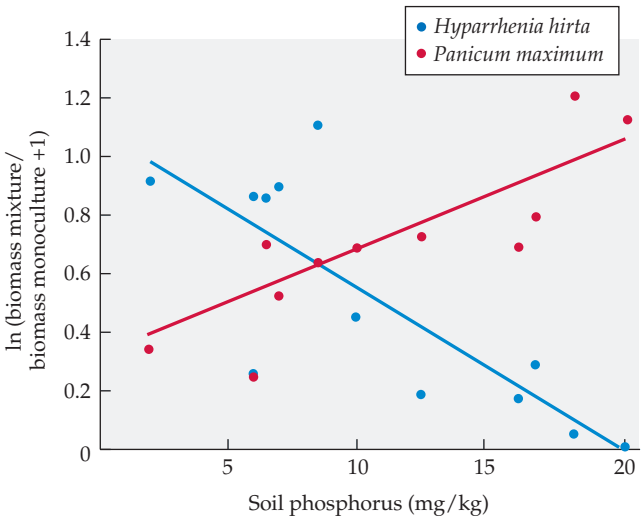



**Figure 10.11** Ratio of biomass in intraspecific competition for *Panicum maximum* (red circles) and for *Hyparrhenia hirta* (blue circles) in soils with phosphorus levels ranging from very low to higher values. At the lowest soil phosphorus levels, *H. hirta* does relatively better in comparison to *P. maximum*, and the opposite is true at high soil phosphorus values (see text). (After R. W. S. Fynn et al. 2005. *J Ecol* 93: 384–394.)

Research by Mark Westoby and colleagues (Westoby 1998; Westoby et al. 2002) has attempted to overcome objections to previous strategy characterizations by proposing objective, widely applicable characterizations that can be used to make worldwide comparisons of plant species in different communities (Box 10A). Westoby's strategies are based on the assumption that there are trade-offs along a small number of axes that define major, easily quantified traits, such as specific leaf area, leaf life-span, leaf size, twig size, canopy height, seed mass, and seed number. (For a similar type of analysis using life history traits, see Figure 7.11.) The goal in developing these characterizations is similar to that of people seeking to define plant functional groups (see Box 12A): reducing the vast diversity of terrestrial plants to conceptual categories so that plants can be grouped in ways that enable one to pose testable hypotheses and make predictions. If the ecology of every plant species must be described uniquely, only very limited progress will ever be made in understanding the ecology of any multispecies units, from plant communities and ecosystems to biomes.

## Are there fixed competitive hierarchies?

Differing views on the nature of trade-offs and strategies are related to differing views on competitive ability, that is, whether it is a fixed characteristic of plant species or it varies according to the environment in which plants live and the other species with which they are interacting. Are some species always competitively superior, and is the rank order of subordinates relatively fixed? This view is implied by Grime's (1977) ideas. Alternatively, others argue that the general rule for plant interactions among species growing together in communities is that dominance in competitive interactions varies with environmental conditions. (We evaluate this question again toward the end of this chapter in light of newer theoretical work.)

The issue of competitive hierarchies is important because it touches on questions regarding the basic structure of plant communities. If consistent competitive hierarchies do occur as a general rule, ecologists should be able to predict the competitive abilities of plant species based on their traits, such as growth form and size (Herben and Krahulec 1990; Shipley and Keddy 1994). Explanations of species diversity would then require an understanding of the factors that prevent competitive exclusion, such as disturbances (see Chapter 13).

## BOX 10A Plant Traits and the Worldwide Leaf Economic Spectrum: Attempts to Simplify Understanding of Plant Diversity

Analyzing the 300,000 or so angiosperms and numerous other taxa to understand fundamental community, regional, and global patterns and processes is a daunting prospect. Yet, we are challenged to understand how changes in climate, land use, global biogeochemistry, and other factors associated with the human domination of the globe will affect extinctions, invasions, range shifts, and the composition of novel communities. One approach has been to simplify plant biodiversity data to just the key traits that might help reach such an understanding (Lavorel and Garnier 2002; Messier et al. 2010). One of the most successful efforts is the conceptualization of the "worldwide leaf economic spectrum," which is based on a mechanistic understanding of plant traits (Chapin 1980; Bloom et al. 1985; Wright et al. 2004; Reich et al. 2006). The basic framework of this theory is that plant leaves have a set of traits (specific leaf area, leaf nitrogen content or carbon:nitrogen ratio, maximum photosynthetic rates, resistance of xylem vessels to cavitation, leaf toughness and life-span, and others) that are associated with one another in specific combinations that form strategies (Reich et al. 2003; Wright et al. 2005). While aspects of the leaf economic spectrum have been criticized on various bases (Bolnick et al. 2011; Wright and Sutton-Grier 2012; Osnas et al. 2013), many of the predicted general patterns seem to hold. These analyses are possible because databases with information collected on plant traits worldwide provide a source of "big data" for researchers seeking to analyze these patterns. Examples include the TRY database (Kattge et al. 2011), with almost 12 million records in 2019, and the GLOPNET database (Wright et al. 2004).

A counterposed view is that competitive dominance varies among environments. If this is the case, then community composition should be determined either by a balance between dispersal and extinction (MacArthur and Wilson 1967) or by the distribution of resources, assuming that species differ in their competitive advantages in competing for different resources (Vandermeer 1969). Models incorporating this perspective assume that consistent competitive hierarchies do not occur across all environments. If, as these models predict, competitive dominance is reversed in different community types, an understanding of community structure would require studies of how the niches of the species differ and of how life history or other trade-offs permit stable coexistence.

If we understand competition to be an effect on fitness or population growth, then competitive success depends not simply on success in acquiring resources, but also on a variety of physiological, life history, and demographic traits. Since all of these components can vary with the environment, we might expect the competitive abilities of plant species to vary as well. John Connolly (1997) pointed out that most of the data that support the existence of competitive hierarchies are largely based on two-species substitutive experiments (described above), which he argued are biased. Biased or not, extrapolation of the results of these greenhouse experiments to natural plant communities is necessarily limited. We need to have more information, particularly from field experiments, before reaching a conclusion about the general existence of competitive hierarchies in nature. Unfortunately, while there have been hundreds of studies of plant competition in the field, few have addressed this issue, and thus the data on this question are still too limited to allow any reliable conclusions to be drawn (Goldberg and Barton 1992).

### Does allelopathy between species explain patterns in nature?

Plants affect their environments chemically in many ways, including altering the pH and balance of nutrients in the soil, altering the physical and chemical composition of the rhizosphere by secreting and taking up materials (see Chapter 4), and shedding parts. One way plants might potentially alter their environments to gain an advantage over competitors is to release toxins that reduce the growth of adjacent plants or even kill them. The existence of such a phenomenon, called **allelopathy**, could be a way for plants to secure uncontested access to resources. It is analogous to interference competition among animals. **Allelochemicals** have been hypothesized to be released in a variety of forms, including root exudates and chemicals leached by rainwater or volatilized from living or dead plant tissues. A large amount of research has been conducted on allelopathy, aimed at characterizing some of these compounds and their modes of action. Many of these studies, however, involve experimental protocols that have little relevance to how plants actually interact in nature, and almost all of them are conducted in controlled environments. Nevertheless, many attempts have been made to explain observed patterns in communities in terms of allelopathic interactions—such as competitive exclusion, negative spatial correlations between species (where the locations of individuals are negatively associated with one another), community dominance, and rapid invasions—as consequences, at least in part, of allelopathy.

However, no such ecological patterns have been conclusively linked to allelopathy in field experiments. Attempts to explain observed ecological patterns in terms of allelopathy have engendered great controversy over many decades. Discussion of allelopathy is difficult because the term is used differently by different writers:

Most plant ecologists today use the term *allelopathy* to mean the direct toxic effects of chemicals released from plants on neighboring plants (Harper 1977; Inderjit and Callaway 2003).

Early authors (e.g., Rice 1974) used the term to refer to both positive and negative effects, whether direct or indirect.

The term has also been used to include indirect effects as well, such as the effects of materials excreted from plants on soil microbes, which then negatively affect other plant species. This approach has led to an active area of research on soil-plant feedbacks, discussed in the subsequent section.

There is no disagreement that plants may have important effects on their chemical and biological environments. Rather, what has been hotly debated is the degree to which allelopathy in the strict sense (plant species producing compounds that have direct or indirect toxic effects on other plants) is widely responsible for causing ecological patterns in nature. We use *allelopathy* here only in the first, strict sense, which is in keeping with its literal meaning.

Demonstrating the existence of allelopathy, much less its importance, involves a number of serious practical problems. Plants certainly produce a wealth of toxic compounds that deter or kill herbivores and pathogens (see Chapter 11). Do these compounds also have negative effects on neighboring plants? There is no doubt that many plants contain substances that can harm other plants when extracted and applied to those plants, but these compounds may be serving strictly or primarily antiherbivore, antibacterial, or antifungal functions. The methods used in testing extractions for allelopathic effects, or using carbon to bind hypothesized allelopathic agents in soil, may not mimic processes that occur under natural conditions. It is difficult to conclusively separate the effects of chemicals hypothesized to be allelopathic from the effects of other factors affecting neighbors' performances and distributions, particularly resource competition and herbivory; in any event, these factors do not necessarily operate separately in nature (Inderjit and del Moral 1997). Similarly, the chemical and physical natures of soils are enormously complex, and it is difficult to elucidate the hypothesized compounds and their actions in the field to determine whether they are released naturally in sufficient quantities to harm neighboring plants. To further complicate the situation, the compounds released by the plant may not be harmful in themselves; rather, their degradation products may be the effective agents. To make matters more difficult, some plant compounds can stimulate microbial growth, which can reduce oxygen availability in the soil, indirectly harming the roots of other plants. Finally, one must be able to explain how the plants avoid autotoxicity: harming themselves as well as their neighbors.

Early research on allelopathy was motivated by the observation that in the California chaparral (a shrub-dominated ecosystem found in mediterranean climates; see Chapter 18), bare patches of ground often surround shrubs of certain species, whereas herbaceous plants are abundant just outside the bare halos surrounding the shrubs, as well as under other shrub species. These shrubs, particularly *Salvia leucophylla* (purple sage, Lamiaceae) and *Artemisia californica* (California sagebrush, Asteraceae), have pungent aromatic scents, indicating the release of volatile chemicals into the environment. In the early 1960s, Cornelius Muller and his colleagues and students (Muller and Muller 1964; Muller 1969) believed that they had found evidence for allelopathy in these two species. Eventually, they attributed many other ecological phenomena in chaparral to allelopathy.

However, in an experiment testing Muller's claims, Bruce Bartholomew (1970) found a surprising alternative explanation for the lack of plants under the shrubs hypothesized to be allelopathic. Bartholomew placed small wire exclosures in the bare halos under chaparral shrubs to keep out herbivorous small mammals. Inside the exclosures, a lush growth of herbaceous plants appeared. Apparently the animals spent most of their time foraging under the cover of the shrubs, seeking protection from predators (particularly birds). Bartholomew concluded that the bare zones were created largely by herbivory. Other researchers raised both technical and conceptual questions regarding the demonstration of and the importance of allelopathy in nature in the 1970s and 1980s (e.g., Harper 1977; Newman 1978; Stowe 1979; Stowe and Wade 1979). A decade or so later, Jon Keeley and his students (1985) conducted further studies on allelopathic inhibition in chaparral, attempting to overcome many of the limitations and criticisms that had affected Muller's work (e.g., evaluating the effects of hypothesized allelopathic chemicals on cultivated species, such as cucumber seedlings). Their experiments showed that of the 22 native species tested, 2 were inhibited by the hypothesized allelopathic chemical leachate, 11 were unaffected, and seedlings of 9 other species actually benefited from the chemicals. Keeley and his coworkers also found that fire was overwhelmingly the primary cue for the germination of seeds of native herbaceous plants. This long and interesting saga of the process of science in an area of controversy, including the role of the personalities of scientists, is discussed in detail by Richard Halsey (2004).

Another area of allelopathy research during the 1960s and 1970s was based on the suggestion that dominant

grassland species could inhibit succession to woody vegetation for many decades by allelopathic interactions with soil microorganisms (Rice 1974). This research was subject to many of the same methodological criticisms as was Muller's.

In the twenty-first century, two well-known researchers in allelopathy, Inderjit and Ragan Callaway (2003), directly addressed many of these methodological problems. They argued that convincingly demonstrating allelopathy requires determination of the concentrations and release rates of hypothesized allelochemicals, followed by experiments in which concentrations of these compounds alone are manipulated. They further argued that experiments must be designed to distinguish between toxic chemical effects on the one hand and microbial or resource effects on the other. Laboratory and greenhouse observations need to be directly linked with patterns in the field, emphasizing the importance of large-scale manipulations of chemical effects.

A major claim (and dispute) about allelopathy was made by Callaway and his associates (Callaway and Aschehoug 2000; Bais et al. 2003), who proposed that allelopathy is the major cause of the success of two important invasive weeds in North American rangelands. *Centaurea diffusa* (diffuse knapweed, Asteraceae) and *C. stoebe* (spotted knapweed; Figure 10.12), both native to Europe and Asia Minor, now infest millions of hectares in the United States. Compounds were isolated from each of these species that have been found in substantial concentrations in the soils in the rangelands they have invaded and do not appear to have other sources. In laboratory studies using the same ranges of concentrations reported in invaded rangeland soils, these compounds had toxic effects on native North American range plants. The compounds in question were also believed to promote nutrient uptake by *Centaurea*. In an interesting twist to the story, these allelochemicals have only weak effects on Eurasian plants from the native regions of the *Centaurea* species, and the *Centaurea* plants themselves are not affected by these compounds. Jorge Vivanco and colleagues (2004) suggested that these observations might reflect evolutionary adaptation to tolerate these compounds in the native environment, but not yet in the new one.

It turned out, though, that the case for allelopathy as a major mechanism for the successful invasion of *Centaurea* species in western North America was not closed. Amy Blair and coworkers (2005), using a then-new technique to assay for the compound reported from *C. stoebe*, found no measurable quantities of the compound in knapweed-infested soils and found only slightly enhanced mortality in one of the native species previously reported to suffer 100% mortality. Katharine Suding and associates (2004) studied competition and resource availability in a heavily invaded rangeland site in Colorado,

Courtesy of G. C. Thelen

**Figure 10.12** *Centaurea stoebe* (the rose/purple-flowered species covering the hillside in the foreground) is widely invasive in rangelands in western North America.

where they examined both the competitive responses and competitive effects of *C. diffusa* and three other species by manipulating local nutrient availability and plant densities. They found that *C. diffusa* was better able to tolerate competition than the other three species studied but that it had little species-specific effect on the growth or survival of its competitors.

What can we conclude? It has been difficult to demonstrate that allelopathy has major effects in structuring species interactions in nature (Inderjit and Callaway 2003), and despite considerable efforts, no field experiment has unambiguously demonstrated allelopathy to be the cause of a community pattern. The question that is most important, however, is not, Does allelopathy occur? or even, Does allelopathy explain patterns seen in natural communities? Considering the number of plant species and the number of compounds each one makes, it would not be surprising if some of them secondarily or indirectly had toxic effects on neighbors, or even direct effects, if some of those toxic effects sometimes resulted in altered community patterns. What we need to know is how common those effects are overall, relative to the effects of other factors (including other biotic interactions), in structuring communities. In the largest context, we need to ask (Inderjit and Weiner 2001): How do plants change their soil environments, and how do these changes affect their interactions with both conspecifics and competitors? We also need to know more about how hypothesized allelopathic chemicals affect competing neighbors mediated by organisms at other trophic levels. These are broadly challenging questions for research, much of which occurs under the name *plant-soil feedbacks* (see Chapter 4). A review by Alison Hale and

Susan Kalisz (2012), for example, points to evidence that species from a number of plant families can secrete substances that inhibit mycorrhizal associations of their neighbors. The interactions between some plant species may thus be more complex than a framework of simple resource competition would suggest.

### Plants can change the environment to the advantage of other plants

Plants can have positive as well as negative effects on neighbors. Sometimes the effects of neighbors are positive at one time and become negative at later stages in the plants' lives. Negative and positive effects can even occur simultaneously; for example, neighbors may enhance survival but decrease growth. Ecological theorists have been criticized for focusing largely on competition between plants, ignoring facilitation—positive interactions among species. Facilitation is an interaction in which the presence of one species alters the environment in some way that enhances some component of fitness of a neighboring species. While, at first glance, facilitation sounds like another term for mutualism, Judith Bronstein (2009) made clear the distinction: facilitation may be mutualistic (e.g., associational defense, see Chapter 11) but can also occur in commensalisms (e.g., one species provides shade that benefits another). As we will see in this section, things can be more complicated—an interaction may be facilitative in some contexts but competitive or exploitative in others. Nor do all mutualisms involve facilitation (e.g., plant-pollinator mutualisms).

Andreas Fitchner and collaborators (2020), working in an enormous manipulative field study on biodiversity in China, examined the effects of the diversity of neighbors on the ability of young trees to withstand drought stress. Trees were planted across gradients of biodiversity and climate. The effects of the species richness of neighbors and the functional traits of the young trees were evaluated. The researchers found that the species richness of neighbors had facilitative effects in reducing drought-induced growth reduction and that these beneficial effects were greatest during drought periods and increased with the taxonomic diversity of neighbors. Trees that had the greatest vulnerability to drought (as measured by leaf traits and vulnerability to xylem cavitation in response to water stress; see Chapter 3) were helped most.

Neighbors' effects may certainly differ depending on other aspects of the environment. Suzanne Boyden and colleagues (2005) studied the interactions between two plantation tree species in Hawaii. They found that the trees' effects on each other's growth and survival were complex, varying from strongly competitive to strongly facilitative, depending on soil nitrogen and phosphorus. The survival of *Falcataria moluccana* (Moluccan albizia, Fabaceae), a nitrogen-fixing tree, was enhanced by neighboring *Eucalyptus saligna* (Sydney bluegum, Myrtaceae) at high soil nitrogen, but competitive effects of *Eucalyptus* decreased the survival of *Falcataria* at low soil nitrogen (Figure 10.13A). In contrast, *Eucalyptus* survival was only weakly affected by either intra- or interspecific competition, and these effects did not vary with soil nutrients. *Falcataria* growth was enhanced by positive interactions with *Eucalyptus* at low nitrogen but decreased by competition with *Eucalyptus* at high nitrogen (Figure 10.13B). Intraspecific effects of *Falcataria* on its own growth were facilitative and increased with conspecific density at low soil phosphorus but were competitive and became more negative with intraspecific density at high soil phosphorus (Figure 10.13C). *Eucalyptus* benefited somewhat from *Falcataria* neighbors at high soil phosphorus but experienced competitive effects from *Falcataria* under low soil phosphorus (Figure 10.13D). *Eucalyptus* had highly negative competitive intraspecific effects on growth that were not changed by soil resources. The mechanisms of these interactions are not yet well understood, but the authors speculated that mycorrhizal connections may mediate the complex interplay between resource availability and competition in determining growth and survival in these trees.

**Nurse plants** enhance the establishment of juvenile plants in a variety of community types. Generally the "nurse" is a mature plant of a different species, and often a different growth form, than the juvenile. Some of the best-known examples occur in deserts (see Figure 13.17). Both shrubs and grasses were found to increase the survival of seedlings and cladodes (vegetative reproductive units) in *Opuntia rastrera* (prickly pear, Cactaceae) in the Chihuahuan Desert (Mandujano et al. 1998). Nurse effects were attributed primarily to protection from herbivory, but also to shading. The massive columnar cactus *Neobuxbaumia tetetzo* (cardón, Cactaceae) in central and southern Mexico becomes established with the help of the nurse shrub *Mimosa luisana* (Fabaceae). The cactus eventually suppresses and may kill the nurse plant as a result of competition for water (Flores-Martinez et al. 1998). None of this happens as an act of kindness, cruelty, or even coevolution; the "nursed" plant benefits from early protection without the "nurse" choosing to protect it or not, and it may repay the nurse's "kindness" by suppressing or eventually killing it.

Nurse plants can facilitate germination, establishment, and seedling growth in many other kinds of environments besides deserts. Experiments demonstrated the important effects of facilitation on the resprouting of herbaceous and woody plants in shrublands in southern Argentina after severe fires (Raffaele and Veblen 1998), where nurse plants increased soil moisture by reducing temperature and light intensity. Similarly, in the pine barrens of Long Island in

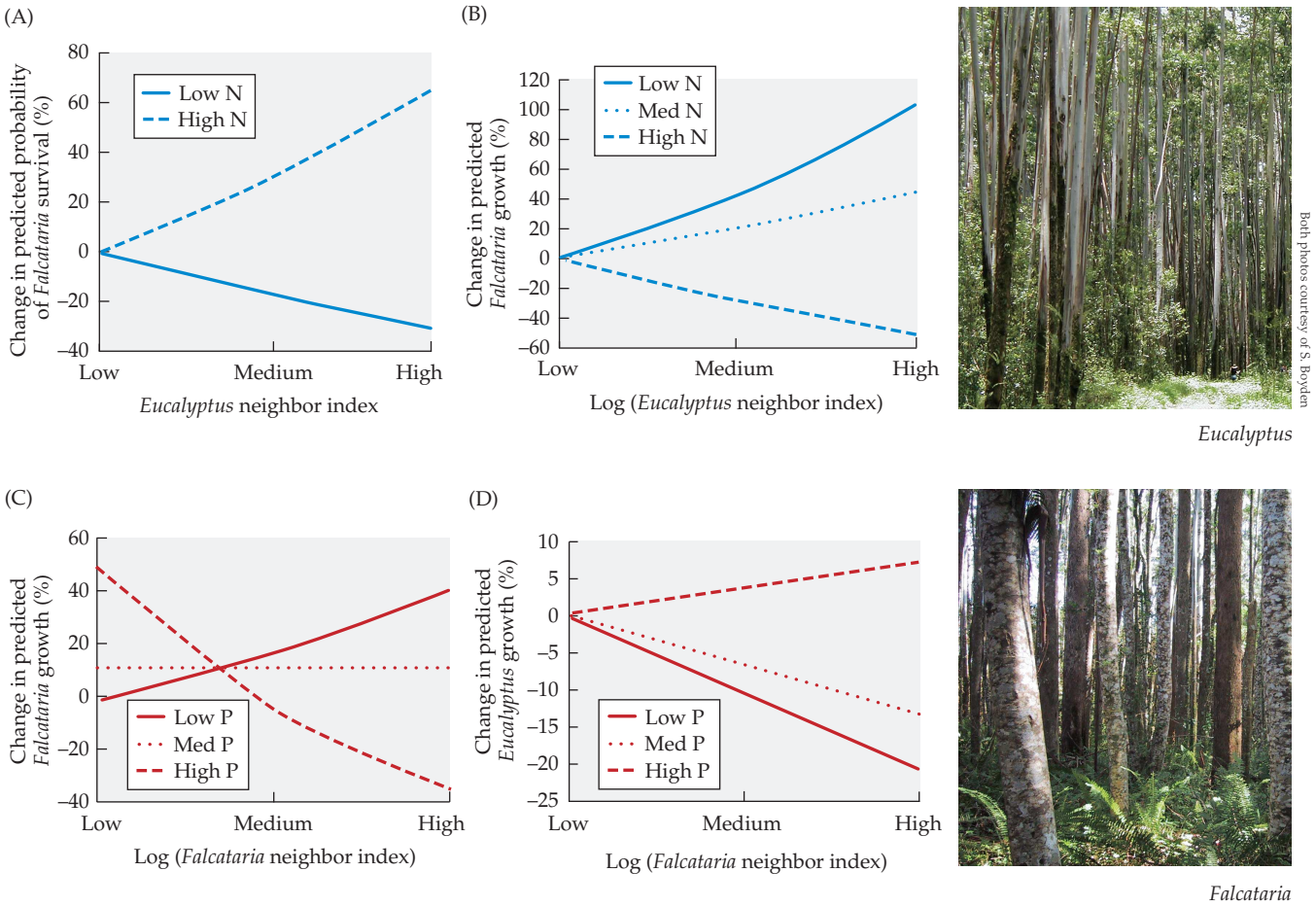



**Figure 10.13** (A) The percentage change in survival predicted for *Falcataria* at low and high soil nitrogen with low, medium, and high densities of *Eucalyptus* neighbors. (B) The change in predicted growth of *Falcataria* at low, medium, and high soil nitrogen with low, medium, and high densities of *Eucalyptus* neighbors. (C) The change in predicted growth of *Falcataria* at low, medium, and high soil phosphorus with low, medium, and high densities of *Falcataria* neighbors. (D) The percentage change in predicted growth of *Eucalyptus* at low, medium, and high soil phosphorus with low, medium, and high densities of *Falcataria* neighbors. Predicted change was based on multiple logistic regression for survival, and generalized linear models for growth. (After S. Boyden et al. 2005. *Ecology* 86: 992–1001.)

eastern New York state, the shrub *Quercus ilicifolia* (scrub oak, Fagaceae) increased the survival of seedlings of *Pinus rigida* (pitch pine, Pinaceae) regenerating after an intense fire but strongly suppressed their later growth (Landis et al. 2005). At high altitudes, such facilitation may be important in areas as diverse as the Colorado Rockies (Wied and Galen 1998), the Andes (Cavieres et al. 1998), and high-elevation forests of Hawaii (Scowcroft and Jeffrey 1999), where nurse plants protect seedlings from frost, reduce radiative cooling (see Chapter 2), and ameliorate summer drought stress. On an active volcano in northern Japan, different shrub types had opposite effects on facilitation of herbaceous understory species, as shown by Shohei Uesaka and Shiro Tsuyuzaki (2005). Patches of a deciduous shrub generally enhanced the germination, growth, and diversity of neighbors, but patches of an evergreen shrub inhibited establishment of herbaceous species. The deciduous shrub trapped wind-dispersed seeds of understory species. Compared with surrounding areas outside of shrub patches, nutrients and soil moisture were much higher under both deciduous and evergreen shrubs, but light levels were reduced, particularly under evergreen shrubs.

Nurse plants may also facilitate establishment on coastal sand dunes, where they act to stabilize the substrate and moderate conditions such as wind, drought, and temperature fluctuations (Martinez and Moreno-Casasola 1998). But shrubs do not always act to facilitate the growth of neighbors. Marcelo Sternberg and colleagues (2004) found no facilitation effects of shrubs on annuals in a Mediterranean dune system in Israel. In Bartholomew's (1970) study discussed in the previous section, shrubs had indirect negative effects on annuals beneath them by sheltering small herbivores.

One intriguing type of facilitation may occur through plant associations with fungi in mycorrhizae (see

Chapter 11). It is known from laboratory experiments that mycorrhizal fungi can form connections with more than one plant, even of different species, and allow transfer of materials between those plants. In many communities, extensive mycorrhizal connections, called **common mycorrhizal networks** (CMNs), are known to link plants of many different species. Carrying out research on CMNs in natural communities presents many technical challenges. It has been hypothesized that CMNs have the potential to alter competitive interactions; facilitate seedling and sapling survival and growth; transfer carbon, nitrogen, phosphorus, and water among plants; and alter diversity patterns and community dynamics (He et al. 2003; Simard and Durall 2004). Besides regulating nutrient flux among individual plants, CMNs also are important in competitive interactions, influencing species coexistence and community structure (Tedersoo et al. 2020). Some authors have cautioned, however, that the evidence that these events are actually occurring and are important in nature is equivocal (Fitter et al. 1999).

Researchers have been especially interested in the possibility that CMNs allow facilitation between mature and juvenile trees in forests, because such positive interactions might counter competitive interactions, with large implications for succession and the maintenance of diversity. Nerre Awana Onguene and Thom Kuyper (2002) carried out experiments on the role of mycorrhizae in mediating plant-plant interactions in forests in Cameroon, in western Africa. They showed that when seedlings of the ectomycorrhizal rainforest tree *Paraberlinia bifoliolata* (African beli, Caesalpiniaceae) were grown near adult trees of their own or other species, they formed mycorrhizae more readily than when grown in isolation from mature trees. Seedlings in contact with roots of mature trees had higher survival and growth rates than seedlings that were not associated with mature tree roots. There was considerable variation in the positive effects associated with connections with different adult tree species. Presumably the seedlings formed CMNs and benefited by receiving materials from the mature trees (although this was not demonstrated). Experiments in a temperate New England forest revealed even more complex interactions, with CMNs of ectomycorrhizal species having positive effects on the growth of juveniles of one ECM species, *Pinus strobus* (white pine, Pinaceae), and negative effects on the survival of those of an arbuscular mycorrhizal species, *Acer rubrum* (red maple, Sapindaceae) (Booth 2004).

## Competitive exclusion sometimes determines species distributions

In one of the earliest experimental studies of species interactions, Arthur Tansley (1917) sought to determine the role of competition in the distributions of two species of *Galium* (bedstraw, Rubiaceae), both small herbaceous perennials, in the United Kingdom. *Galium saxatile*, the heath bedstraw, is found primarily on sandy soils, while *G. sylvestre* (probably the species now called *G. sterneri*), the limestone bedstraw, is found primarily on calcareous, limestone-derived soils. Tansley grew plants of both species alone and together in each of the two soil types in large wooden boxes outdoors. When grown alone in calcareous soil, limestone bedstraw grew normally, but heath bedstraw grew slowly and had yellowish leaves, indicating nutrient deficiency. When grown alone in sandy soil, heath bedstraw grew vigorously, while limestone bedstraw survived but grew poorly. When the two species were grown together in calcareous soil, limestone bedstraw overtopped heath bedstraw and eliminated it from the mixture. When they were grown together in sandy soil, heath bedstraw became dominant but did not completely eliminate limestone bedstraw during the experiment. Tansley concluded that while each species appeared to be adapted to the soil in which it lived in nature, competition also played an important role in determining the restriction of the two species to different soil types (see the discussion of fundamental and realized niches in Chapter 9).

But competitive exclusion is not always a good explanation for species' distributions. M. B. Richards and his colleagues (1997) examined the importance of competition and adaptation to soil type among six shrubs in South Africa belonging to the Proteaceae. These species grow in the extraordinarily diverse fynbos of southern Africa (see Box 19A). Large numbers of apparently ecologically similar species coexist (high α-diversity; see Chapter 12), and there is great variation in species among communities (high β-diversity). Sharp discontinuities exist among communities, each dominated by different species. Richards and colleagues asked, What is the role of competition in determining these boundaries? They chose three transects along which they compared the influences of soil type and interspecific competition in determining species' distributions. Each transect crossed a sharp community boundary where there was a transition from one distinct soil type to another. There was a different species pair in each transect in which one dominant species replaced the other along the transect. In a 3-year experiment, seeds of both species were planted in monoculture and in mixture at three sites along each transect. Interspecific competition had negative effects on growth, but these negative effects could not explain species distributions or the sharp differentiation of the communities. There was an effect of competition, but it was small compared with that of differences in soil type among communities. Adaptation to soil conditions strongly affected both seedling growth and survival at two of the three experimental sites. The researchers concluded that soil type, not competitive exclusion, may be the critical factor determining the distributions of these species.

The ultimate outcome of competition may be different from what short-term results suggest. *Lythrum salicaria* (purple loosestrife, Lythraceae) is an invasive species in North America that appears to be displacing native wetland species (see Figure 14.4). Some ecologists have argued, however, that evidence that *L. salicaria* is actually displacing native species is weak. Tarun Mal and colleagues (1997) carried out a 4-year field experiment to examine competition between *L. salicaria* and *Typha angustifolia* (cattail, Typhaceae), a dominant native wetland species. *T. angustifolia* was initially competitively superior, but in the second and third years of the experiment, the species were relatively evenly matched. By the fourth year of the study, *L. salicaria*, the invader, became competitively dominant, displacing *T. angustifolia*. The researchers attributed this result to differences in life history strategy between the two species. *T. angustifolia* has large rhizomes with substantial stored resources, which might give it an initial competitive advantage, but the high costs of producing new ramets and the strong suppressive effects of *L. salicaria* led to the eventual competitive replacement of *T. angustifolia*. When asking questions about the population- or community-level effects of competition, studies and metrics need to account for the entire life cycles of individuals.

## 10.4 Competition and Facilitation May Vary along Environmental Gradients

One way to examine the importance of competition and facilitation is to look at how species interactions change along gradients of productivity. Are there particular kinds of habitats in which competition is predictably strong, determining community composition, or predictably weak and unimportant? Is facilitation more important in some kinds of habitats than others? Ecologists agree that competition is intense in productive, nutrient-rich habitats, at least when disturbance and herbivory are low. In these environments, plants are able to develop large canopies quickly, and competition is thought to be primarily for light. The relative intensity and importance of competition and facilitation across a range of habitats, however, remain matters of debate.

### There are conflicting models of how productivity affects the importance of competition and facilitation

Grime (1977, 1979) proposed that competition is unimportant in unproductive environments and that success in these environments is dependent largely on the ability to tolerate abiotic stress (low nutrient levels, drought, or cold, for example), rather than on competitive ability. He further argued that in environments where disturbance frequently reduces plant biomass, competitive exclusion should be prevented. The dominant plants in such environments should not be competitively superior, but rather should possess traits that allow them to withstand or recover from disturbance and recolonize rapidly following disturbance.

Most subsequent discussion has focused on unproductive environments. Edward Newman (1973) disagreed with Grime's characterization, arguing that competition is important in low-resource as well as high-resource environments but that the resources for which plants compete differ—light in productive environments, nutrients and water in unproductive environments. Later work by David Tilman (1987) reinforced and developed Newman's ideas. Tilman argued that competition in low-productivity environments would be for belowground resources (primarily nitrogen), whereas competition in high-productivity environments would be primarily for aboveground resources (light). Rein Aerts (1999) reiterated Grime's argument in part, maintaining that selection in nutrient-poor habitats would favor traits that reduce nutrient losses rather than those that enhance the ability to compete for nutrients, resulting in slow growth rates.

As with competition, is facilitation especially likely under some circumstances? Mark Bertness and Ragan Callaway (1994) predicted that positive interactions should be particularly common under conditions of high abiotic stress or high levels of herbivory. Under conditions in which the physical environment is favorable for growth and there is less herbivory, they predicted, competitive interactions should be important in structuring communities—the stress-gradient hypothesis. Many people have studied positive interactions among a wide variety of plants in a wide array of communities since those predictions were made, but as yet no agreement has been reached about the general role of facilitative interactions relative to competition or herbivory. While this hypothesis is distinct from the predictions of Grime and Tilman regarding competition, they are often discussed together.

Still another view on how species interactions and productivity intersect comes from Paul Keddy (1990), who created a plant community version of the centrifugal theory of community organization that Michael Rosenzweig and Zvika Abramsky (1986) had proposed for desert rodent communities. The centrifugal model proposes that there is a core habitat type preferred by all species in a region, presumably with ideal growing conditions. Other habitat types, called peripheral habitats, are defined by particular negative conditions (stresses or disturbance) to which only some of the species are adapted (Figure 10.14). Interspecific competition is most intense, Keddy suggested, in core habitat and is more relaxed in peripheral habitats because fewer species are adapted to their particular conditions. The peripheral habitats serve as refuges, preventing competitive

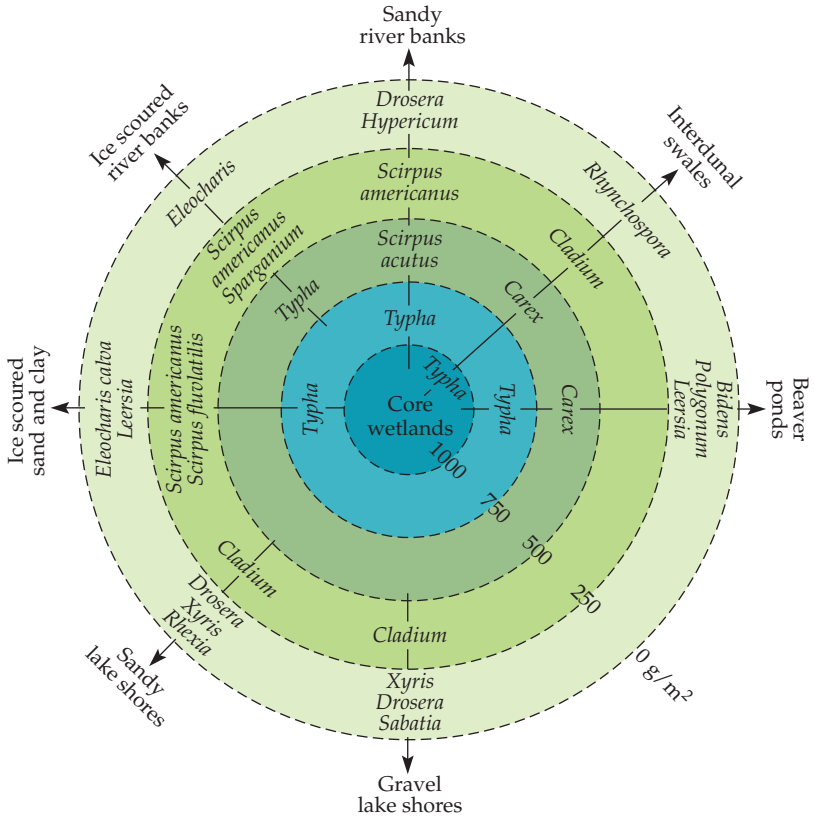



**Figure 10.14** The centrifugal model of plant community organization, applied to the distributions of a number of wetland species in Ontario, Canada. *Typha latifolia* (cattail, Typhaceae) occupies core habitat, while other species become more prominent as one moves toward more extreme conditions in more peripheral habitats. The rings refer to habitats with different productivities (expressed in g/m$^2$), and the genus and species names indicate the dominants in different habitat types. For example, among the communities with lowest productivity, *Xyris*, *Drosera*, and *Sabatia* dominate on gravel lake shores, while *Eleocharis* and *Leersia* dominate on areas of sand and clay that experience ice scouring. (After D. R. J. Moore et al. 1989. *Biol Conserv* 47: 203–217.)

exclusion in the landscape as a whole. In wetlands in Ontario, Canada, for example, all species prefer sites with high fertility and low disturbance rates (core habitat), while the peripheral habitats are defined by infertile soils and disturbances such as ice scouring. The core habitat is dominated by *Typha latifolia* (cattail, Typhaceae), and different species dominate as one moves toward more extreme conditions in each peripheral habitat type.

One of the problems with trying to resolve the debate about the relative intensity and importance of competition and facilitation in productive and unproductive habitats is that real-world environments can be unproductive for very different reasons. Many of the hypotheses about competition in unproductive environments have often implicitly focused on low nutrient levels. But inadequate water supply is one of the most important factors limiting productivity globally. Low productivity may also be due to cold temperatures, short growing seasons, saline soils, or toxic materials in the soil that inhibit growth, such as heavy metals.

Not only are plants likely to have very different adaptations to drought, cold, and heavy metals than to low nutrients, but species' interactions under each of these conditions are also likely to be very different. Deborah Goldberg and Ariel Novoplansky (1997) predicted that the effects of competition in nutrient-poor environments would differ substantially from its effects in water-stressed environments. Water availability, they pointed out, is pulsed: even in arid environments, water is freely available for short periods of time, but there are long interpulse intervals in which water is partly or completely unavailable. They hypothesized that in arid environments growth and competition should be limited to the periods of high water availability (the pulses). Growth and competition in low-nutrient soils, in contrast, should not be limited to pulses of short duration.

There are also conceptual and statistical difficulties in evaluating these ideas. The hypotheses about competition and facilitation are easy to state verbally, but how should we measure the importance of these interactions? As with measuring competition itself (see "How we quantify competition can affect experimental results" above), a number of different ways of measuring importance and intensity have been proposed. Charles Welden and William Slauson (1986) initially raised this problem but did not propose any single way of quantifying importance or intensity. Lauri Oksanen and his associates (2006) showed that RCI has some severe statistical problems and recommended not only that it not be used, but that studies using RCI should be excluded from meta-analyses. Mark Rees, Robert Freckleton, and their collaborators (2012) also showed that various indices of competition (such as RCI, ACI, and LRR) are related to productivity in ways that make them problematic for evaluating Grime's (1977, 1979) hypothesis. A more general point was made by Robert Freckleton and coworkers (2009): any measure of the importance of competition must be explicit about what it is measured with respect to (e.g., population growth or community composition). They showed that none of the indices used are directly related to any quantity about population growth or community composition and questioned whether a single general index of importance could actually be useful.

Mark Rees (2013) proposed a different approach to studying this problem: develop a mechanistic model of resource competition that describes rates of individual growth and changes in resource availability on the time scale relevant to an experiment. This makes it possible to define an intensity of competition in terms of the ratio

of biomass accumulation between competitive and noncompetitive treatments. Rees went on to ask how this ratio depends on productivity, multiple competitors, and nonlinear resource uptake. One reason it has been difficult to evaluate hypotheses about how species interactions may change with productivity is that they have been posed too simply and vaguely, and they do not lend themselves to meaningful testing. Indeed, Rees's model makes more complicated predictions, for example, that intensity of competition should increase greatly with productivity among subordinate species but much less so among dominant species. Increases in numbers or biomass of competitors in more productive environments should lead to a greater intenstity of competition, an effect that can be missed in controlled-environment experimental settings where numbers are likely to be fixed and biomass will depend on pot sizes. There are many ways in which Rees's approach might be extended reasonably. For example, it could be extended to demographic performance by coupling it with additional equations describing reproduction as a function of biomass. In any case, explicit mechanistic models of competition may prove helpful in resolving some of the ambiguities and contradictions about competition along productivity gradients.

Courtesy of Mark Rees

Mark Rees

## Experimental evidence provides a mixed picture about the roles of competition and facilitation along productivity gradients

The evidence about whether the intensity of competition varies along productivity gradients remains contradictory and ambiguous. Scott Wilson and Paul Keddy (1986) compared the competitive abilities of six species that are dominant at different points along a productivity gradient on the shore of Axe Lake, Ontario. Plants were collected along a gradient that ranged from wave-disturbed, nutrient-poor beaches with low standing biomass to sheltered, nutrient-rich sites with dense vegetation. Plants were grown in competition in plastic beakers at a protected site in the field, using substrate from the favorable (sheltered, nutrient-rich) end of the gradient. The researchers found that competitive ability (measured as both competitive response and competitive effect) was positively correlated with mean position along the productivity gradient (Figure 10.15). They interpreted these results as evidence that species with high competitive ability occupied nutrient-rich, undisturbed sites, while species with low competitive ability were displaced to disturbed sites with poor soils, where competitive exclusion was prevented by wave action and low soil nutrients. One limitation of this study, however, is that competition was not actually quantified along the gradient, but only with the nutrient-rich substrate and in a single site under somewhat artificial conditions. Nor did the study estimate either the intensity or importance of competition. It is not clear how one can interpret the results in terms of the predictions of the models discussed above.

Jessica Gurevitch (1986) carried out a field study of competition along an environmental gradient in southeastern Arizona. She hypothesized that *Hesperostipa neomexicana*, a $C_3$ grass, was limited to arid ridgetops by competitively superior $C_4$ grasses. This is precisely the

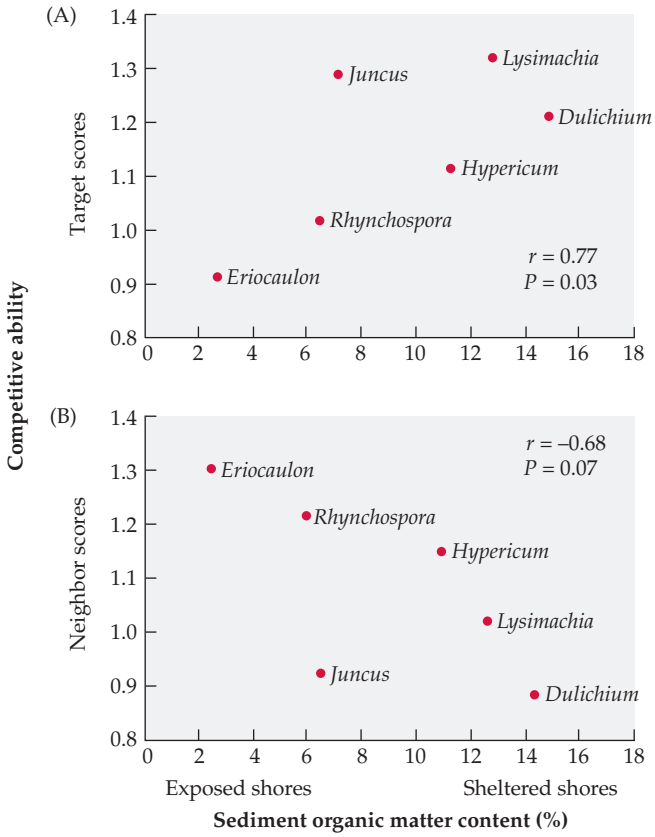
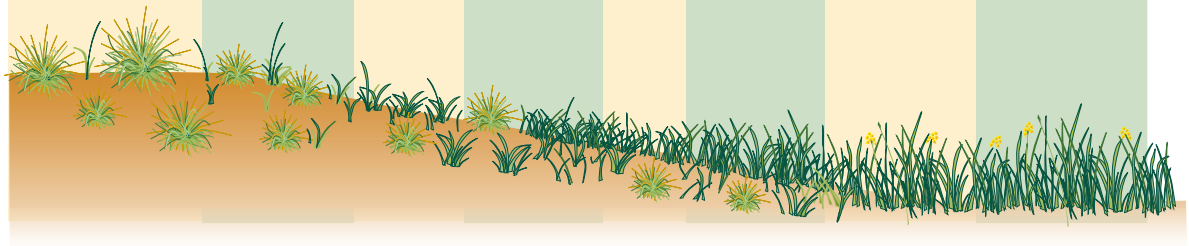



**Figure 10.15** Competitive abilities of six wetland plant species grown in a garden experiment. These species are naturally found along a gradient from exposed, nutrient-poor lake shores to sheltered, nutrient-rich sites (shown on *x*-axis), corresponding to sediment organic matter. (A) Competitive abilities expressed as target scores, defined as the mean relative growth (increase in dry mass) of the target species when grown in the presence of all neighbor species; this score is similar to competitive response. (B) Competitive abilities expressed as neighbor scores, defined as the mean relative growth (increase in dry mass) of all neighbor species in the presence of the target species; this score is similar to competitive effect. Competitive effects and responses were not positively correlated. (After S. D. Wilson and P. A. Keddy. 1986. *Ecology* 67: 1236–1242.)

opposite of what one would expect if physiology were determining species distributions, as $C_4$ species should be better able to tolerate the unfavorably hot, dry conditions on the ridge crests. She removed neighbors from around target *H. neomexicana* individuals at three sites along a gradient from a ridge crest to a moister lower slope, and she found that competition affected growth of mature plants, flowering, seedling establishment, and survival. When competitors were removed, growth and flowering for the mature *H. neomexicana* plants were greatest on the lower slope, where their abundance was lowest. Competition had the smallest effect on estimated population growth rates on the ridgetops, where *H. neomexicana* was most abundant, and increasingly greater effects downslope. The largest effects were at the lowest sites where *H. neomexicana* was present (Figure 10.16). These results strongly suggested that competition was a major factor in determining the distribution of

(A)

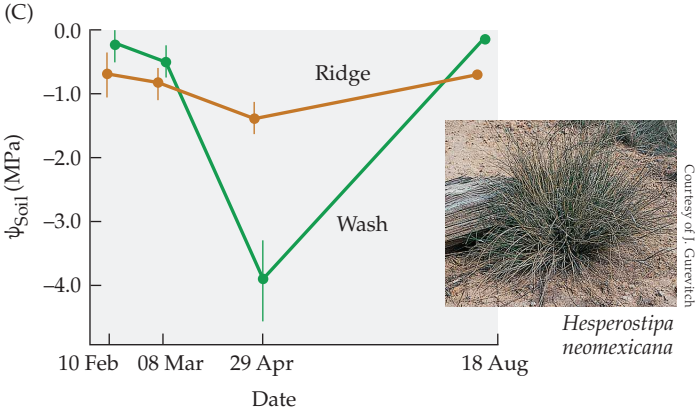
| | Ridge crest | | Midslope | | Lower slope | | Wash | |
|---|---|---|---|---|---|---|---|---|
| | *Hes* | $C_4$ | *Hes* | $C_4$ | *Hes* | $C_4$ | *Hes* | $C_4$ |
| Biomass (g/m$^2$) | 140 | 16 | 76 | 111 | 1 | 282 | 0 | High, not measured |
| Cover (%) | 20.5 | 9.5 | 8.5 | 24.5 | 3.0 | 37.0 | 0 | ~100% |



(B)

| Hesperostipa: | Ridge crest | | Midslope | | Lower slope | |
|---|---|---|---|---|---|---|
| | Control | Removal | C | R | C | R |
| Seedlings/m$^2$ | 4.55 | 9.36 | 0.58 | 2.04 | 0.10 | 0.35 |
| Seedling survival | 0.47 | 0.79 | 0 | 0.56 | 0 | 0.50 |
| Growth $\left(\frac{\text{final size}}{\text{initial size}}\right)$ | 0.77 | 1.97 | 0.69 | 2.58 | 1.38 | 4.28 |
| Number of flowers | 1.50 | 13.30 | 4.10 | 28.70 | 35.50 | 58.90 |
| Population growth (λ) | 0.93 | 1.04 | 0.83 | 0.92 | 0.59 | 0.88 |

(C)




(D)

Courtesy of J. Gurevitch

**Figure 10.16** Results of a removal experiment examining the effects of competition on *Hesperostipa neomexicana* at three topographic positions in a southeastern Arizona grassland (elevation 1400 m). (A) Means for inital biomass and cover ($N$ = 40 plants) for *H. neomexicana* and its $C_4$ grass neighbors (combined) at the three experimental sites (ridge crest, midslope, and lower slope) and in the wash below them. The drawing diagrams the cover and the distribution of *H. neomexicana* (with paler leaves) and $C_4$ grasses in the three topographic positions. (B) Experimental results for control (no removals) and removal (all neighbors removed) treatments at the three experimental sites. Means for seedlings/m$^2$ (in 1980 plus 1981), seedling survival, mature plant growth, and number of flowers produced by mature plants ($N$ = 20) over the 20 months of the experiment are shown. (C) Soil water potential ($\psi_{soil}$) at the ridge crest and wash at 15–20 cm depth (where most roots are located) over a drying and rewetting cycle of 6 months in 1980. (D) *H. neomexicana* growing at midslope at the experimental site; plants are in flower. (After J. Gurevitch. 1986. *Ecology* 67: 46–57.)

*H. neomexicana* along this gradient of productivity and environmental favorability.

Theresa Theodose and William Bowman (1997) suggested the existence of the opposite pattern, in which competition prevented a species from a more productive area from growing in a resource-poor site. The perennial *Deschampsia cespitosa* (hair grass, Poaceae) is common in moist alpine meadows in the tundra of the Front Range of Colorado but is rare in dry meadows. The dry meadows are dominated by a sedge, *Kobresia myosuroides* (Cyperaceae). The authors hypothesized that *D. cespitosa* was prevented from growing in the dry environment by competition with *K. myosuroides*. An earlier study had demonstrated that *K. myosuroides* was kept out of the moist meadows by deep winter snow. To test these ideas, they transplanted individuals of each species, as well as two-species pairs of plants, to a dry meadow, either clipping the existing vegetation (largely *K. myosuroides*) at ground level or leaving it intact. *D. cespitosa* had a greater increase in survival in response to vegetation clipping than did *K. myosuroides*, and soil moisture was substantially depressed in plots with intact vegetation compared with those in which vegetation was clipped (Figure 10.17). The researchers concluded that interspecific competition with *K. myosuroides* excluded *D. cespitosa* from the dry meadows by depressing soil moisture below the drought tolerance of *D. cespitosa*. The *K. myosuroides*, with greater tolerance of drought, was able to survive during periods of low moisture.

One problem with this conclusion, however, is that the mortality of *K. myosuroides* (28%) was actually higher than the mortality of *D. cespitosa* (15%) in intact vegetation. Furthermore, most of the *D. cespitosa* plants survived in the intact vegetation of the dry meadow. Therefore, it is difficult to argue convincingly that competitive exclusion was the result of high mortality in *D. cespitosa*. The growth of *D. cespitosa* in intact vegetation was also greater than the growth of *K. myosuroides*, and the growth of both species was about equally affected by vegetation clipping. (This is a good example of why it is important to actually look at the data in a paper, rather than relying solely on the paper's abstract in considering its conclusions.) Nothing is known about the effects of competition on reproduction or establishment of these species in these environments. So, although this study clearly demonstrated intense and statistically significant effects of competition in this resource-poor habitat, more work needs to be done to conclusively demonstrate that competition leads to the exclusion of *D. cespitosa* in the dry meadows.

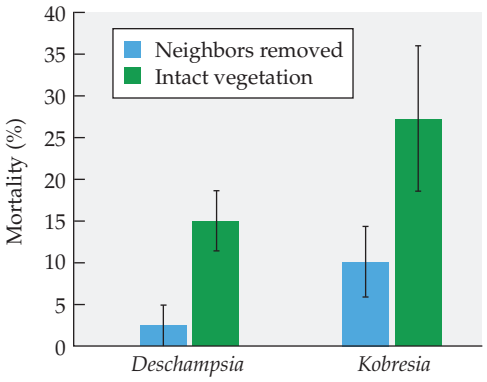



**Figure 10.17** Mortality of *Deschampsia cespitosa* (Poaceae) and *Kobresia myosuroides* (Cyperaceae) in a dry meadow when transplanted into intact vegetation and with neighbors clipped. Values are mean mortality (± 1 standard error) based on four plots, each of which had ten experimental plants per species. (After T. A. Theodose and W. D. Bowman. 1997. *Oikos* 79: 101–114.)

## Research syntheses provide some help in interpreting the evidence

While the results of individual studies are apparently contradictory, there have been several attempts to gain a better overview both experimentally and through synthetic analyses. In a single, coordinated cross-continental field experiment, R. Reader and associates (1994) evaluated the intensity of (diffuse) competition from neighbors for transplanted individuals of *Poa pratensis* (Kentucky bluegrass, Poaceae; see Figure 6.6) at 12 locations in Europe, North America, and Australia. Each site encompassed a range of standing biomass (a surrogate for productivity). The researchers reported the results using two indices of competition intensity, relative competition intensity (RCI) and absolute competition intensity (ACI). There was some suggestion that ACI increased as neighbor biomass increased, but RCI showed no clear relationship to neighbor biomass (Figure 10.18). The researchers concluded that there was no convincing evidence to support the hypothesis that competition increased along a gradient of increasing neighbor biomass, when measured across a wide range of sites and productivities. Since there are flaws in both ACI and RCI, other approaches to measuring competition intensity might offer different interpretations of the results, however.

In a synthetic review of root competition based on trenching experiments in forests worldwide, it was concluded that seedling growth on moist, nutrient-rich soils was largely limited by light and not by root competition for nutrients but that competition for belowground resources was important on infertile soils as well as in more arid habitats (Coomes and Grubb 2000).

Deborah Goldberg and collaborators (1999) took a more quantitative approach to synthesizing results over a greater range to reach more general conclusions than is possible in an individual experiment. They carried out

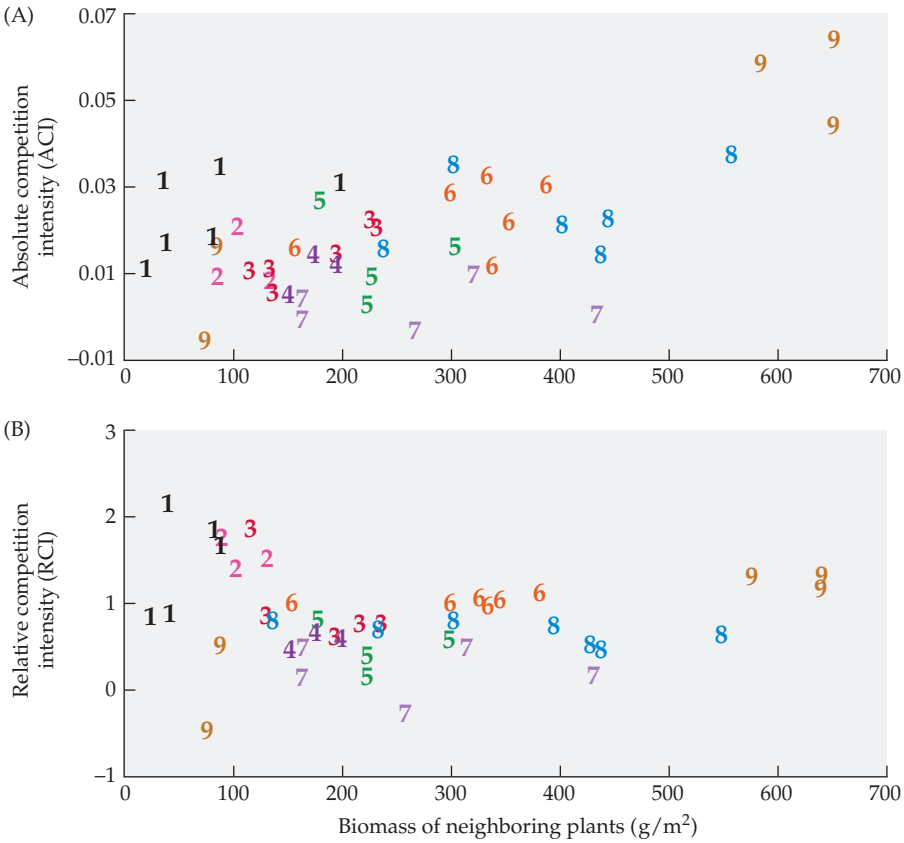



**Figure 10.18** Results from a large set of field competition experiments in which Kentucky bluegrass (*Poa pratensis*) was planted in 44 plots across nine sites in locations across the world. Each point represents a plot; the numbers indicate the particular sites. At each site, competition was studied over a gradient of neighbor biomass, so there is more than one experimental outcome shown for each site. Competition was measured as (A) absolute competition intensity (ACI) and (B) relative competition intensity (RCI). (After R. J. Reader et al. 1994. *Ecology* 75: 1753–1760.)

a **meta-analysis**, or quantitative synthesis (Box 10B), of 14 papers reporting a large number of outcomes of competition experiments. The synthesis examined competition and facilitation among plants along productivity gradients, using vegetation biomass as a surrogate for productivity. Contrary to theoretical predictions, there was a strong negative relationship between competition intensity and productivity when the measure of competition intensity used was the log response ratio (LRR) for both final plant biomass and survival (Figure 10.19). (Grime's theory had predicted a positive relationship, and Tilman's no relationship, as discussed above.) Consistent with Tilman's predictions, no relationship was found for growth. When competition intensity was measured using RCI, they found no clear relationship between RCI and vegetation biomass except, again, a negative relationship for survival. Given the inconsistent outcomes, it is difficult to interpret these results in the context of current theory. Clearly, however, the hypothesis of increasing competition intensity with increasing productivity is not supported as a general pattern, at least when competition is measured using these indices.

The stress-gradient hypothesis addresses a related question from a different perspective. Rather than focusing on productivity gradients, this hypothesis considers environmental stress gradients. The stress-gradient hypothesis predicts that as environmental stress (typically abiotic) increases, competition between plants declines and facilitation becomes more important (Bertness and Callaway 1994). Many papers have been published examining the evidence for the stress-gradient hypothesis. A meta-analysis by Fernando Maestre and his coworkers (2005) led them to conclude that neither facilitation nor competition increased with abiotic stress and suggested that new models are needed. A reanalysis of these data by Christopher Lortie and Ragan Callaway (2006), however, pointed to a different conclusion: they found that facilitation often played a significant role along stress gradients. The difference between these studies was due not to statistical technique, but rather to differences in the criteria for inclusion in the meta-analysis. Maestre and colleagues (2005) did not quantify gradient lengths, and these often differed among studies. Other criteria that differed in the two meta-analyses were that a number of studies included in the original meta-analyses did not involve stress gradients, focused on invasive species, or were not peer-reviewed. This points to the importance of clear, explicit, biologically meaningful criteria for selecting studies in meta-analyses. It also underlines the importance of reading studies carefully rather than just accepting their conclusions!

## Can we resolve the conflicting results?

What explains the lack of agreement among these studies, and what can be done to achieve a better understanding of the importance of competition over a range of environments? Rees's model-based analysis

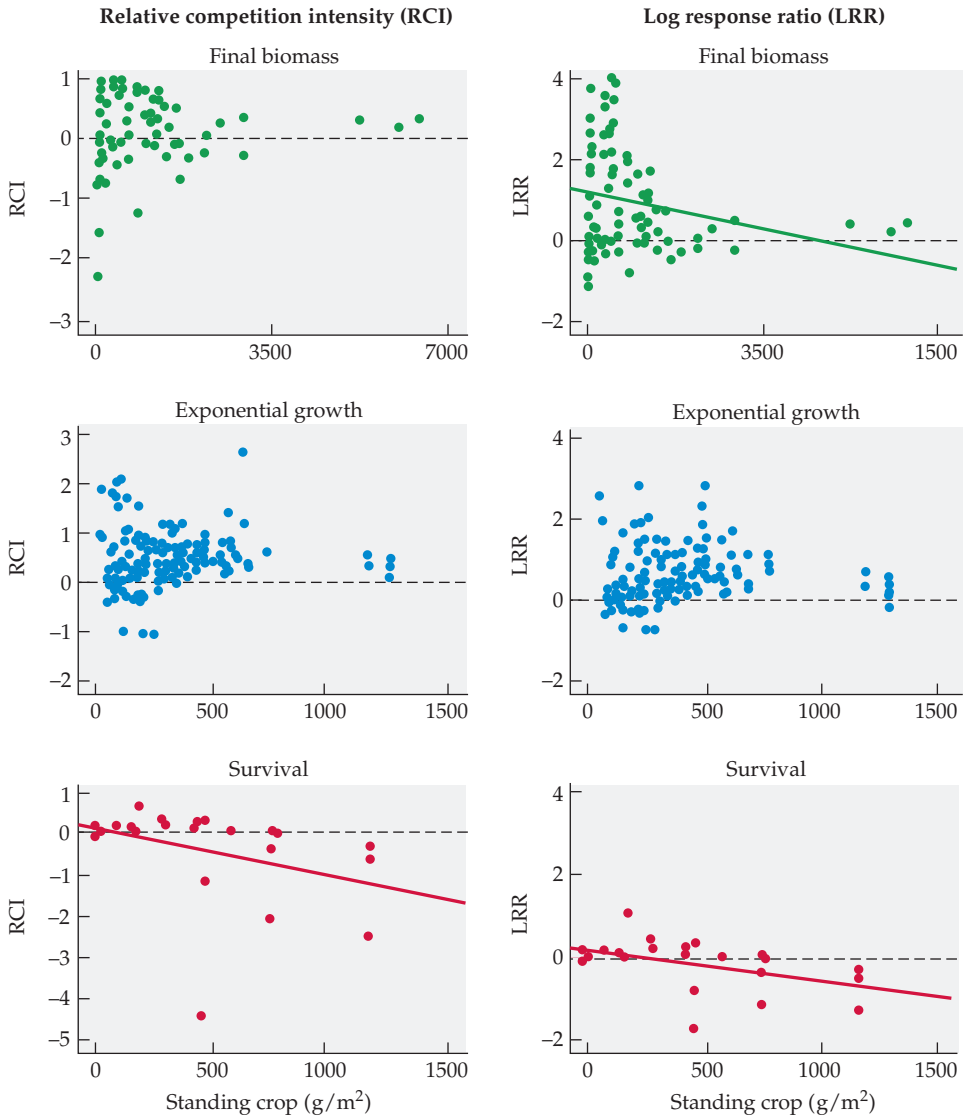



**Figure 10.19** Values for two measures of response to competitors—relative competition intensity (RCI) and log response ratio (LRR)—across 14 separate published studies over a range of community productivity (estimated by standing crop, g/m²). Regression lines are shown only where there was a statistically significant relationship between the variables. Positive values of RCI or LRR indicate that competition is occurring, and negative values indicate that neighbors have a beneficial effect on target plant performance. (After D. E. Goldberg et al. 1999. *Ecology* 80: 1118–1131.)

suggests that a key problem is the overly general nature of the predictions from both Grime and Tilman. The fact that a large number of studies have been conducted using metrics for species interactions that are arbitrary, logically flawed, and statistically questionable may also contribute to conflicting results. Performing experiments in a relatively consistent way across many sites, as Reader and colleagues did, is a big step in the right direction, as is using modern quantitative synthesis techniques to bring together large numbers of independent studies, as Goldberg and her colleagues, and Lortie and Callaway did. It is clear from these efforts at synthesis that conclusions can vary greatly depending on how the experiments are designed, how long they are maintained, what is measured, and how the results are analyzed.

Some of the inconsistencies among the conclusions of different researchers are almost certainly a result of these artifacts. More profoundly, most studies on plant competition look at individual growth (or other measures of the responses of individual plants) in response to competition, rather than estimating population responses. Comparisons of RCI, for example, cannot show whether competition restricts where a species is found; this index simply does not provide enough information. If the question is what limits the distributions of species or determines community composition, studying population-level responses is the only way to get meaningful results. A necessary step is to distinguish between the importance or significance of competitive interactions and their intensity. Competition is *important* or ecologically significant at the community level if it plays a major role in determining community composition. Ecologically significant competitive interactions play a substantial role in determining species coexistence within a community. Competition can be *intense*—that is, it can have major effects on various aspects of individuals' performances—without having much effect at the community or even at the population level.

## BOX 10B Research Synthesis, Systematic Reviews, and Meta-Analysis: Tools for Summarizing Results across Studies

In this chapter, and throughout this book, we often refer to results based on meta-analyses. What are they and where did they come from? Modern research synthesis is the application of the scientific method to the review of scientific studies on a particular question. It may surprise you to realize that ordinary narrative reviews are not repeatable, transparent, or unbiased, and they follow no particular protocol. Research syntheses utilize systematic reviews—a rigorous approach to searching and selecting studies for review—and meta-analysis, the statistical synthesis of the results of a set of studies addressing a particular scientific question. These approaches were first developed in the social sciences and medicine (and are the cornerstone of evidence-based medicine) and entered ecology in the early 1990s. In ecology, as in other disciplines, meta-analysis was controversial when first introduced, but this powerful methodology has now become an important part of the ecologist's toolkit. Many statistical and methodological innovations have been made in meta-analysis since it was first introduced to ecology (Koricheva et al. 2013). Meta-analysis depends on determining an effect size for each dependent variable in a study that summarizes the outcomes, and then analyzing them across all of the studies. The more precise studies, usually those with larger samples, are weighted more heavily than less precise ones. The meta-analysis (and more complex meta-regression approaches) determines how heterogeneous the outcomes of the different studies are and the factors that may explain that heterogeneity. Because heterogeneity is so ubiquitous in ecology (see Chapter 1), meta-analyses in ecology typically find a great deal of heterogeneity among studies, while also identifying and quantifying sources of commonality (Gurevitch et al. 2018). For example, a meta-analysis might find that despite variation among studies, one group of organisms responds one way to an experimental treatment, while another group responds quite differently. The meta-analyses and meta-regressions cited in this chapter and others in this book offer examples of the wide range of questions that such statistical analyses might address in summarizing the outcomes of multiple studies. In all cases, a meta-analyst seeks to determine what a group of studies on a particular ecological question or set of questions tell us, following rigorous statistical methods.

### Models of plant competition can help us to better understand competitive processes and the role of competition in species coexistence

One way to examine the role of competition in community composition is to develop and test multispecies models. A model that describes competition in additive and response-surface experimental designs was developed by Andrew Watkinson and his colleagues (Firbank and Watkinson 1985; Law and Watkinson 1987). This model is an extension of Equation 10.1, for two-species mixtures. In this two-species model, the intensity of competition is estimated by the competition coefficients α and β. For species A, mean weight per plant in mixture, $w$, is given as

$$w_A = \frac{w_{mA}}{1 + a_A(N_A + \alpha N_B)^b} \tag{10.7}$$

and density-dependent mortality is described by

$$N_{sA} = \frac{N_{iA}}{1 + m_A(N_{iA} + \beta N_{iB})} \tag{10.8}$$

where the subscripts A and B specify which of the two species a given variable represents; $w_m$ is the average weight per plant in the absence of competition (that is, $w_{mA}$ is the mean weight for species A in absence of competition); $a_A$ and $b$ are fitted parameters (see Equation 10.1); α and β are competition coefficients that measure the average effect of an individual of species B on an individual of species A, and vice versa; and $N_i$ and $N_s$ are the initial and final densities (Firbank and Watkinson 1990). Analogous equations are used for species B.

In this model, the mean weight for species A is its mean weight in the absence of competition, divided by 1 plus the total competitive effects of neighbors. The $(N_A + \alpha N_B)$ term gives the standardized number of neighbors because α allows us to equate individuals of the two species in terms of their effects on species A. Raising this term to the power $b$ allows density to have nonlinear effects, and the entire quantity is multiplied by $a_A$, which measures the effect of these standardized neighbors on mean weight. The equation for numbers of individuals is analogous.

R. Law and Andrew Watkinson (1987) extended this model by relaxing the requirement for a constant value for the competition coefficients and allowing those coefficients to vary with the frequency and density of the two species. Their model describes rather than predicts competitive outcomes because the outcomes are determined by fitting parameters empirically—the values used in the equations are determined by the results obtained in the experiment. Thus, while this model has been used extensively to describe competitive interactions, it does not provide a set of testable predictions about species coexistence.

Conceptual and verbal models have also helped to elucidate the nature of species coexistence in communities. Alexander Watt (1947) introduced the idea of the plant community as a mosaic of patches. These patches, he maintained, are dynamically related to one another. There is a tension between predictable order and chance events that tend to disrupt that order, which together result in the structure of the community. Competitive interactions played a major role in Watt's view of patch dynamics (Figure 10.20). Watt specifically excluded successional change from this conceptual model of the maintenance of community structure.

Many contemporary views of plant communities are remarkably similar to Watt's perspective (Pickett and White 1985). Based on data showing much predictability in which species (or open space) replace one another in desert communities, Joseph McAuliffe (1988) used a transition matrix model (mathematically similar to those used to analyze population growth; see Chapter 8) to predict the long-term fraction of cover of each species. Eddy van der Maarel and Martin Sykes (1993) extended Watt's perspective in the carousel model they developed to describe the dynamics of species on very small scales in species-rich grasslands on the island of Öland in southern Sweden.

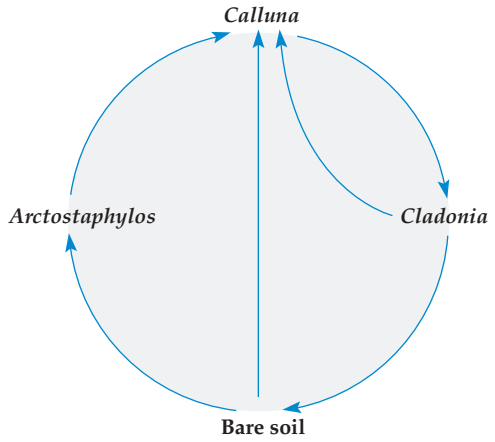
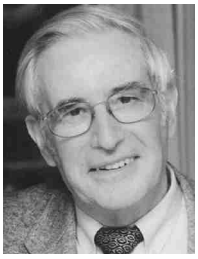



**Figure 10.20** Watt's model of vegetation dynamics in the dwarf callunetum. This plant community, at about 800 m altitude in the Cairngorm Mountains of Scotland, has a natural pattern of double strips of *Calluna* (heather, Ericaceae) and *Arctostaphylos* (bearberry, Ericaceae) separated by windswept bare soil. Watt explained the dynamics of this pattern as follows: The young shoots of *Calluna* exclude lichens, but as they age, their competitive ability diminishes, and the area is invaded by *Cladonia* (a mat-forming lichen). The constant wind prevents the *Cladonia* mat from persisting, however, and in time it disintegrates, leaving bare soil. *Arctostaphylos* then invades by vegetative spread, eventually achieving complete occupation of the patch. Young shoots of *Calluna* in time spread in from the margins, eventually competitively replacing the *Arctostaphylos*. (After A. S. Watt. 1947. *J Ecol* 35: 1–22.)

They emphasized the small-scale mobility of plants, both as individuals, as parts of plants die and shoots colonize adjacent gaps, and as propagules. The *carousel* is the cycle of species replacement, which may go around quickly or slowly.


Courtesy of Eddy van der Maarel

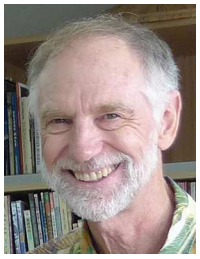
Eddy van der Maarel

Cellular automata models represent a different approach to modeling competition spatially. Cellular automata are computer simulations of species dynamics on a regular, uniform grid, in which interactions occur among individuals located on adjacent positions on the grid. These models have been used for a variety of processes at a variety of scales, from spatially dependent molecular interactions to the development of spiral galaxies, as well as spatially dependent ecological processes. Heiko Balzter and colleagues (1998) demonstrated the usefulness of this approach for plant communities with a preliminary model of the dynamics of three plant species—*Lolium perenne* (ryegrass, Poaceae), *Trifolium repens* (white clover, Fabaceae), and *Glechoma hederacea* (ground ivy, Lamiaceae)—growing in a lawn over a 3-year period. This approach may eventually prove valuable in making predictions about species' abundances and conditions for coexistence in a spatially explicit context.

## Modern coexistence theory is a framework for understanding how competition affects coexistence

Competition can affect where plants live, but how is it that so many species of plants—all competing for combinations of the same basic resources—coexist in many communities? It would be hopeless to do empirical experiments like the ones discussed above on even a small portion of the communities in the world, to try to understand even a tiny fraction of Earth's plant communities; we need models to try to form a general picture and to help us organize our thinking. Models of plant competition seek to explain the vast diversity of plant life by identifying the circumstances under which competitors can coexist. Some models also try to explain the mechanisms by which competitors interact and the characteristics of winners and losers. Many basic textbooks in ecology use a single model of competition, the Lotka-Volterra model of two-species competition, to introduce a discussion of the way species compete. While some ecologists argue that this approach is useful for understanding the outcome of local competition, many plant ecologists believe it is not very useful for describing competition among plants—among other reasons, because plants do not experience the average effects of density in the same way mobile animals do (see Chapter 8).

While many different sorts of models have been developed to understand competition, most of them can


Courtesy of Peter Chesson

Peter Chesson

be understood by using modern coexistence theory, an overarching framework developed by Peter Chesson (2000, 2018) with several collaborators. The theory itself can involve mathematical techniques beyond the scope of this book, but we present a verbal outline here.

In a community of multiple species, if each can grow when at low population density, those species can coexist. This condition for adding new species to a community, or genotypes to a population (Turelli 1978), arises in many contexts in ecology and in evolutionary biology, including conditions in which multiple genotypes or phenotypes can be maintained in populations (Grainger et al. 2019). (Notice that in many such papers, *invasion* does not refer to non-native species that spread and become dominant; *invasion* here means adding any species that increases diversity to a community.) Chesson showed that for many competition models there are two components to the contribution an average individual makes to the population growth rate when the population is small (the low-density per capita growth rate). The first component is the differences in average fitness between each species and its competitors; these are called **equalizing mechanisms**. Anything that reduces the differences among average fitnesses makes it easier for a population to have a positive low-density per capita population growth rate. The second component, **stabilizing mechanisms**, consists of the differences between the niches of the competing species. Anything that increases the differences between species' niches tends to have the effect of causing each population to limit its own growth more than that of other species—thus contributing to coexistence. Both equalizing and stabilizing mechanisms play important roles in coexistence. Particular models developed by plant ecologists have taken a variety of approaches to modeling competition using this theoretical framework among plants.

## Models within the framework of modern coexistence theory have stimulated research and discovery

Models of the community-level consequences of plant resource uptake were first proposed by Robert MacArthur (1972) to contrast with the Lotka-Volterra competition model. The latter model treats the growth of each population as a function of the densities of the two populations, while MacArthur's model treats population growth as a function of resource availability. In turn, resource availability depends on the rate at which resources are supplied (say, by chemical weathering of phosphate-laden rock) and the rates at which they are taken up by the competing species. MacArthur showed that coexistence occurred under certain combinations of resource uptake and supply rates. If competition is for multiple resources (e.g., different soil nutrients), stable coexistence is predicted when competing plant populations are competitively superior at different ratios of essential resources (if several other conditions are also met). This theory, popularized by David Tilman (1982) and called the resource-ratio model, has been very influential in plant ecology, but its predictions are not well supported by experimental tests. In seeking to explain natural patterns of multispecies coexistence, Tilman extended this model to account for the coexistence of more than two species by including patches of different resource availability. The differing growth rates of the competitors, which are consequences of differences in their $R^*$ values, are fitness differences. Because there are no stabilizing mechanisms, the model leads to competitive exclusion unless the species have identical $R^*$.

In a review of the literature, Thomas Miller and his collaborators (2005) found that although over 1300 studies cited Tilman's central papers, only 26 provided well-designed experimental tests of the seven theoretical predictions of the resource-ratio model (for a total of 42 tests). One prediction—that species dominance should change with the ratio of resource availabilities—was supported in 13 of 16 experimental tests (many of them on aquatic species rather than on terrestrial plants), but there were too few tests to begin to draw conclusions about the other six predictions. More experiments that explicitly test the predictions of the model are needed. A response by J. Bastow Wilson and his coworkers (2007) disputed the numbers of papers reviewed by Miller and colleagues that provided support for the resource-ratio hypothesis but pointed out that support for the Tilman model was confined to bacterial and phytoplankton studies. In any event, it seems clear that the model has been cited far more than it has been tested.

In more recent models, species coexistence may not be determined by conditions at equilibrium. Instead, the outcome of competition is determined by stochastic or nonlinear dynamics. Nonlinear models, for example, can have a fully deterministic outcome (entirely determined by the terms in the model, with no random elements), but this outcome may not be a single equilibrium point, but instead result in population cycles or chaos (a technical and mathematical topic that is fascinating but beyond the scope of this book). In stochastic models, chance events (such as weather fluctuations) have a key role in determining competitive outcomes.

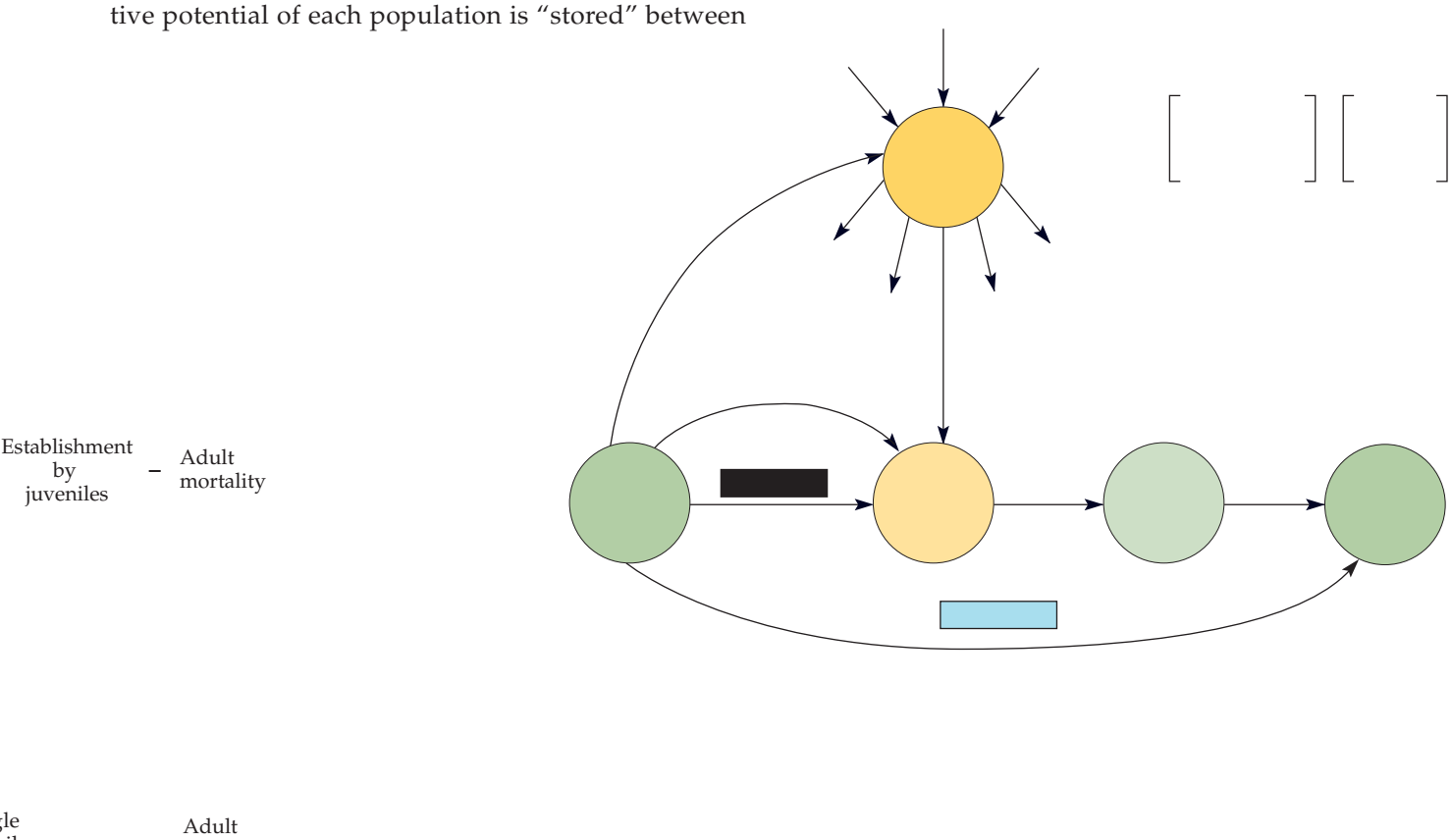
Peter Sale (1977) first informally introduced the *lottery model* as an explanation for the coexistence of fish species on coral reefs, but this approach has subsequently been used for understanding coexistence of many different kinds of species, including plants. In this model, newly recruited fish of various species

randomly obtain a feeding territory on a reef on a first-come, first-served basis; this "lottery" for space allows the maintenance of diversity. Lottery models have encouraged many ecologists to think about the importance of spatial and temporal variation in species coexistence. When the lottery model was fully analyzed mathematically by Peter Chesson and Robert Warner (1981), this approach became very influential in ecologists' thinking. It has stimulated both empirical and theoretical research, especially for plant communities. An intriguing contribution to this work is a lottery model focusing on spatial patchiness in plant communities developed by Avi Shmida and Stephen Ellner (1984). In addition to the assumption that the outcome of competition among juveniles for space is determined by chance, their model incorporated asymmetric competition (juveniles cannot displace adults) for microsites, nonuniform seed dispersal, and larger-scale spatial and temporal patchiness (Figure 10.21). An important result was that it predicted species coexistence without requiring differences in habitat or resource use (that is, without differentiation in the species' niches).

Can temporal, as well as spatial, variation lead to species coexistence? Peter Chesson and his collaborators have shown that the answer is yes, under some circumstances. If recruitment (the establishment of seedlings) of competing species responds differently to environmental fluctuations—so that, say, one species largely becomes established only in years with warm, wet summers and the other in years with cold, rainy springs—the species can coexist stably. This phenomenon is called the **storage effect** (Warner and Chesson 1985; Chesson and Huntly 1989). The reproductive potential of each population is "stored" between generations, and the different **temporal niches** of the species allow their coexistence. This idea of temporal niches is similar to Grubb's (1977) idea of a regeneration niche (discussed above). Storage can occur in any population with age structure, such as perennial plants or annual plants with seed banks, as long as there is sufficient temporal variation with periods of time that favor each species. A number of studies of desert annuals have suggested that the storage effect may help to explain the diversity of these communities (Ehleringer et al. 1991; Pake and Venable 1995; Chesson et al. 2004).

Lottery models assume that competition is **diffuse**—that is, that many species are competing simultaneously, so the competition between any particular pair of species is fairly weak. If some species are very similar to one another, they may compete much more strongly. Colleen Kelly and Michael Bowler (2005) studied the consequences for coexistence when competition is focused (two species compete directly with each other) rather than diffuse. Using a lottery model in which one species is more sensitive to environmental variation than the other, they predicted that the less common (more sensitive) species of a competing pair should be more aggregated spatially. According to their model, coexistence can occur if the less common species is also the better competitor. Results from studies of a number of pairs of related tree species in dry tropical forests in Mexico support both predictions, as well as others from their model (Kelly and Bowler 2002, 2005).

A distinctly different kind of model of species interactions is the neutral model (Hubbell 2001). Adapted largely from models of genetic drift (see Chapter 9),



the neutral model begins with the baseline assumption that coexisting species are ecologically equivalent: adding an individual to a community from any competing species has exactly the same effect as adding an individual from any other species (including members of one's own species). The neutral model is, then, a case in which there are no stabilizing mechanisms and the species have identical average fitnesses (Adler et al. 2007). Some key ideas in the neutral model are derived from the theory of island biogeography (see Chapter 15), in which all species in a regional pool are treated as though they have the same chance of dispersing to a new island or of going extinct on the island if they are already there. If the species are competitively equivalent, long-term community composition is mainly determined by speciation, randomly occurring extinction, slow average extinction rates (due to the species' competitive equivalence), and dispersal.

Observations of the relative abundances of tree species in tropical rainforests provided early motivation for the neutral model. These communities typically have many rare but few common species, as shown when species are ranked in order of abundance and rank is plotted against actual numerical abundance (see Figure 14.10). The neutral model predicts exactly such rank-abundance curves. The problem is, many non-neutral models also make predictions that are indistinguishable from this pattern (McGill 2003) (Figure 10.22). Many ecologists have analyzed their data sets on relative abundances of species in light of the neutral model, although there has been considerable debate about its validity (Clark and McLachlan 2003; Volkov et al. 2004; Etienne and Olff 2005; Purves and Pacala 2005). If the predictions of models are indistinguishable from one another, one can evaluate them only by directly testing their assumptions. There is considerable evidence against the central assumption of the neutral model that species in a community are competitively equivalent (Purves and Pacala 2005). While there is still disagreement about whether neutral models can be useful, a paper by a number of prominent neutral theorists (Rosindell et al. 2012), in discussing uses of the neutral theory, agreed with other ecologists that the "real world is not neutral." Most interest in the neutral model is in using it within the context of modern coexistence theory to understand what stabilizing and equalizing mechanisms there are, and how important each is in permitting stable coexistence.

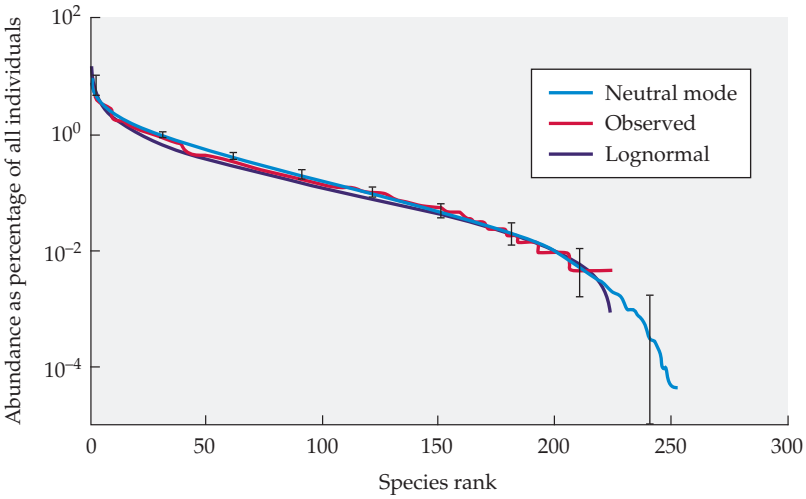



**Figure 10.22** Rank-abundance curves for observed abundances of tree species on Barro Colorado Island, Panama, and the best-fitting curves under the neutral model and an alternative model assuming a lognormal distribution. The error bars are for 1 standard deviation around the predicted value for the neutral model. Both models mimic the data, but the lognormal provides a better fit. (After B. J. McGill. 2003. *Nature* 422: 881–885.)

## New research can extend our understanding of coexistence

Within existing communities, is intraspecific competition typically stronger than interspecific competition? It turns out that this is not a simple question. Peter Adler and his coworkers (2018) reviewed studies of plant competition and found 26 that adequately quantified both intra- and interspecific competition, and in which both had negative effects. Their analysis concluded that intraspecific competition was on average four to five times stronger than interspecific competition—strong evidence for the utility of modern coexistence theory. However, Matteo Detto and his associates (2019) criticized the way Adler and colleagues carried out their analyses, so the conclusion may be premature. In a related study, Sean Tuck and his collaborators (2018) conducted an experiment with a community of European sand dune annuals that had previously been the subject of a number of studies (e.g., Mack and Harper 1977). They fit models for population growth of the study species and then asked how well those models predicted the responses of each species to removal of competing species. Their models underpredicted those responses by at least 50%. The researchers suggested that this underestimate was because the estimates were made only where the species are actually found. If species are confined to realized niches by competition, competition will be estimated as much weaker than it actually is because it is not observed in other microhabitats. Both revised statistical methods and new experimental and analytic approaches may be needed to properly assess the strength of competition.

New theoretical challenges are also on researchers' agendas. Peter Chesson (2018) outlined a list of important problems for modern coexistence theory, including the need to extend the theory to multispecies interactions and to include facilitation and other types of interactions, such as natural enemies. Stephen Ellner and his coworkers (2019) have proposed methods that may make it easier to apply the theory in empirical settings.

The development of modern coexistence theory has led to advances in ecological network theory (Losapio et al. 2019), which strives to understand and analyze interactions among multiple species simultaneously. Network theory grew out of the mathematics of graph theory, and these methods are used in many settings beyond ecology. Ecological networks have considered trophic and plant-pollinator networks. Several studies have analyzed competitive and facilitative interactions in communities, including plant-plant interactions (Godoy 2019; Losapio et al. 2019). One property of networks that has been tackled is the question of transitivity. If species A outcompetes B, and B outcompetes C, it is natural to assume that A outcompetes C. However, that is not always the case—this sort of transitive relationship applies in the mathematical systems we are all familiar with ($X > Y$ and $Y > Z$ implies $X > Z$), but real plant communities can be much more complicated! Intransitivity in a network, in contrast, implies that there is no such strict hierarchy: species A may be competitively dominant over species B, and species B over species C, but species C may be competitively superior to species A. In any case, where it does occur, intransitivity will act as a stabilizing mechanism; Chesson (2018) outlined methods for analyzing intransitive competition with modern coexistence theory. A study by Santiago Soliveres and his collaborators (2015) concluded that intransitive competition is very widespread, but this has been questioned because the study was based on patterns of co-occurrence in natural communities, not on direct evidence from experimental studies.

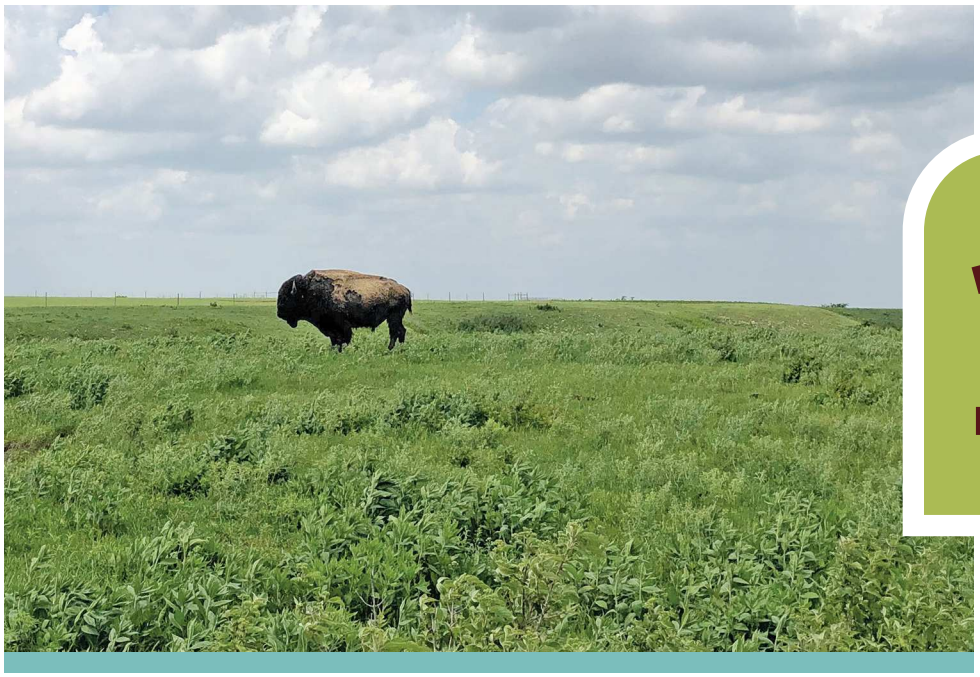
A unique approach to addressing the role of transitivity in plant communities was carried out by Nicole Kinlock (2019), who constructed a meta-analysis of networks from experimental studies of plant interactions . Kinlock's analysis provides a more convincing conclusion: she found some intransitive networks (as well as some cases of facilitation), but they were relatively uncommon in plant communities. She also found strong competitive interactions in many networks, but facilitation also occurred, particularly in indirect interactions and even in less stressful environments. Patterns of connectance and imbalance among competitors were complex and nuanced. Overall, this has been a very rich and promising approach to the analysis of multispecies communities in the context of modern coexistence theory.

Studies of competition and coexistence usually treat species as though they are composed of identical individuals, at least in their potential. This is, of course, an approximation, and factors like genetic variation and variation among microsites can actually render this approximation misleading. James Clark and his coworkers (Clark 2010; Clark et al. 2010, 2011) examined large data sets on individual variation within a number of species of North American trees and showed that estimates of competition based on the average performance of individuals within a species can severely underestimate the actual competitive effect of that population on another species. In addition to changing estimates of the strength of interspecific competition, such variation within species also can magnify the strength of intraspecific competition. We expect that further studies along these lines will advance our understanding and promise further insight.

Competition and its consequences have been studied for a long time. Much is known about mechanisms of competition between individuals, and conceptual advances have led to new theoretical approaches to the study of coexistence that point pathways to answering some old and thorny problems in ecology. This is an exciting time in the study of mechanisms that build plant communities.

## Summary

- Competition is defined as a reduction in fitness due to shared use of a limiting resource. This reduces the growth rate of populations.
- Competition among seedlings often results not only in greatly reduced average sizes, but also in highly unequal size distributions among individuals, called size hierarchies, which ultimately translate into large differences in survival and reproduction.
- The resources for which plants compete include light, water, mineral nutrients, space, pollinators, and seed dispersers. Different mechanisms are involved in competition for different resources. In addition, plants may compete for sites in which to establish and grow.
- Most resource competition occurs among adjacent individuals. Thus, local density, not the total density of a population or community, determines the intensity of competition for an individual.
- In most settings, larger individuals have a competitive advantage over their smaller neighbors.
- Many different experimental designs have been used to assess competition, including various ways of manipulating densities and frequencies. These experiments have been carried out in greenhouses, transplant gardens, and natural populations.
- Plant strategies, characteristic suites of traits that are hypothesized to be most advantageous under particular sets of circumstances, may play a role in competitive outcomes in different environments and may involve trade-offs among traits that confer advantages under different conditions.
- Allelopathy—the release of chemicals to inhibit or kill competing neighbors—has long been believed to be an important mechanism of plant interactions and has been hypothesized to explain some species distributions, but experimental evidence in the field for allelopathy as an important driver of community patterns has been limited to date.
- Plants can have positive interactions with neighbors (facilitation). Some plants provide physical environments that permit plants of other species to establish nearby.
- Common mycorrhizal networks (CMNs) can provide living fungal links between plants of the same and different species in communities, and they have been hypothesized to enable facilitative interactions between plants.
- Ecologists have used different approaches to modeling competition among plants and its consequences.
- Modern coexistence theory provides an overarching framework under which many models of coexistence can be analyzed. One set of models, the neutral theory, is a particular case within this framework. There is now considerable agreement that communities are not actually neutral, although there is still disagreement about the utility of the neutral theory.
- Empirical studies of the relative strength of interspecific and intraspecific competition—a central prediction of modern coexistence theory—are ongoing, and they present a number of practical challenges.



Courtesy of J. Gurevitch

11

# Herbivory and Other Trophic Interactions

Green plants are the foundation of almost all terrestrial food webs. In Chapter 10, we looked at the ways plants compete with other plants and the consequences of plant competition. Plants also have a wide range of trophic interactions with a wide diversity of other organisms. Trophic interactions involve consumption of, or dependence on, energy and nutrients by one group of organisms on another group of organisms. Plants are eaten by **herbivores**—organisms that rely on eating plants for nutrients and energy—and they may be infected by pathogens or live symbiotically with endophytic microorganisms. They form symbioses with fungi called mycorrhizae, and they may parasitize or be parasitized by other plants. We touch on some of the ways in which plants interact with other organisms in other chapters—with nitrogen-fixing bacteria (see Chapter 4), pollinators (see Chapter 8), and seed dispersers (see Chapter 13). In this chapter we consider how plants interact with and defend themselves from herbivores and pathogens, the crucial role of mycorrhizal interactions for terrestrial life, and other trophic interactions between plants and other organisms. We begin with that most basic of interactions, animals eating plants.

All animals (including humans, of course) ultimately depend on plants (or plant eaters) for food and therefore for their existence. Yet, as pointed out in a very short and very famous paper in 1960 (Hairston et al. 1960), even casual observation seems to reveal a green world teeming with (generally uneaten) plants. Why is the world so green? Conversely, we might ask, what are the consequences of herbivory from the plants' perspective? As we will see, plants do a lot to control how much they are eaten and who they are eaten by. We will return to the question posed by Nelson Hairston and colleagues later in this chapter.

Above: The American bison is the symbol of the American West and the quintessential herbivore of the Great Plains. Large and small herbivores shape grasslands and other plant communities in numerous ways.

**Herbivory** is the consumption of all or part of a living plant. Some ecologists use the term *predation* when an herbivore eats and kills an individual. **Seed predators**, or **granivores**, are herbivores that consume seeds or grains (the one-seeded fruit of grasses—plants in the family Poaceae), killing the individual embryonic plant within. **Grazers** are herbivores that eat grasses and other low-growing plants (generally not killing them, with some important exceptions), while **browsers** eat leaves from trees or shrubs. **Frugivores** are herbivores that consume fruits, either with or without damage to the seeds.

Plants are consumed by organisms from a variety of kingdoms: animals, fungi, bacteria, and even other plants. Generally, though, we mean animals eating plants when we talk about herbivory, although the animals can range from nematodes to insects to birds and mammals. Herbivory can have ecological effects at the level of the individual plant, the population, the community and landscape, and the ecosystem. For instance, herbivores can alter patterns of coexistence of plant species, influence the geographic range of a species, and change the rate of nutrient cycling. Herbivory can also influence the evolution of plants. Ecologically and evolutionarily, some of the most important herbivores are grazing mammals and insects. However, other types of herbivores, such as birds, mollusks, and nematodes, can be very important in particular systems.

We begin this chapter by examining the effects of herbivory on individuals, populations, and communities, briefly touch on landscape effects, look at the complex topic of how plants defend themselves against herbivores and pathogens, and consider how those defenses may have evolved. People have used herbivores to control populations of undesirable plants—called biological control—and we will take a look at some of the consequences of those efforts. We also consider some of the evolutionary consequences of the interactions between plants and herbivores.

## 11.1 The Effects of Herbivores on Individual Plants Depend on What Is Eaten

What do herbivores do to individual plants? Herbivores can consume an entire plant, causing the death of the individual. By eating seeds, granivores—such as some ants, rodents, and birds—kill individual (embryonic) plants. Alternatively, herbivores may eat only some parts of a plant, damaging, removing, or destroying those parts, but not necessarily killing the plant. White-tailed deer (*Odocoileus virginianus*, Figure 11.1A) in the northeastern and midwestern U.S., for example, selectively prefer the newest leaves or parts of shoots of woody species when food is abundant in summer but will eat buds, twigs, acorns, and even tree bark in winter when food is scarce. These choices can have very different impacts on the plants they feed on and on forest communities. Insects may eat the least toxic leaves or parts of leaves. Herbivores can also live on or within a plant and consume some of the plant's resources. For example, aphids are insects that extract dissolved sugars and other nutrients directly from the phloem by penetrating it with their mouthparts (Figure 11.1B). Leaf miners are insect larvae of various species of moth or butterfly that live inside leaves, consuming the internal tissues (Figure 11.1C). Gall makers cause plants to change their development to form specialized structures within which larvae can grow and feed (Figure 11.1D). Both of these latter feeding strategies are ways that the herbivore can avoid being eaten itself. Pathogenic microorganisms can parasitize a plant, depleting its resources over time. Some plants are parasites on other plants, tapping water, sugars, proteins, and other resources for their own use.

The effects of herbivory on a plant depend, among other things, on what parts of the plant are consumed. Removal of or damage to roots can reduce or prevent the plant's uptake of water and mineral nutrients and can make the plant more vulnerable to being toppled by wind, flooding, or soil erosion. Consumption of leaves reduces the photosynthetic surface area, while removal of phloem sap may reduce the energy and materials available for growth and reproduction. Consumption of leaves, stems, and twigs may alter the competitive relationships among neighboring plants. Removal of meristems may alter the growth form of the plant. If the apical meristem of an herbaceous plant is eaten, this can change the plant form, from tall and straight to low and bushy, which can make it vulnerable to shading by neighbors but also make it less likely to be blown over. Consumption of flowers, fruits, and seeds may reduce the plant's fitness by reducing its potential contribution to the next generation. Of course, to the new individual in each seed, consumption means death. Alternatively, fruits are often consumed without damage to well-protected seeds with hard seed coats, in which case the frugivore may disperse the seeds to potentially favorable locations by defecating them far from the parent plant.

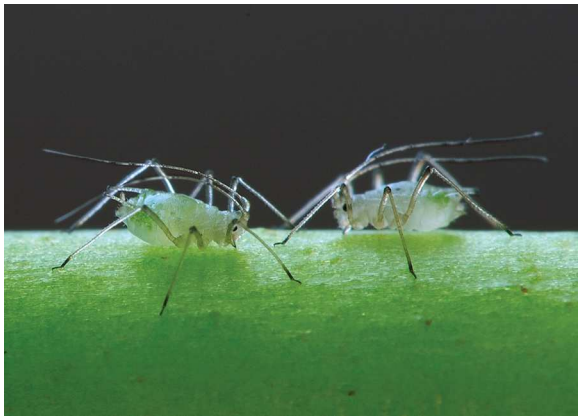
The life history stage at which the plant is attacked or damaged is also important. Seedlings are particularly vulnerable to herbivores. One mouthful for the herbivore can kill a seedling but have little effect on a more mature plant. Grazing on grasses that have just begun to flower can critically affect their ability to produce seeds, whereas heavier grazing after seeds have been shed may have less of an effect on population dynamics.

There are two common adaptive responses to herbivory: resistance and tolerance. **Resistance** to herbivory is the ability of a plant to avoid being eaten; **tolerance**

(A)

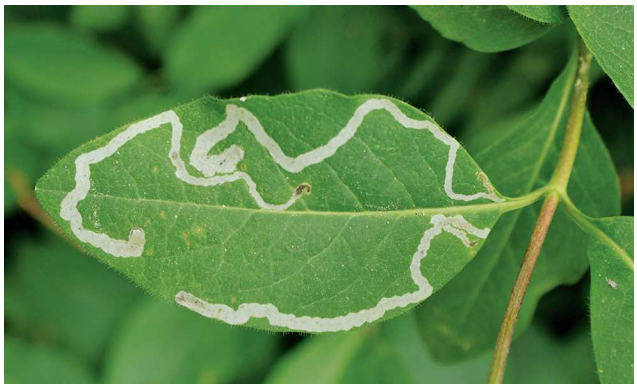
Clinton Robertson and Charles Robertson/CC BY-SA 2.0

(B)



Luc Viatour/CC BY SA 3.0

(C)



Kenraiz Krzysztof Ziarnek/CC BY-SA 4.0

(D)

Courtesy of Warren G. Abrahamson

**Figure 11.1** (A) The large populations of white-tailed deer are creating very large changes in the forest of the eastern U.S. (B) Aphids are small sap-sucking insects in the order Hemiptera. (C) Trails produced by a leaf-mining fly larva in *Lonicera periclymenum* (European honeysuckle, Caprifoliaceae). (D) A gall on the stem of *Solidago altissima* (tall goldenrod, Asteraceae) made by the larva of the fly *Eurosta solidaginis*.

of herbivory is the ability to minimize reductions in fitness due to herbivory. Jake Weltzin and Guy McPherson (1999) showed that there was a high level of tolerance in seedlings of *Prosopis glandulosa* (honey mesquite, Fabaceae): while artificial herbivory (clipping) reduced biomass production, seedlings that were defoliated five times still had a survival rate of at least 75%. Both tolerance and resistance to herbivory often change over the life of an individual plant, as resource allocations and defenses are altered (Boege and Marquis 2005). We discuss mechanisms of resistance later in this chapter.

How much do herbivores eat? It has been estimated that about 10% of the leaves of forest trees are lost every year to herbivores (Coley and Barone 1996). Herbivory is greatest in dry tropical forests, somewhat less in tropical rainforests, and least in temperate forests. In the tropics, young leaves tend to be eaten more readily than mature leaves, which are likely to be tougher and sometimes more toxic. As you might expect, there is tremendous variation among species, locations, and years in the degree of damage caused by herbivores (discussed below).

Can herbivory ever actually help plants to grow or reproduce? In the 1980s and early 1990s, a group of scientists postulated the existence of **overcompensation**, in which plants purportedly respond to herbivory by growing even more than before an attack (McNaughton 1983). (If a plant's hypothesized extra growth in response to herbivory resulted in no net difference between grazed and ungrazed individuals, it would be called **compensation**.) These researchers suggested that overcompensation was due to the coevolution of plants and herbivores,

particularly grasses and mammalian herbivores. The saliva and urine of bison, for example, were thought to contain substances that stimulated growth in grasses (Detling et al. 1980). These ideas were highly controversial and received a great deal of attention; it seemed difficult for many ecologists to believe that being eaten could actually be a good thing for plants.

While there was some experimental evidence for overcompensation, when taken as a whole, the idea was not well supported by the available data (Belsky 1986; Belsky et al. 1993). One of the possible explanations for reports of overcompensation was that researchers had measured only aboveground plant parts, while underground reserves may have been depleted to stimulate the observed aboveground growth. Long-term herbivory might result in significant depletion of these underground reserves, ultimately harming the ability of the plant to recover from subsequent bouts of herbivory. Another explanation was that in dense plant stands, if herbivores eat only shaded, unproductive understory leaves, there may indeed be no reduction in the photosynthetic capacity of the plant, and thus no negative effects of herbivory. Joy Belsky and her associates (1993) argued that overcompensation seemed to occur mainly in those experimental treatments most favorable to plant growth, such as a combination of high nutrient availability and reduced competition. A meta-analysis of experiments of the effects of herbivory on plant growth and reproduction concluded that although herbivory usually harms individual plants, overcompensation may occur in some cases for grasses and other plants with basal apical meristems growing at high resource levels and for some dicots at low resource levels (Hawkes and Sullivan 2001). While instances of overcompensation continue to be reported, herbivory in most cases reduces both growth and reproductive output, and resource addition increases both growth and reproduction (Figure 11.2).

## 11.2 Herbivores Can Alter Plant Population Composition and Dynamics

When does herbivory affect the rate of growth or decline of a plant population? Can the presence of herbivores—or their exclusion—have a decisive role in determining what plant species are able to coexist in a community? The extent to which herbivores affect plant population dynamics remains a controversial question. What factors control insect herbivore populations was the subject of a heated series of debates during the mid-twentieth century among ecologists (Egerton 2015). On one side, John Nicholson (1933, 1950, 1954) argued that biotic factors controlled insect populations, later generalizing this argument to density-dependent contol. On the other side, Herbert Andrewartha and Charles Birch (Andrewartha and Birch 1954; Andrewartha 1957; Birch 1957) argued that density-independent factors like weather largely controlled insect population numbers. Based on the general observation that "the world is green" (that is, plants are not all consumed), Nelson Hairston, Frederick Smith, and Lawrence Slobodkin (1960) later argued that populations of herbivores had to be limited by their predators—**top-down population control**—rather than by food supply—**bottom-up control**. They also suggested that they might be limited by density-independent, abiotic factors such as weather, but their argument relied only on logic, not on data.

But what actually limits plant populations? Do herbivores determine plant densities in top-down control, or are plants limited from the bottom up by their own resources? The argument depends on which trophic level is being regulated by top-down or bottom-up factors, the plants or the herbivores. The top-down argument is that herbivores are maintained at such low densities by their own predators that they rarely exert negative effects on plant populations (e.g., Strong et al. 1984), while the bottom-up argument makes the case that plant populations are limited by abiotic factors such as water, light,

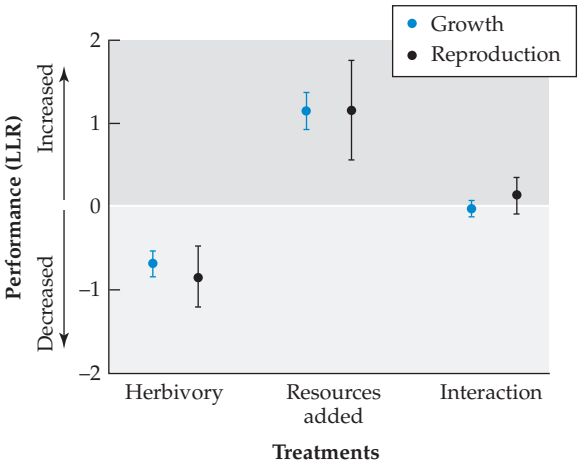



**Figure 11.2** The effect of herbivory on both growth and reproduction is generally negative according to the results of a meta-analysis of the average effects of herbivory on growth across 82 studies, and on reproduction in 24 studies. Plant performance is measured as the log response ratio (LLR) (here, natural log of average performance in experiments with herbivores [or resources] divided by performance in control treatments); error bars are 95% confidence limits. Resource addition improved plant performance, but there is no evidence of an interaction between resource addition and herbivory across studies; adding resources does not change the response to herbivory. (After C. V. Hawkes and J. J. Sullivan. 2001. *Ecology* 82: 2045–2085.)

and soil nutrients, not by herbivores (Hairston et al. 1960; Slobodkin et al. 1967; Hairston and Hairston 1993).

In contrast, others have argued that herbivores do play an important role in controlling plant populations. In all likelihood, top-down herbivory regulates plant population dynamics in some cases, while bottom-up processes dominate in others. In a quantitative synthesis of herbivore-exclusion experiments, Shihong Jia and colleagues (2018) showed that overall, herbivores reduce plant abundance, biomass, survival, and reproduction. Herbivory did not increase the number of plant species on average across studies, but did in temperate grasslands as well as studies where large animals were excluded, animal exclusion resulted in higher numbers of plant species.

One of the most obvious situations in which plant populations are dramatically affected by herbivores is the killing of large stands of forest trees by insects. Outbreaks of lepidopteran larvae can cause massive defoliation, in some instances resulting in tree mortality. Repeated defoliation of oaks in the northeastern and midwestern United States by gypsy moths has caused die-offs of large numbers of trees (Davidson et al. 1999) (see Figure 13.8). Trees in more mesic areas appear to be more likely to die from a single episode of gypsy moth defoliation than those in drier sites.

Bark beetles (family Scolytidae) are an important cause of conifer mortality in western North America and the southeastern United States, now spreading to other regions as climates warm (Figure 11.3) (Powers et al. 1999). Conifer mortality due to bark beetles has also become an increasingly major issue in the coniferous forests of northern, central and eastern Europe and Siberia. While most bark beetles inhabit branches and trunks of trees already undergoing stress, such as those damaged by a lightning strike or drought, some species attack and kill healthy trees. The beetles enter the tree by chewing a hole through the bark into the cambium (the actively growing layer of cells under the bark) and lay their eggs there. After hatching, the larvae feed on the cambium, destroying it and the vascular tissue.

Courtesy of the USDA Forest Service Archives

**Figure 11.3** Scolytid beetles cause extensive damage to conifers in western and southeastern North America. These larval *Dendroctonus pseudotsugae* are living in the cambium of *Pseudotsuga menziesii* (Douglas fir, Pinaceae).

Coniferous trees (particularly pines) usually respond to boring by beetles by oozing sap (also called pitch) into the wound, either suffocating the beetle or killing it with monoterpenes in the sap. These chemical also prevent the beetle eggs from hatching. The monoterpenes are present at low concentrations constitutively, but beetle attack induces production of high, toxic levels. The sap also physically "pitches out," or pushes, the beetles out of the hole. However, a massive attack by a large number of beetles reduces the ability of the tree to "pitch out" the beetles by overwhelming the tree's ability to produce the sap itself as well as the toxic compounds in the sap. The beetles mobilize massive attacks by attracting other beetles with pheromones, while the monoterpenes in the sap inhibit the ability of the beetles to produce the pheromones. Who knew that chemical warfare was so complex and widespread in conifers? Different species of bark beetles often have symbiotic fungal partners that interact in complex ways with the beetles and the trees. The trees' toxic monoterpenes that kill the beetles also inhibit or kill the fungi. We return to the effects of these fungi later in this chapter.

Bark beetles that are responsible for widespread tree mortality in North America include the Mexican bark beetle, the western pine beetle, the spruce beetle, the southern pine beetle and the western balsam bark beetle; the European spruce bark beetle has killed tens of millions of ha of conifers in Europe and Eurasia. As their names suggest, bark beetles tend to be fairly selective, generally specializing on one or a few species of conifers. However, while the southern pine beetle primarily attacks *Pinus echinata* (shortleaf pine, Pinaceae), *P. taeda* (loblolly pine), *P. palustris* (longleaf pine), and *P. elliottii* (slash pine), it has very recently spread northward with warming climates. Warmer temperatures increase winter survival of the beetles and speed up their development. Southern pine beetles have become a major problem in the pine barrens of New Jersey and Long Island, New York, targeting *P. rigida* (pitch pine, Pinaceae, closely related to the other pines this beetle targets), after first appearing in 2001 in New Jersey and 2014 on Long Island. They are now killing hundreds of thousands of trees on tens of thousands of hectares of protected preserves and forests there. The mountain pine beetle (*Dendroctonus ponderosae*) generally attacks *P. contorta* (lodgepole pine), a dominant conifer at lower elevations in the Rocky Mountains of the U.S. and Canada. At higher elevations, *P. albicaulis* (whitebark pine)

dominates. This pine species has generally been free of mountain pine beetle attacks because the colder winters kill the beetles. Ken Raffa and his associates (2013) discovered that climate change has made whitebark pine vulnerable to attack by the beetles as temperatures warm and the beetles move up in elevation, much as southern pine beetles are moving northward in the eastern U.S. However, whitebark pine only recently has been exposed to the beetles. Unlike lodgepole pine, which evolved with bark beetle attacks over many generations, whitebark pine does not release as much resin when attacked, and its resin has lower concentrations of the toxic monoterpenes. Therefore, it is much more vulnerable to attack, as well as being less successful in counteracting the pheromone release that the beetles use to attract the swarms that can overwhelm tree defenses. Globally, warmer temperatures, dry summers, stressed trees, and in Europe, extensive monoculture conifer plantations have all been implicated in the unprecedented increase in bark beetles and the death of conifers across large areas.

**Chronic herbivory**—herbivory that is sustained over long time periods—can have major effects on plant demography. *Pinus edulis* (pinyon pine) trees subjected to chronic herbivory have lower rates of growth and an altered shape, and they produce male cones almost exclusively (Whitham and Mopper 1985). In a study of *Eucalyptus pauciflora* (snow gum, Myrtaceae) and *E. stellulata* (black sally) in subalpine habitats in the Snowy Mountains of Australia, Patrice Morrow and Valmore LaMarche (1978) found that trees treated with insecticide experienced large increases in growth, suggesting that chronic herbivory can greatly reduce lifetime fitness. Similarly, by spraying insecticide on experimental plants several times a year, Nadia Waloff and O. Richards (1977) showed that chronic herbivory reduced the seed yield of the British shrub *Sarothamnus scoparius* (broom, Fabaceae) by about 75% over 10 years. In a study of the semelparous plant *Cirsium canescens* (Platte thistle, Asteraceae; see Figure 7.6) in the sandhills of Nebraska, U.S., Svata Louda and Martha Potvin (1995) excluded inflorescence-feeding herbivores, leading to increases in the number of seeds, seedling density, and ultimately the number of flowering adults. It is likely that similar stresses are imposed by herbivory on many other plant species.

## Herbivores can change where plants grow

The spatial distribution of plants can also be affected, or even determined, by herbivores. The role of granivorous rodents in the distribution of *Achnatherum hymenoides* (Indian ricegrass, Poaceae), a perennial grass with large, nutritious seeds, was studied in the desert of western Nevada. Indian ricegrass is common in sandy soils but rare in adjacent rocky habitats (Breck and Jenkins 1997). The grass was able to survive and grow in both soil types when planted experimentally, although the plants grew taller in the sandy soils. However, rodents cached seeds only in the sandy areas, and this seed dispersal behavior appeared to be an important positive factor in determining the distribution of this plant.

The distribution of the shrub *Hazardia squarrosa* (previously *Happlopappus squarrosus*, sawtooth goldenbush, Asteraceae) in California is strongly affected by granivory. Svata Louda (1982) showed that *H. squarrosa* was most abundant in an inland transitional zone between coastal and interior climates but produced more seeds closer to the ocean (Figure 11.4). This pattern appears to be mainly a result of increased insect granivory on the plants growing closest to the ocean.

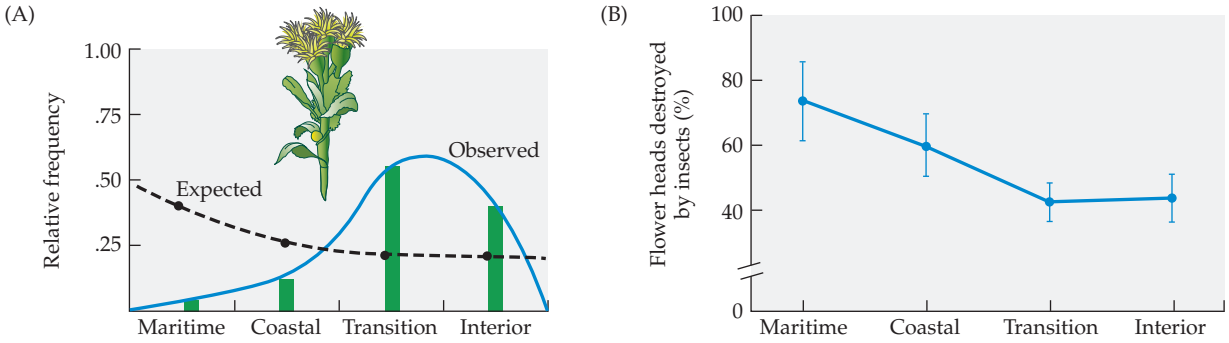



**Figure 11.4** (A) Observed (bars) versus expected (dashed line) relative frequencies of *Hazardia squarrosa* (previously *Happlopappus squarrosus*, sawtooth goldenbush, Asteraceae) over a climatic gradient in San Diego County, California. The observed distribution is based on the presence or absence of plants along 15,250 roadside segments, each 167 m in length. The expected frequency is derived by assuming that relative population size should be proportional to relative seed production. (B) Percentage of flower heads destroyed by insect granivores in each climatic zone; error bars are 95% confidence intervals. (After S. M. Louda. 1982. *Ecology* 52: 25–41.)

Donald Strong and his colleagues (1995, 1996) have unraveled a complex set of interactions that apparently underlie the population dynamics of a shrub, *Lupinus arboreus* (yellow bush lupine, Fabaceae), along the central coast of California. Large patches of this woody perennial periodically die off and are eventually regenerated from seed, so the population fluctuates over time. The plants are killed primarily by subterranean caterpillars of the ghost moth, *Phymatopus californicus*, which bore into the roots. However, an insect-killing nematode, *Heterorhabditis hepialus*, with its symbiotic bacterium, *Photorhabdus luminescens*, is a highly effective predator on the ghost moth caterpillars. Consequently, lupine-dominated areas that are heavily colonized by the nematode are protected from attack by the caterpillar, while those areas without the nematode experience massive periodic die-offs. Few other studies have reported control of plant populations by underground herbivores, not to mention such trophic complexity, but then, few studies have looked for either.

### Herbivory on seeds has both negative and positive consequences for plant populations

Granivory can have important consequences for plant populations. In some populations, granivores consume a large fraction of the seeds. *Tachigali versicolor* (suicide tree, Fabaceae) is a large tropical forest tree from Central America with large (500–600 mg) seeds. It is unusual in being monocarpic (flowering once then dying, see Chapter 7). In a detailed study of the seeds and seedlings produced by two large adult trees, Kaoru Kitajima and Carol Augspurger (1989) found that 51% to 83% of seeds were killed prior to germination, depending on the tree and the distance of the seed from the tree. The important granivores were bruchid beetles (which ate between 13% and 38% of seeds) and vertebrates (which ate from 0% to 59% of seeds). Seedlings had somewhat lower mortality rates in their first 2 months (24% to 47%, again depending on the tree and distance from the tree). Seedling mortality was primarily due to herbivory (6% to 17% of seedlings) and disease (3% to 25% of seedlings).

The chemical properties of seeds can deter or enhance granivory. Seeds of *Erythrina* (coral bean tree, Fabaceae), for example, contain compounds that are strongly neurotoxic in vertebrates; as a result, they are rarely eaten. On the other hand, studies of squirrel granivory on oaks (Fagaceae) provided some surprises. *Quercus rubra* (red oak, Fagaceae) and closely related species have acorns with high tannin concentrations, while the acorns of *Q. alba* (white oak) and its close relatives are low in tannins. One might expect that gray squirrels (*Sciurus carolinensis*) would therefore prefer white oak acorns. However, Peter Smallwood and Michael Steele (Smallwood et al. 2001; Steele et al. 2001) showed that the interaction is more complicated. Acorns of the red oak group remain dormant until spring. Acorns from the white oaks germinate in autumn and are therefore less desirable for the squirrels to cache over winter. While squirrels might eat white oak acorns when they encounter them, they prefer to collect and cache red oak acorns in quantity for the winter. Moreover, when they do cache white oak acorns, the squirrels generally remove the embryos first, preventing them from germinating. Thus, more red oak than white oak seedlings emerge from squirrel caches, and the squirrels are more effective as dispersers of red oaks—not because the squirrels avoid the tannin-laden red oak acorns, but because they actively prefer to collect and store them!

Many plants—especially woody plants—have large and erratic variability among years in the size of the seed crop produced, and this variability is generally synchronized among most of the plants in a population. This phenomenon, called **masting**, has been widely explained as an adaptation to granivory. Various ecologists have hypothesized that during masting years, the large numbers of seeds overwhelm the capacity of granivores to eat them all, allowing at least some seeds to survive. This explanation has been questioned by other ecologists, who argue that other factors (in particular, wind pollination) may be more important in selecting for masting (see Chapter 7).

### People use insect herbivores for biological control

**Biological control** is the deliberate use of herbivores or pathogens by humans to control populations of undesirable plant (or other) species. Many biological control agents (like many of the plant pests they are intended to control) are introduced from other continents. Successful instances of biological control of plant pests offer strong examples of control of plant populations by herbivores. One well-known example is the introduction of the moth *Cactoblastis cactorum* to Australia to control the introduced and invasive prickly pear cacti *Opuntia inermis* and *O. stricta* (Cactaceae). These species and three other Opuntias were brought to Australia in the late 1700s for cochineal production, and were subsequently planted more widely for cattle feed and living fences. By the 1920s, at their peak, the cacti had spread to cover vast areas in Australia, rendering them useless for farming and sheep grazing (Figure 11.5). The moth was introduced in 1925 from Argentina, where its caterpillars were found to be specialist herbivores on prickly pears. By 1935, the prickly pear populations in Australia had been decimated, and they have remained at low levels since that time.

Unfortunately, the success of *Cactoblastis* in Australia is colored by recent problems with this species in North America. *C. cactorum* was introduced to the Caribbean

(A)

(B)

Photographs by A. P. Dodd and the Commonwealth Prickly Pear Board

**Figure 11.5** (A) A dense stand of *Opuntia* in Queensland, Australia, prior to the release of *Cactoblastis*. (B) The same stand 3 years later.

in 1957 to control *Opuntia* species that had spread, primarily as a result of overgrazing of the perennial grasses. From there, the moths invaded Florida and are now spreading rapidly in the southeastern United States, where they are affecting several endangered native *Opuntia* species (Johnson and Stiling 1998). Travis Marsico and his associates (2011) used mitochondrial gene sequence data to determine the origin of the Florida moths and found that they are closely related genetically to the moths found near the original 1925 source site in Argentina, likely colonizing Florida multiple times from the Caribbean. There is much concern that this invasion will spread to the southwestern United States and Mexico, where the moths can devastate native cacti. *Opuntia* species are ecologically important in both countries, and they are also economically important in Mexico.

Specialist insects have been introduced to control *Lythrum salicaria* (purple loosestrife, Lythraceae), a highly invasive plant species that now dominates vast areas of wetlands in eastern and central North America (see Figure 14.4). Purple loosestrife may be limited by herbivores in its native habitat in central Europe. In those environments, it never reaches more than 5% cover and remains a minor component of wetland vegetation, in dramatic contrast to its spread in North America, where its specialist herbivores were absent. Bernd Blossey and his colleagues (2001) found that purple loosestrife is attacked by many different insects where it is native. These researchers imported populations of a root-mining weevil (*Hylobius transversovittatus*) and two leaf-feeding chrysomelid beetles (*Neogalerucella calmariensis* and *N. pusilla* [formerly *Galerucella calmariensis* and *G. pusilla*]) and released them in the United States and Canada in 1992 and 1993. These insects initially appeared to be successfully eliminating the dense populations of purple loosestrife and allowing the recolonization of native wetland plants. However, in a follow-up study in Ontario, Canada 20 years after release of the biocontrol insects, Excedera St. Louis and colleagues (2020) found no evidence that native plant communities had recovered following release of the biocontrol beetles. The beetle biological control was associated with greater herbivore damage, but there was no evidence for reduced invasive plant density or biomass or restoration of native species richness in their study populations, raising questions about the efficacy of this biocontrol effort.

An agent introduced to control the severely invasive exotic plant *Carduus nutans* (musk thistle, Asteraceae) and its relatives in the midwestern United States has proved to have serious negative effects on rare native thistles. *Rhinocyllus conicus* (flowerhead weevil) was introduced to control *Carduus* species after studies indicated that this weevil prefers *Carduus* to native thistles in the genus *Cirsium.* However, despite the weevil's preference for the invasive musk thistle, studies by Svata Louda and her associates (1997) showed that the weevil is also consuming and negatively affecting three native thistle species—*Cirsium canescens*, *C. centaureae*, and *C. undulatum*. Observed levels of *R. conicus* infestation were as great or greater on the native thistles as on the exotic thistles and were also greater than the levels of native insect infestation on the native thistles. Infested plants set many fewer seeds. In the case of *C. canescens*, which is already uncommon and restricted in its habitat, this new herbivore may present a danger to population persistence, as a reduction in seed production is expected to reduce its population growth

rate. The native *C. canescens* may actually be protected to some degree by the maintenance of the invasive *Carduus* populations in the same habitat, as the invader may attract the weevils away from the native thistle. However, this strategy has its own risks, as the weevils may "spill over" onto the natives some of the time.

George Heimpel and Matthew Cock (2018) encourage a new perspective on the introduction of biocontrol insects, where benefits and risks of introductions are explicitly weighed prior to introduction so that we can benefit from the introduction while reducing risks. In reviewing all insect biocontrol introductions, they found that 227 introductions have been carried out since 1888, when the first one (in California) was made. Of these, at least 10 had serious unintended negative consequences such as attacks on native species, leading to a focus on risks associated with biocontrol introductions (Simberloff and Stiling 1996; Heimpel and Cock 2018). The problem was that while this greatly reduced the risks associated with biocontrol releases, it also had the effect of greatly reducing potential benefits. Biocontrol has numerous environmental and other advantages over chemical controls using pesticides and herbicides, and there have been numerous striking successes such as that of *Opuntia* in Australia as well as other more modest gains.

## 11.3 Herbivores Affect Plant Communities in Different Ways

### Herbivore behavior can change plant community composition

Herbivores show preferences and exhibit behavior with respect to what they eat and how and when they eat it. Their behavior can have profound consequences for plant species richness and abundance. One important way in which herbivore behavior can affect plant community composition is the extent to which animals behave as generalists versus specialists with regard to what they eat. At one end of the spectrum, a pure generalist herbivore eats plants in the same proportions at which they are present in the community. At the other extreme, a pure specialist herbivore eats only plants (or even only specific plant parts) belonging to a single species or to a small group of closely related species. Highly specialized herbivores are therefore usually the most desirable choices for biological control agents because they minimize effects on nontarget species.

Most herbivores are somewhere between these two extremes, preferring certain foods to others but able to eat a variety of plants. Thus, drawing a strict contrast between specialist and generalist herbivores would be misleading. However, generalist herbivores tend to promote or maintain species diversity because they keep faster-growing, dominant species from outcompeting others. The effects of specialists, on the other hand, depend on the roles that their preferred food species play in the community. It is likely that a specialist on a potentially dominant species will have a very different effect on the community than a specialist on a less common species. Thus, herbivory can either increase or decrease the diversity of the plant community. The outcome depends not only on patterns of consumption, but also on interactions between herbivory, plant competition, and abiotic factors such as soil moisture, nutrients, and light levels. The effects of herbivory can vary over spatial scales and over the course of time.

Nancy Huntly (1987) studied the foraging behavior of pikas and the consequences of their foraging for the plant community. Pikas, a small, territorial relative of rabbits and a strong contender for the world's cutest mammal, live on high-altitude, rocky talus slopes in western North America. Pikas forage outward from their dens. They are generalists but prefer certain plant foods to others. Pikas do not hibernate; instead, they collect large hay piles during the summer for winter use (Figure 11.6). Huntly experimentally excluded pikas from plots of vegetation at different distances from their dens. She found that the foraging animals had large effects on plant species composition close to their dens, where they spent the most time, with decreasing effects farther away in the surrounding meadows.

Courtesy of C. Ray

**Figure 11.6** Pikas (*Ochotona princeps*), relatives of rabbits and hares, are common in high-altitude habitats in western North America. Pikas feed on almost all the plant species that grow around the rock piles where they live. Because they forage from a central location—a den—their effect on the plant community is strong, but it becomes weaker at greater distances from the den. As shown here, pikas collect hay piles during the summer for winter use.

Seed-eating and vegetation-eating small mammals have been shown to affect plant community structure in a number of arid environments. In the semiarid shrubland of northern Chile, Julio Gutiérrez and his colleagues (1997) used fences and netting to selectively exclude small mammals (principally the herbivorous degu, *Octodon degus*; Figure 11.7) from large experimental plots, and predatory birds (particularly owls) from other plots in addition to having unfenced control plots. The exclusion of degus resulted in an increase in the cover of shrubs and perennial grasses, an increase in the diversity of perennial species, and a decrease in annual plant cover. The researchers found some indirect effects of predator exclusion on the vegetation (presumably by allowing increases in herbivores). They also documented strong effects of weather as well as interactions between weather and herbivore effects.

While the areas with arid and semiarid vegetation of Chile and Argentina are primarily home to herbivorous and insectivorous species with few granivores, in striking contrast, seed-eating rodents dominate North American deserts. James Brown and his colleagues (Brown and Munger 1985; Guo and Brown 1996; Brown et al. 1997) conducted a series of long-term experiments in which they excluded different small mammals (particularly heteromyid rodents) and ants from plots in the Chihuahuan Desert of eastern Arizona. Over time, removing either rodents or ants caused substantial changes in plant species composition. Where rodents were removed, large-seeded species increased and small-seeded species decreased. Where ants were removed, the opposite results were found.

Valerie Brown and Alan Gange (1992) carried out an innovative field experiment to ask how the effects of root-feeding insects on the plant community might differ from the effects of foliage feeders in an early successional field. To test this, they used insecticides to kill either aboveground or belowground insects. They found that both aboveground and belowground herbivory by insects had major (but different) effects on the timing and direction of succession (see Chapter 13). In this study, the aboveground insects were largely sap-feeding Hemiptera, which preferred perennial grasses. Their herbivory suppressed perennial grass colonization, slowing succession. The removal of the insects led to a luxuriant growth of the grasses, which then shaded and replaced the lower-growing herbaceous dicots, leading to a steep decline in species richness. Underground, chewing insects in the Coleoptera and Diptera fed primarily on the roots of the dicots. Reducing belowground insect numbers led to the persistence of annual dicots and a great increase in colonization by perennial dicots and consequently a great increase in species richness. The ordinary presence of these root-feeding insects thus speeds up succession by reducing the dicots and increasing the colonization of the field by grasses.

Courtesy of L. Ebensperger

**Figure 11.7** The degu (*Octodon degus*) is a small rodent that is a principal herbivore in Chilean deserts.

Herbivory is not the only way in which plant-eating animals affect plant communities. Other kinds of herbivore behavior can also change the environment and have strong effects on plant communities. Mammalian herbivores, in particular, create gaps and patches when making burrows and trails, and they trample vegetation around available water sources. Domesticated cattle can cause severe damage to vegetation along streams and watercourses by repeated trampling. Elephants consume and trample enormous amounts of plant material. Tree canopy cover in Serengeti National Park in Tanzania, for example, was reduced by about 50% by elephants (Pellew 1983).

Not all such effects of herbivores are negative. In tallgrass prairies in the midwestern United States, native bison create depressions where they roll in the dust. In spring these areas contain small temporary pools of water, and in summer they are inhabited by annual species that otherwise would be excluded by perennial grasses (Collins and Uno 1983). Large herbivores can also affect plant distributions and species richness through strong effects on soil nutrients caused by their urination and defecation. In a study at Yellowstone National Park, David Augustine and Douglas Frank (2001) compared soil characteristics and community characteristics between ungrazed grasslands and grasslands grazed by large herbivores—elk, bison, and pronghorn. Species richness and diversity were greater in the grazed grasslands, particularly at very small extents (Figure 11.8). These herds move over large areas as they graze. In contrast, in grazed rangeland, intensive grazing by cattle that are kept at high densities in limited areas for long periods of time can lead to large decreases in plant diversity.

Large mammals can sometimes shape ecosystems in complex ways. In Yellowstone National Park, wolves prey on elk, and elk eat aspen (see Figure 9.11). In the park, *Populus tremuloides* (trembling aspen, Salicaceae)

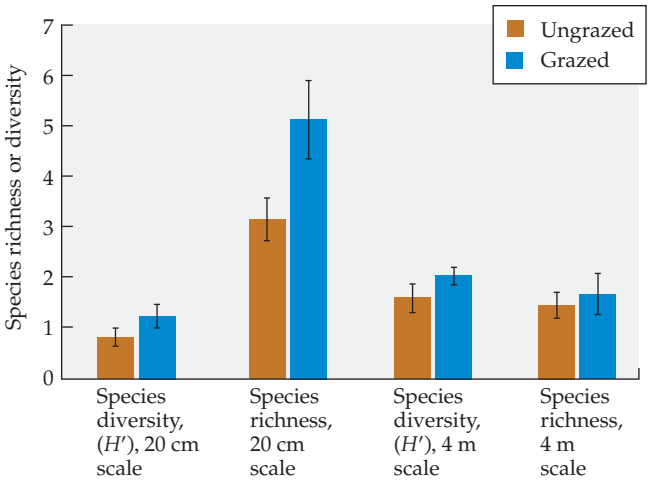



**Figure 11.8** Plant species diversity (*H′*), as measured by the Shannon-Wiener index (see Chapter 12), and plant species richness are greater in grazed than in ungrazed grasslands in Yellowstone National Park. The effect was greater for comparisons of 20 × 20 cm quadrats than for 4 × 4 m quadrats, implying that the effects were occurring at a very small spatial scale. This scale effect was contrary to expectations because urination and defecation by large mammals is thought to lead to a very patchy distribution of soil nitrogen and nitrogen mineralization rates. Species richness is the mean number of species per plot on the 20 cm scale, and the mean number of species per plot divided by 10 on the 4 m scale. Error bars are ±1 standard error. (After D. J. Augustine and D. A. Frank. 2001. *Ecology* 82: 3149–3162.)

© Sergey Gorshkov/Minden Pictures

**Figure 11.9** Snow geese grazing and flying over tundra, Wrangel Island, Russia, in the Arctic Ocean. Wrangel Island was probably the last place on earth inhabited by wooly mammoths, who lived there for many thousands of years after they were extinct elsewhere; the grazing effects of wooly mammoths would have been indeed impressive.

have been in decline in recent decades, coinciding with other pronounced changes to these plant communities. Elk browsing on the aspen appears to have profound effects on its persistence and regeneration. Wolves were reintroduced to Yellowstone in 1995–1997. Some ecologists thought that wolves could save the aspen if the predator (wolves) limited the behavior of the herbivore (elk) due to predator avoidance, resulting in a postive effect of the predator on the primary producers (aspen). This was called the 'landscape of fear' hypothesis (Kauffman et al. 2010). Since wolf reintroduction at least some riparian plant communities have experienced a recovery (Beschta and Ripple 2013). However, the data have not strongly supported the hypothesis. The elk eat the aspen when the wolves are not active, and wolf predation risk has no effect on elk stress levels, body condition, or pregnancy (Kohl et al. 2018). While the findings remain controversial, aspen decline is unlikely to be halted by the wolves.

Even birds can have dramatic effects, particularly when they reach very high numbers. Lesser snow geese (Figure 11.9) are distributed pan-globally in wetlands in the subarctic and Arctic in Europe, Asia, and North America. In North America, they breed on the coastal marshes of the tundra of Hudson Bay and James Bay in Canada during the summer, migrating in autumn to the Gulf Coast of Louisiana and Texas and other regions. Their numbers have exploded, from midwinter counts of about 0.8 million in the 1960s, to 3 million in the mid-1990s, to well over 12 million by 2019 (Abraham and Jefferies 1997; Koons et al. 2019). These increases are due to a variety of factors, primarily changes in land use by humans that have expanded food sources (especially agricultural grains) in the fields where they overwinter and during migration, and protection from hunting in wildlife refuges in winter. Efforts to control their numbers by increasing hunting have been a dismal failure (too few hunters taking too few geese). Ever larger flocks of geese return each summer to the tundra in excellent nutritional condition and capable of rearing large numbers of young. Snow geese are grazers but also kill plants by pulling them up by their roots or rhizomes. Large flocks of the birds can have dramatic effects on the tundra landscape, not only changing plant species composition but also leaving entire areas bare of plants and destroying biotic soil crusts (see Chapter 4). This has resulted in catastrophic changes to Arctic salt marshes, with well over 35,000 ha of productive salt marsh transformed into bare sediment along 2000 km of coastline of the Hudson Bay in Canada (Jefferies et al. 2006; Koons et al. 2019). Their winter and summer habitat use has expanded both in size and in the variety of habitats they occupy, and their population is therefore unlikely to be controlled by food availability as they continue to spread to new environments. As salt marsh and other wetland habitats have been catastrophically transformed, over 200 other species of water birds and numerous other species have been threatened due to the loss of habitat.

## Herbivory might result in apparent competition among plants

One type of animal-mediated interaction among plants is **apparent competition**: a density-dependent negative interaction between species that might at first appear to be due to competition for resources but is actually due to a shared predator or herbivore (Holt 1977) (see Chapter 10). Apparent competition might occur if, for example, as the combined density of two plant species increased, they increasingly attracted the attention of herbivores. Thus, each species would suffer from the other's presence and abundance, but because of herbivory rather than competition.

Although it is often invoked, apparent competition among plants has been demonstrated only in a few cases. One intriguing study by Laura Sessions and David Kelly (2002) showed that a native New Zealand fern, *Botrychium australe* (adder's tongue, Ophioglossaceae), was declining in abundance due to apparent competition with an introduced grass, *Agrostis capillaris*. The grass may serve as a refuge for an introduced slug. The native fern survives well after fire. However, fire also increases populations of the introduced grass, which leads to an increase in slug density. Increased slug populations following a severe fire led to severe defoliation and mortality of the fern due to the positive effects of the greater grass cover on the slugs.

Apparant competition is not an ecologial process per se. Rather, it is a failure by the researcher to recognize that population densities are being controlled by a factor other than competition. It might be that apparent competition is rare in plants because plant population densities are rarely controlled by herbivores, or it might be that plant ecologists simply have not looked for apparent competition because they assume that top-down control is rare.

## Domesticated and introduced herbivores can shape plant communities

One of the classic stories in the ecology of herbivory is that of rabbits, chalk grassland vegetation, and a rabbit disease, myxomatosis. Chalk grasslands are highly diverse plant communities found on limestone-derived soils in southern England. These grasslands alternate with woodlands, and they have been used for centuries to graze sheep and, beginning in the twentieth century, cattle. European rabbits (*Oryctolagus cuniculus*) were brought to England by the Romans but were not established in the wild until medieval times, when they escaped from the estates of royalty and aristocrats where they had been introduced as a source of food and for sport hunting. They gradually became widespread, with numbers approaching 100 million feral rabbits in Britain between the early 1800s and 1950, due to direct and indirect human influences, including climatic amelioration, predator reduction, and increases in agriculture. They became a major crop pest and had large results on native vegetation (Sumption and Flowerdew 1985). Rabbit densities varied greatly from one spot to another.

What effects have the rabbits had on the high-diversity vegetation of the chalk grasslands? The father of modern experimental plant ecology, Arthur Tansley, observed that where rabbit populations were high, the turf was typically chewed down to a height of 1 to 2 cm, while that of sheep-grazed grasslands was typically 5 to 10 cm in height. He hypothesized that if sheep grazing and rabbit feeding were prevented, the grasslands would revert to forest. To test the effects of the grazers, Tansley (Tansley and Adamson 1925) fenced plots of vegetation to exclude rabbits and sheep. At first, there was a great increase in the growth of plants inside the exclosures, and many plants that never succeeded in reproducing outside the fences flowered abundantly. After some time, however, perennial grasses, the preferred food of the rabbits, grew up and shaded shorter dicot species. Total biomass and average vegetation height increased substantially. The palatable grasses became more dominant, while less competitive plant species declined in abundance but did not disappear altogether. In addition, plants not characteristically found in the chalk grasslands invaded some of the ungrazed areas. However, the predicted large-scale colonization by woody species did not occur, perhaps because the exclosures were too small or too far away from sources of tree seeds (see Chapter 15).

Sir Arthur George Tansley

This experiment was repeated unintentionally on a larger scale some 30 years later, when a viral rabbit disease, myxomatosis, was accidentally introduced into Great Britain that wiped out over 99% of the rabbit populations, including in the chalk grasslands. Immediately following the decimation of the rabbit population in 1954, many rare plant species were found in the chalk grasslands that had not been seen in years (Thomas 1960, 1963). These species had been selectively grazed by rabbits, never becoming large enough to flower and be noticed, or had been grazed as soon as they grew past the seedling stage. Various rare orchids and other showy species, such as *Helianthemum chamaecistus* (frostweed, Cistaceae) and *Primula veris* (primrose, Primulaceae), appeared and flowered in abundance. Some of these species had been common 100 years before, prior to the great increase in rabbits. Other species decreased as the rabbits disappeared, either because they were outcompeted or because they had been favored by the nitrogen

from the rabbits' urine, and a number of small legumes and annual plants became extinct locally. Tall grasses began to become more prominent, and woody species began to invade (as Tansley had predicted), reducing species diversity in some regions. A **trophic cascade** (when population or species composition changes at one trophic level result in major changes at other trophic levels) resulted in alterations in community composition, including increases in invertebrate species and in herbivores, such as voles (*Microtus agrestis*), and declines in some of the predators on rabbits (Sumption and Flowerdew 1985). When the small surviving rabbit populations evolved resistence in the early 1960s, their numbers recovered, and the vegetation largely returned to its previous state for the next half century. However, since the mid-1990s, a new disease, rabbit viral hemorrhagic disease, has again cut their numbers, particularly in Scotland. It remains to be seen how this new decrease will affect the plant communities in those areas.

Striking pictures of the effects of grazing are offered by fence lines where one side is heavily grazed and the other ungrazed or lightly grazed (Figure 11.10). Like pikas, cattle are selective generalists, eating many species, preferring some, and avoiding others. They usually avoid woody and spiny species, as well as those with toxic or noxious defensive chemicals, but may eat them when preferred forage is scarce. Cattle can poison themselves, for example, by grazing on plants in genera such as *Digitalis* (foxglove, Scrophulariaceae; see Figure 7.1) and *Astragalus* (locoweed, Fabaceae). Grazing by cattle can have dramatic effects on community composition, particularly when their densities are high or when grazing occurs at particularly sensitive times of the year for plant recovery and regeneration (e.g., when grass seeds are ripening). Over time, particularly with heavy grazing, preferred plants such as palatable and nutritious grasses decline in abundance and are replaced by less edible species, drastically changing the composition and appearance of the plant community. Heavy overgrazing leads to bare patches of ground, weed invasion, and severe erosion, especially on slopes. The landscape itself can be changed, with deep ravines and gullies replacing rolling, grass-covered slopes, as a result of long-term damage to the plants that once held the soil in place. These problems tend to be more severe in arid environments but can occur even in mesic habitats. Other grazing animals, including sheep and goats, can cause similar effects. Problems caused by overgrazing are widespread in western North America but also occur in Africa, the Mediterranean region, and Australia, among other places. Overgrazing by domestic animals has contributed to turning vast areas of marginal grasslands into shrublands or deserts.

An introduced insect herbivore, the hemlock woolly adelgid (*Adelges tsugae*), is currently causing dramatic changes in forests in the eastern United States (Figure 11.11). This insect, a specialist on hemlocks, was brought from Asia to the Pacific Northwest in 1924 and moved to the eastern United States in the 1950s. It has spread rapidly in forests from New Hampshire to the mountains of Georgia and Tennessee. Trees are killed within several years of becoming infested by these tiny phloem-feeding, aphid-like insects. Because hemlocks are dominant in, or a major component of, many eastern forests, this insect invasion is resulting in substantial changes in community composition as many large trees are killed, opening gaps for colonization by other species.

White-tailed deer have caused major changes in forests in the eastern and midwestern U.S. as their population numbers have exploded since the mid-twentieth

Courtesy of S. Scheiner

**Figure 11.10** A fence line in northern Arizona, U.S.A. The area at the left has been grazed by cattle; the area at the right has not been grazed.

Courtesy of M. Montgomery/ USDA Forest Service

**Figure 11.11** Small, white egg sacs on the underside of this hemlock twig are a sign of infestation by the hemlock woolly adelgid.

century. White-tailed deer are quintessential edge species and have thrived as people have fragmented the forests of these regions and in the sprawling suburbs surrounding cities. They have denuded forest understories by their browsing and slowed or prevented forest succession and regeneration. Experiments by Andrea Dávalos and associates (Dávalos et al. 2015) show that the combined effect of white-tailed deer, earthworms, and invasive plants decreased the ability of rare plant species to persist and recruit in forest environments. They also found that the interaction of browsing by white-tailed deer and earthworm invasion greatly facilitated the success of the highly invasive species *Alliaria petiolata* (garlic mustard, Brassicaceae), *Berberis thunbergii* (Japanese barberry, Berberidaceae), and *Microstegium vimineum* (Japanese stiltgrass, Poaceae). Many other researchers have found similar effects of white-tailed deer in altering forest communities.

### How important is herbivory in shaping the natural world?

Most ecologists agree that herbivorous insect populations can be controlled by food and abiotic factors—density dependence and density independence—determined by the time, place, history, and conditions, and that stochasticity also plays a large role in controlling insect population sizes. These factors then influence the effects that the insects have on plant communities. In a meta-analysis of experimental studies on herbivory, David Bigger and Michelle Marvier (1998) concluded that, on average, herbivory causes a substantial reduction in plant biomass in natural plant communities.

Contrary to the casual observations of many people, invertebrate herbivores such as insects often have a much greater effect on plant populations and communities than do vertebrates. As we have seen, plant defenses play an important role in protecting plants and limiting the success of herbivores, although both abiotic factors and predation clearly also can be important in limiting herbivore populations. Confining ourselves to the question with which we began, however—whether herbivory is important for plant communities—the answer seems clear: herbivores often have large effects on plant populations and communities.

## 11.4 Plants Defend Themselves against Herbivores by Different Means

Because established plants cannot move, they cannot escape by running away from herbivores, but have to "sit there and take it". Their immobility results in natural selection for being able to defend themselves from damage or death from herbivory. Selection may result in plants that are tougher, less palatable, and generally better defended. There are a great many different types of plant defenses, just as there are a great many different ways that plants can be attacked.

### Plants use a variety of physical defenses to protect themselves

The physical and mechanical defenses of plants include obvious structures such as thorns and spines, which probably serve best to discourage mammalian browsers and birds but do little to deter insects. Single-celled plant hairs—**trichomes**—serve many functions, including protection (Figure 11.12A). Insects are deterred by leaf hairiness and can be impaled by some trichomes. Other trichomes secrete noxious compounds that can deter vertebrates or sticky substances that impede insects. Species of *Urtica* (stinging nettles, Urticaceae) have brittle, elongate trichomes that break off when brushed, leaving a pointed fragment that pierces the skin and injects a painfully irritating fluid (Figure 11.12B).

(A)

Courtesy of D. McIntyre

(B)

© A. Syred/Science Source

**Figure 11.12** Trichomes are important physical defenses for plants. (A) A haze of trichomes covers the stem and leaves of this *Solidago* plant (goldenrod, Asteraceae). (B) Scanning electron micrograph of trichomes on the leaf of *Urtica dioica* (stinging nettle, Urticaceae). The long trichome in the center is a stinging hair; when its tip is broken off by some action (such as the touch of a human hand or foot), it easily penetrates the skin and injects the neurotransmitter acetylcholine, along with histamines that produce allergic reactions. The shorter trichomes do not have these protective chemicals.

Trichomes of this kind have evolved independently in four plant families: the Urticaceae, Euphorbiaceae, Loasaceae, and Boraginaceae.

Other physical defenses include materials that make entry into the plant body difficult, such as thick bark on trunks or roots or the tough coats that protect seeds and some fruits, such as nuts. Various cells and tissues that make up the plant body probably also serve defensive functions, such as the cap of thick **sclerenchyma** (a plant tissue with lignified cell walls) that surrounds the vascular bundles carrying food and water in young stems, and the waxy cuticle on the surface of leaves, which both reduces water loss and protects against fungal attack.

Leaf toughness presents an important mechanical barrier to chewing insects, mammalian browsers, and other herbivores, as well as interfering with tissue digestibility. Toughness is primarily a result of the content, type, and placement of sclerenchyma fibers (elongated cells), although short **sclereids** (sclerenchyma cells with thick, lignified, much-pitted walls) may be responsible for the hardness of some structures. Toughness may also depend on other thick-walled cell types, such as xylem and **collenchyma** (supporting tissue made of living elongated cells with irregularly thickened walls). Newly formed, expanding leaves are considerably less tough than older ones because the structural tissues involved in making leaves tough would interfere with the leaf's ability to grow to its full mature size and thus do not develop until the leaf is mature. For this reason, new leaves are vulnerable to attack and must depend on alternative means of protection, such as toxic chemical defenses.

Grasses are unpalatable to most generalist herbivores because they sequester large amounts of silica, which makes them difficult to chew and difficult to digest (picture chomping down on ground-up bits of glass, or large amounts of sand). The silica is contained in specialized epidermal cells as well as in other plant parts. A few other taxa also use silica as a deterrent to herbivory, including horsetails (Equisetaceae) and palms (Arecaceae).

Another characteristic that has been suggested to act as a defense against herbivory is the nutritional quality of leaf tissue (see Chapter 4). Eating leaves with a low nitrogen and water content results in poor growth and survival for herbivores. Leaves with a higher nitrogen and water content are usually preferred by herbivores over those with a lower content (all else being equal). This fact poses an evolutionary dilemma for plants: The metabolic enzymes responsible for growth and photosynthesis contain nitrogen. Reducing the concentration of nitrogen-containing compounds in the leaves to discourage herbivory could result in reduced photosynthesis and growth.

Nitrogen is almost always limiting to plants, but it is even more limiting to herbivores. Animals need much more nitrogen to function than plants do, and they maintain much higher concentrations of nitrogen in their bodies (see discussion of ecological stoichiometry in Chapter 4). Plants growing in soils with little available nitrogen may have high concentrations of carbon relative to nitrogen in their tissues, which reduces their nutritive value. As global $CO_2$ concentrations increase because of human emissions (see Chapter 16), many herbivore populations are expected to decline because of the resulting decline in the carbon-to-nitrogen ratios, and thus the nutritional value, of many plants. Peter Stiling and his associates (1999) found that several species of leaf miners (see Figure 11.1C) consumed more tissue from oak leaves grown under enriched $CO_2$ conditions but nevertheless had greater rates of mortality, suggesting reduced nutritional value of the food source (Figure 11.13).

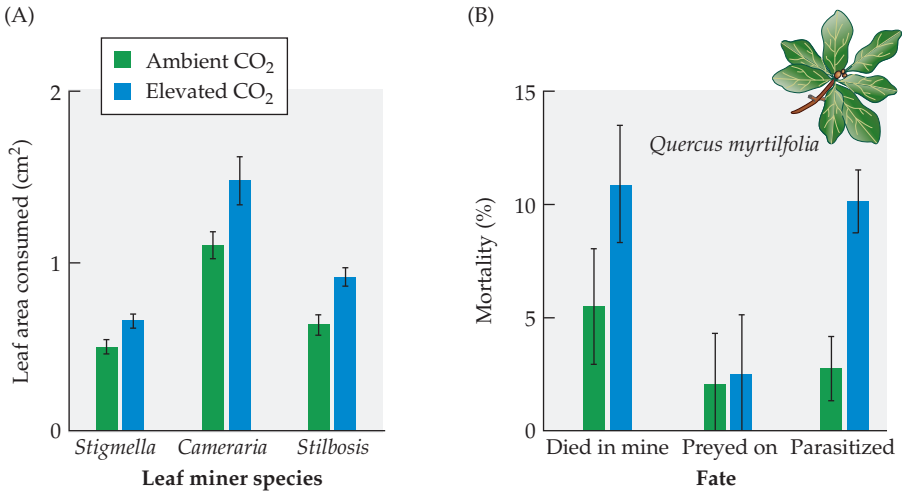



**Figure 11.13** Results of an experiment comparing the performance of leaf miners of the genera *Stigmella*, *Cameraria*, and *Stilbosis* on *Quercus myrtifolia* (myrtle oak, Fagaceae) grown under ambient versus enriched $CO_2$ conditions. (A) The insects removed more tissue from the plants grown under enriched $CO_2$ conditions (error bars are ±1 standard error). (B) Nevertheless, the insects feeding on those plants had higher rates of mortality from a number of causes (error bars are ±1 standard deviation). (After P. Stiling et al. 1999. *Ecol Appl* 9: 240–244.)

## Plants have evolved a wide range of chemical defenses against herbivores

Plants can marshal an arsenal of chemical weapons in response to herbivory. Chemical ecologists distinguish between **primary metabolites** and **secondary chemicals**, also called **secondary metabolites**. Primary metabolites (such as sugars, amino acids, and DNA) are compounds necessary for the basic metabolic functions of the plant (such as cellular respiration and photosynthesis). Secondary chemicals are not typically directly involved in growth, development, and reproduction. They constitute a broad group of compounds that serve a wide variety of functions, often important for their ecological roles, including defense and attraction of pollinators. Secondary chemical compounds are generally found only in particular species or groups of species, and often only in specific organs or tissues. Some primary metabolites can also function as defenses in certain plant groups. Until the last third of the twentieth century, secondary chemicals were believed to be waste products by plant physiologists and biochemists who were unaware of their ecological functions. The role of these compounds was largely discovered by chemical ecologists and evolutionary biologists investigating plant-insect interactions.

The three major categories of defensive secondary chemicals are **phenolics**, **alkaloids**, and **terpenes**. These categories are not exclusive, as some large molecules contain subunits of more than one type. Additional defensive compounds include toxic proteins and amino acids, protease inhibitors, and cyanogenic compounds.

Phenolics include a large variety of chemicals consisting of an aromatic ring with an attached hydroxyl group, –OH (Figure 11.14). Probably the most important defensive phenolic compounds in angiosperms and gymnosperms are the **tannins**, which reduce the digestibility of plant tissues. They are packaged in cell vacuoles and are present in high concentrations in the leaves of many woody plants, such as those in the Fabaceae, Fagaceae,

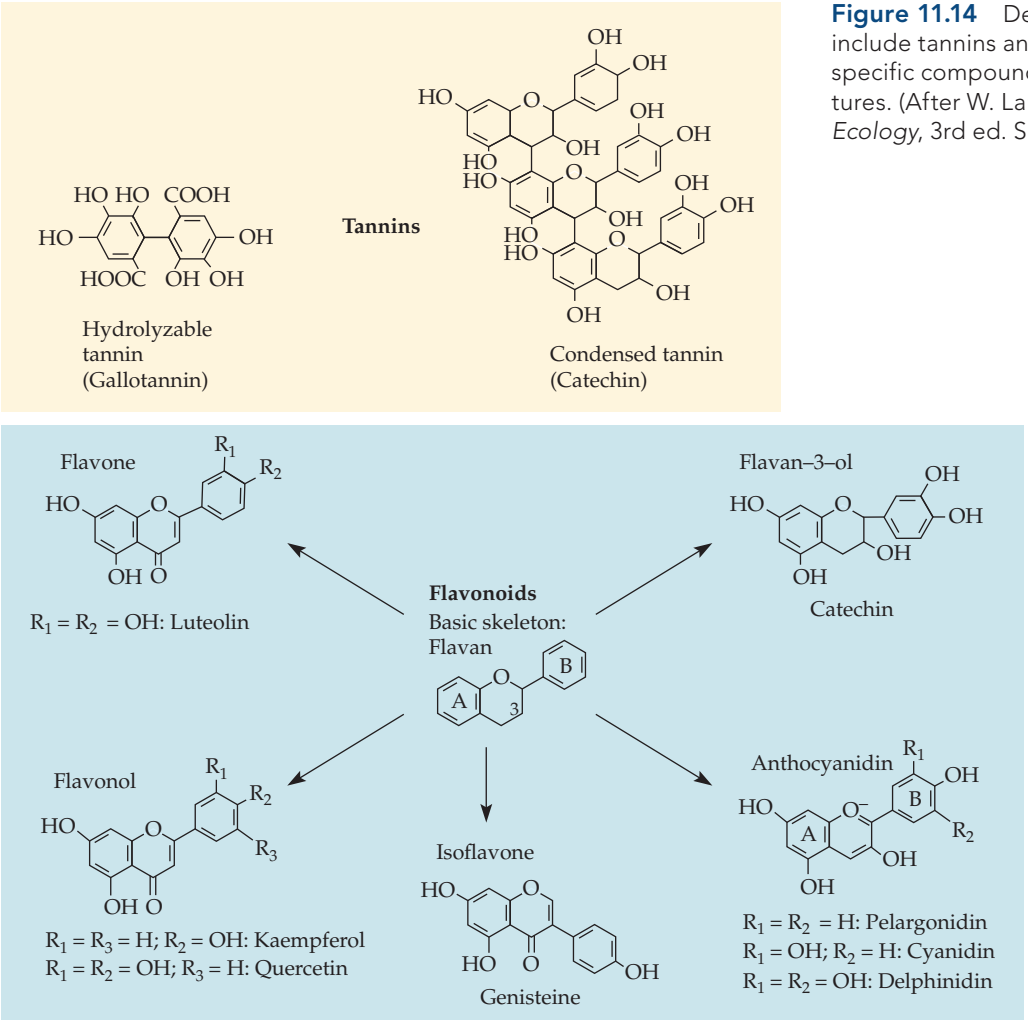



**Figure 11.14** Defensive phenolic compounds include tannins and flavonoids. Examples of specific compounds are given below the structures. (After W. Larcher. 1995. *Physiological Plant Ecology*, 3rd ed. Springer: Berlin.)

Myrtaceae, and Polygonaceae. Another important group of phenolic compounds are the **lignins**, which impregnate secondary (woody) cell walls, giving them structural strength as well as providing a barrier to attack by herbivores and pathogens. Other phenolics include poisonous saponins as well as many non-defensive compounds including the flavonoids and anthocyanins, pigments that give flowers and fruits their colors.

The alkaloids are another broad group of compounds and include many that are used as pharmaceuticals (Figure 11.15). Some 10,000 alkaloids have been isolated and their structures analyzed. Alkaloids are relatively small molecules that contain nitrogen. They have a bitter taste, and many are toxic to herbivores. Alkaloids are generally highly specific to the plant species or group of species in which they are found. They are often effective in small quantities—as in the cases of cocaine, nicotine, and caffeine—although in some cases particular plants or plant parts produce them in high concentrations.

Terpenes are found in all plants, and an individual may contain many different terpenes (Figure 11.16). Terpenes play a wide variety of functional roles in plants. They are composed of multiple units of the hydrocarbon isoprene ($C_5H_8$), and they may be large or small, depending on how many isoprene units they contain. In sunlight isoprene is emitted by the leaves of some plants (including *Eucalyptus* species), sometimes in large quantities, and may serve to protect them from heat damage. The **latex** (milky sap) found in members of the spurge (Euphorbiaceae) and dogbane (Apocynaceae) families contains toxic terpenoid defenses. The oils that are responsible for the characteristic flavors and scents of the mints (Lamiaceae) are terpenes; they deter herbivores and reduce the growth of bacteria and fungi. Cardiac glycosides can cause heart damage in vertebrates and are poisonous to many insects; these terpenes are produced by many unrelated plant species, including species in the Scrophulariaceae (such as *Digitalis*, foxgloves) and the Apocynaceae (such as *Asclepias*, milkweeds, and *Apocynum*, dogbanes). Phytoecdysones are terpenes that mimic insect molting hormones and interfere with the development of insect larvae; they are produced by a number of taxa, including ferns, cycads, and some angiosperms.

Plants in the Brassicaceae (mustard family) contain characteristic secondary compounds, chiefly glucosinolates, which are effective repellents of most generalist herbivores. These mustard oil precursors (responsible for the "bite" of the many cultivated species in this family) almost completely deter mammals and nonadapted insects. Glucosinolates are not as effective against insect herbivores that have become specialized on this family, however, so only high concentrations can offer some protection against these herbivores. Svata Louda

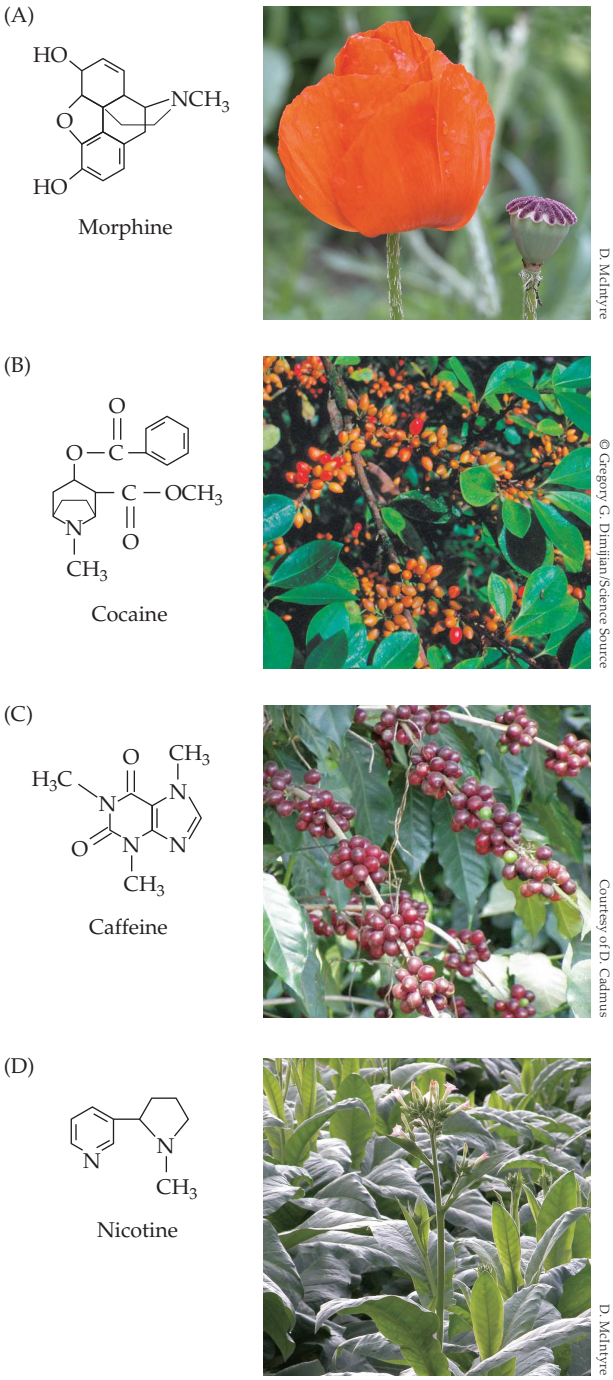



**Figure 11.15** Examples of various alkaloids used by humans, and their sources. (A) Opium comes from the milky sap of seed pods of *Papaver somniferum* (opium poppy, Papaveraceae); morphine is derived from opium. (B) Cocaine is found in the leaves of *Erythroxylum coca* (coca, Erythroxylaceae). (C) Caffeine is found in coffee "beans," the fruits of *Coffea* (coffee, Rubiaceae), as well as many other plants. (D) Nicotine comes from the leaves of *Nicotiana tabacum* (tobacco, Solanaceae) and other species in this genus.

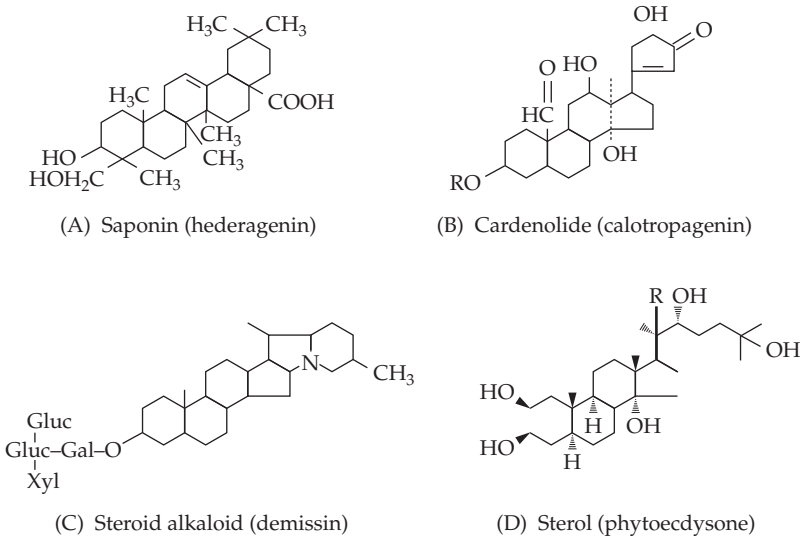



**Figure 11.16** Triterpenes play important roles in plant defenses against herbivores. (A) *Hedera helix* (English ivy, Araliaceae) contains the saponin hederagenin in its leaves and fruits. (B) *Asclepias* (milkweeds, Apocynaceae) contains the cardiac glycoside calotropagenin. (C) *Solanum demissum* (potato, Solanaceae) produces the steroid alkaloid demissin in its leaves. (D) A number of taxa produce phytoecdysones, compounds that mimic insect hormones and interfere with insect metamorphosis. (After W. Larcher. 1995. *Physiological Plant Ecology*, 3rd ed. Springer: Berlin.)

and James Rodman (1996) found that glucosinolates not only can affect the interaction between plants and their herbivores, but also can indirectly control the distribution and abundance of the plants. *Cardamine cordifolia* (bittercress, Brassicaceae) is native to the Rocky Mountains of North America, where it grows only in shaded, moist forest edges. Bittercress may be excluded from sunny sites by greater vulnerability to chronic herbivory at those sites. In a series of experiments, plants grown in sunny sites experienced greater water stress, which caused changes in their biochemistry, including a reduction in their glucosinolate concentrations. Consequently, those plants had more insect herbivores and suffered greater damage. Thus, herbivory may be controlling the plant's distribution by causing differential damage among microhabitats. A prominent hypothesis put forward the 1970s was that plant apparency determines the types of defenses employed by plant species (Feeny 1975). According to this idea, some plants remain one step ahead of herbivores by having short lives and high dispersal abilities, so that the next generation disperses to another location before they are found. These species were hypothesized to be less well defended chemically than long-lived, more apparent plants. However, in a meta-analysis of 30 years of experimental data, Angela Smilanich and colleagues (2016) found that most of the predictions of the apparency hypothesis were not supported. Rather, there were often beneficial effects of secondary metabolites on herbivores who presumably have evolved adaptations to plant defenses.

## Plant chemical defenses can be constant or be induced by herbivory

**Constitutive defenses** are those that are present in a plant regardless of herbivore damage. They may be present throughout the life of a plant or may change as an individual grows and matures. **Induced responses** are elicited by an herbivore attack. If these responses serve to protect the plant (whether or not they harm the herbivore), they are called **induced defenses**; if they have a negative effect on the attacking herbivore (whether or not they defend the plant), they are termed **induced resistances** (Karban and Baldwin 1997). Both physical and biochemical defenses can be induced, although most of the research in this area has concerned biochemical responses. Changes in physical defenses in response to herbivory include the secretion of additional mucilage and fiber in various *Opuntia* species (prickly pear, Cactaceae), which walls off insect herbivores, and an increase in silica concentrations in grasses, which makes them harder to chew. Silica from ingested grasses has been found in the fossilized dung of dinosaurs from 70 million years ago (Prasad et al. 2005).

There has been a great deal of discussion regarding the relative benefits of constitutive versus induced defenses (Karban and Baldwin 1997; Lerdau and Gershenzon 1997). Historically, most research on plant responses to herbivory has focused on constitutive defenses, but beginning in the 1980s, interest in induced defenses greatly increased. It is hypothesized that chronic herbivory should select for constitutive defenses, whereas low average rates of attack with occasional intense bouts of herbivory should select for induced defenses, because the cost of producing defensive compounds should make it beneficial to a plant to produce the smallest effective amounts of these compounds over its lifetime.

Many different classes of chemicals have been found to act as induced defenses. Ian Baldwin and his associates (Baldwin 1988, 1991; Karban and Baldwin 1997) have done much to increase our understanding of the process of induction. Herbivore attack is communicated throughout the plant by a group of plant hormones called jasmonates. In some cases, chemical defenses are induced by highly specific cues, such as caterpillar saliva. In other cases, any mechanical damage is sufficient. The toxic alkaloid nicotine, for example, is produced in

response to mechanical damage to the leaves of *Nicotiana* (tobacco, Solanaceae). Nicotine mimics the neurotransmitter acetylcholine, blocking the receptor for that neurotransmitter in both insects and vertebrates. At low doses, nicotine stimulates the nervous system, but at higher doses it is a depressant, ultimately capable of causing paralysis and death. Smokers become addicted to the excitatory effects of low doses of this potent nerve toxin; caterpillars or grazing mammals may be paralyzed and killed by the same toxin. Small amounts of nicotine are present in undamaged tobacco plants. After leaf damage occurs, nicotine levels can increase four to ten times, reaching concentrations that can kill an herbivore after a single meal.

Nicotine is synthesized in the roots of tobacco plants in response to leaf damage and transported via the xylem to the leaves. The signal indicating that the leaves have been damaged, cuing the roots to produce nicotine, is apparently a hormone transported from leaves to roots through the phloem. Greater leaf damage induces more nicotine production. Because of the long-distance nature of this response and the time needed for biosynthesis, it takes from ten hours to several days for nicotine to increase to its maximum levels. What might be the advantages of belowground synthesis of nicotine, given the vulnerability of the plant during this prolonged wait? If nicotine production occurred largely in the leaves, herbivores consuming those leaves might severely compromise the plant's ability to defend itself from further attack. The site of the synthesis of this defensive compound is thus protected below ground from leaf-eating herbivores. Even if large amounts of the aboveground parts of the plant are damaged or destroyed, the plant can continue to defend itself.

The classic example of an apparently induced defense is the production of cyanogenic glycosides by *Trifolium repens* (white clover, Fabaceae) when the leaves are damaged; this response can also be induced by frost damage. In fact, these compounds are already present in the plant but are activated upon chewing by herbivores. Cyanogenic glycosides are an effective deterrent against snails, which are voracious herbivores in regions with mild, wet winters. The ability to produce these defensive compounds is controlled by two genes, and populations of white clover in Great Britain are polymorphic for the trait (Dirzo and Harper 1982a, b). In the presence of snails, plants that had the ability to produce cyanogenic glycosides had higher survival rates than those that did not. In the absence of snails, however, those plants had reduced growth and reproduction compared with acyanogenic (non-cyanide-producing) plants. As predicted by evolutionary theory, areas with high snail densities had a high proportion of cyanogenic individuals, while those with fewer herbivores had a preponderance of acyanogenic individuals.

In a study of *Lepidium virginicum* (pepperweed, Brassicaceae), Anurag Agrawal (2000) showed that induced defenses were effective against a generalist herbivore, but not against an herbivore that specialized on that plant (Figure 11.17). Whether provided with a choice of food items or not, generalist caterpillars did substantially more damage to plants whose defenses had not been induced. However, specialist caterpillars produced the same amount of damage in induced and noninduced plants.

The production of sap by conifers in response to attack by bark beetles is another example of an induced response. In addition to the physical defense provided by the sticky, suffocating pitch, this response has a biochemical function. Coniferous sap contains substantial amounts of various terpenes (including monoterpenes and sesquiterpenes), as well as phenolics and other compounds, that are particularly toxic to the symbiotic fungi that attack with their beetle partners. Kenneth Raffa and colleagues (Raffa and Berryman 1987; Raffa 1991) found that wounding a tree mechanically caused

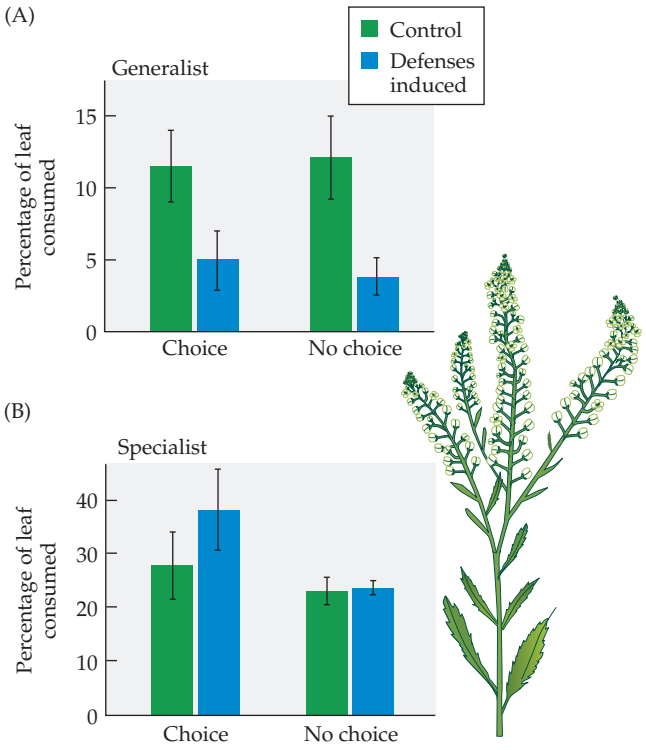



**Figure 11.17** Induced defenses in *Lepidium virginicum* (pepperweed, Brassicaceae) were effective against (A) caterpillars of generalist herbivores (noctuid moths), but not against (B) caterpillars of the specialist butterfly *Pieris rapae* (Pieridae). The defenses were induced by allowing a fixed number of larvae to feed on plants several days before the experimental trials. The amount of damage was the same whether or not the caterpillars had a choice of induced or noninduced plants. Error bars are ±1 standard error. (After A. A. Agrawal. 2000. *Ecology* 81: 1804–1813.)

the production of small amounts of monoterpenes, while trees attacked with living fungi responded with massive amounts of monoterpenes. Trees responded most strongly to those particular fungi associated with the bark beetle species that ordinarily attack them.

## Evolutionary consequences of plant-herbivore interactions

You may be wondering, after reading about the arsenal of defenses possessed by plants, why all herbivores do not simply starve to death. There are several reasons why they do not. First, fortunately for the rest of the biological world, plants are not uniformly well defended. Species, populations, individuals, life stages, and plant parts vary in their defenses, and the less well defended are often sought out and depended on for food. Second, some herbivores have evolved various means of overcoming plant defenses, leading to what has been termed a "coevolutionary arms race" between consumed plants and their would-be consumers.

The behavior, biochemistry, and morphology of herbivores may contribute to their ability to overcome plant defenses. Some insects, for example, are capable of carefully avoiding the ducts and canals that carry toxic latex in certain plants. Even more surprisingly, some chrysomelid beetles neatly incise these canals "upstream" before beginning to feed, preventing the flow of the toxic compounds to the leaf they are consuming. Giraffes, with their tough mouths and phenomenally long tongues, freely browse *Acacia* (Fabaceae) in the African savanna, undeterred by the plants' abundant and sharp thorns (Figure 11.18). In the northeastern United States, white-tailed deer avidly consume vines of *Smilax glauca* (catbrier, Liliaceae), at least when the stems are young, despite their ferocious hooked spines; it is not clear how the deer avoid being "clawed." Clearly, one cannot always depend on broad generalizations in making predictions about patterns of herbivory. One must understand the biology of the species involved.

©Danita Delimont/Alamy Stock Photo

**Figure 11.18** Giraffes browse *Acacia* (Fabaceae) trees despite the long, sharp thorns that protect these species.

A number of leaf-eating herbivores have symbioses with microbial species that can digest cellulose. Almost no animals can digest cellulose, a major component of plant tissue, but various bacterial species can. Ruminants (including deer, cattle, and antelopes) depend on bacterial fermentation of their food, which occurs in a specialized stomach compartment called the **rumen**; the animals receive nourishment from the fermented material and from digesting some of the bacteria. Cellulose-digesting bacteria are also symbiotic with other groups of mammals and, most notably, with termites.

Grasses and grazers provide examples of mutual adaptations of plants and their herbivores. While many generalist leaf eaters are effectively discouraged from eating grasses, a large group of mammals has evolved to depend on them as a primary food source. In most mammals, the teeth cease to grow in adults, but those species that are adapted to eating grasses have teeth that grow continuously. As the silica in the grass erodes their teeth, new growth replaces the worn material. Grasses are highly adapted to being grazed as well. Grasses, like all plants, grow from regions of actively dividing, undifferentiated cells called meristems (see Chapter 6). The meristems of grasses are generally located at ground level, out of the way of a hungry grazer's teeth. Thus, the plants are able to rapidly regrow the tissue lost to grazers, at least under ideal conditions. Heavy grazing, however, may result in destruction of these meristems, or in plants being torn up by the roots. Low resource levels or repeated grazing can also compromise the ability of the plants to produce new leaf tissues.

Many specialist herbivores are able to detoxify or sequester secondary compounds of their preferred host plants that would deter or kill other herbivores. It has long been known, for example, that the caterpillars of monarch butterflies (*Danaus*) are specialists on milkweeds (*Asclepias*, Apocynaceae) (Figure 11.19). The milky latex of milkweeds contains cardenolides, which are highly toxic to most insects and act as cardiac poisons in vertebrates. While most herbivores avoid these plants, monarch butterfly larvae are capable of sequestering the toxins, accumulating them in their own bodies to protect themselves from the birds that are their own predators.

However, there is more to this classic evolutionary ecology story. Stephen Malcolm and Myron Zalucki (1996) found that *Asclepias syriaca* (common milkweed, Apocynaceae) plants rapidly spike their levels of cardenolides after they are wounded, but drop those levels immediately afterward. This specialized pattern of induction, reasoned the investigators, allows the plants to kill small monarch larvae, which are sensitive to the cardenolides. But the subsequent decrease reduces the supply of cardenolides to the larger, older caterpillars, which have the

Courtesy of S. Scheiner

**Figure 11.19** A monarch caterpillar on *Asclepias syriaca* (common milkweed, Apocynaceae).

ability to sequester the chemical. Other specialist herbivorous insects have also evolved the ability to sequester plant secondary compounds and use them for their own protection, such as cucumber beetles that use cucurbitacin from cucumbers to deter invertebrate predators.

## 11.5 Plants Are Involved in Many Kinds of Trophic Interactions

### Some plants are parasites of other plants

Over 4000 plant species in 16 families are parasites on other plants, but parasitic plants are hardly ever abundant in nature. Parasitic plants become connected to the phloem or xylem of a host plant and obtain their water, carbon, minerals, and other materials from the host. They reduce the fitness of the host, thereby functioning somewhat like herbivores. Although parasitic plants do not usually cause the outright death of the host, they may reduce its competitive ability and reproductive output.

Parasitic plants may live almost entirely within the tissues of the host, or they may be found largely on the surfaces of leaves, branches, or other organs. There are obligate parasites, which can survive only by parasitizing other plants (see Box 12A), and **hemiparasites**, which can live either as parasites or independently. Hemiparasites may be photosynthetic, supplying some of their own needs, or obtain all of their energy and nutrition from the host. For example, *Castilleja indivisa* (Indian paintbrush, Orobanchaceae; Figure 11.20), like many other species in its family, is a root hemiparasite; this family accounts for over 40% of all parasitic species.

Mistletoes, a common name for a convergent group of species in the Santalaceae, Loranthaceae, and Misodendraceae, constitute about a quarter of all parasites and 75% of obligate parasitic species. They often have distinctive life histories and growth habits. The

Courtesy of the USDA NRCS Texas State Office

**Figure 11.20** *Castilleja indivisa* (Indian paintbrush), like many other species in the Orobanchaceae, is a hemiparasite, able to live independently or as a parasite on the roots of other plants.

mistletoe *Tristerix aphyllus* (Loranthaceae) is a Chilean species that lives within a columnar cactus, *Echinopsis chilensis*; only the flowers emerge through the cactus stem to reproduce. In parts of North America, mistletoes are common on oaks, some pine species (Figure 11.21), and *Prosopis* species (mesquites). Mistletoes are generally dispersed by birds. They often have sticky, mucilaginous seeds that adhere to the birds, which then carry them to other plants. Other families with ecologically important parasites include the broomrapes (Orobanchaceae) and dodders (Cuscutaceae; see Figure 6.4). Some parasitic plants, particularly the dodders

Courtesy of J. Schmidt/NPS Photo

**Figure 11.21** Mistletoes are a diverse group of obligate parasites. This mistletoe is growing on a pine in Yellowstone National Park.

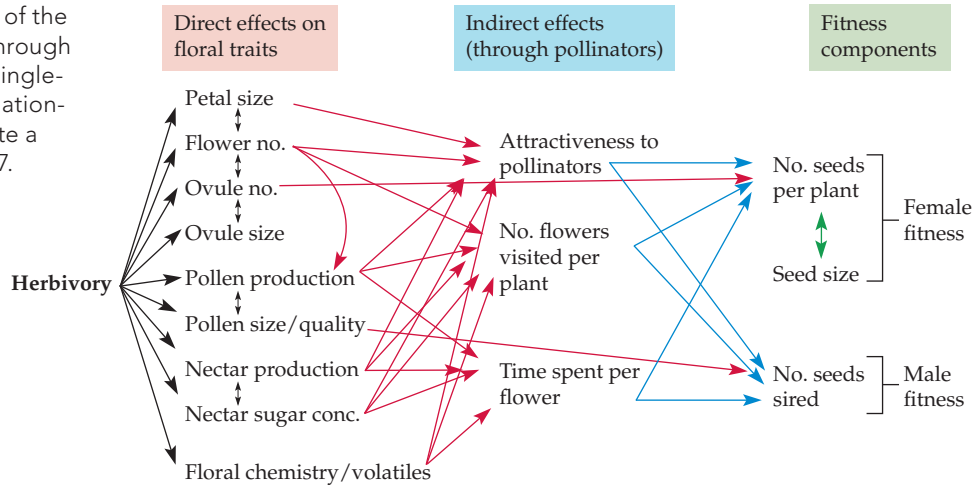



**Figure 11.22** Conceptual model of the effect herbivory exerts on fitness through its effect on floral characteristics. Single-headed arrows indicate a causal relationship; double-headed arrows indicate a correlation. (After S. Y. Strauss. 1997. *Ecology* 78: 1640–1645.)

and some mistletoes, cause substantial economic damage to crops and forest trees and may affect wild plant populations, but in general parasitic plants do not have major ecological effects.

The parasite load on individual plants is quite uneven within populations. At any given site, a few individuals may have a high load of parasitic plants, while most have only a few or even none. One hypothesis to explain this pattern comes from the observation that birds that specialize on the seeds of the parasites defend heavily laden trees. The birds may be depositing the seeds back on those trees, creating a positive feedback loop in which parasitized individuals become ever more heavily parasitized over time.

Parasitic plants can tap not only the host's nutrients and water, but sometimes its antiherbivore defenses as well. Sharon Strauss (1997) suggested a conceptual model of the effects of herbivory on plant fitness (Figure 11.22). Her ideas were tested in a study of the annual plant *Castilleja indivisa* (see Figure 11.20). Lynn Adler and her associates (2001) grew *Castilleja* plants on hosts belonging to two different strains of *Lupinus albus* (Fabaceae): one strain that produces a high concentration of alkaloids and one that produces a low concentration. Half of the *Castilleja* plants from each treatment were then sprayed with an insecticide, which both deterred herbivores and limited pollinators. Thus, the experiment consisted of high and low plant alkaloid levels combined with pesticides or without them. The researchers were able to estimate the effects of herbivory on *Castilleja* fitness (measured as seed production), both directly and indirectly through its effects on pollination. They found that the alkaloids taken up by *Castilleja* plants from the high-alkaloid strain of *Lupinus* had a positive effect on *Castilleja* fitness because the alkaloids reduced herbivory on buds, leaving more flowers and resulting in increased pollinator visitation rates.

## 11.6 Plants Interact with Pathogens, Endophytes, and Mycorrhizae in Complex Ways

Plants interact with bacteria, fungi, and viruses in a wide variety of ways. Plants suffer from diseases just as animals do. Pathogens can affect all life history stages and attack all plant parts. Fungi, water molds, bacteria, and viruses are the most common infectious agents in plants. Water molds (Oomycota) were once classified as fungi but are actually more closely related to brown algae. Plant diseases can kill individuals or reduce their size and fitness. They can have a sudden onset, with dramatic effects on populations, or they can be chronic, persisting in a population for decades. They can affect all the plants in a local community or over the entire distribution of the species, or they can affect only certain individuals, leaving many others untouched. Plant pathogens can interact with herbivores in a variety of ways, perhaps most importantly when herbivorous insects act as vectors to transmit them.

Plant pathogens range from highly host specific to broadly generalist, infecting a wide diversity of taxa; some pathogens require or may utilize two or more alternate host taxa. They can alter plant population dynamics and the composition of plant communities and can have large indirect effects on human communities as well (Box 11A). Pathogens can also cause evolutionary changes in populations. They can increase genetic diversity in populations and be a factor in maintaining species diversity in communities (Gilbert 2002). The ecology of plant-herbivore interactions has received much greater attention than the ecology of plant-pathogen interactions, which has largely been the province of agricultural scientists studying crop diseases. The twenty-first century, however, has seen

## BOX 11A "Broken" Tulips and the Tulip Mania of the 1600s

Tulips are native to central and western Asia and were imported to Europe in the late sixteenth century. By the early seventeenth century they had grown wildly in popularity, particularly in the Netherlands, where the price of the bulbs for these vividly colored flowers was driven higher and higher. Flowers with a rare color pattern of stripes, flames, intricate featherings, streaks, and a swirling mix of brilliant colors—called "broken" tulips—were particularly coveted, leading to frenzied speculation and spectacularly high prices. Tulips, including those with the broken color patterns, were the subject of many famous Dutch paintings. For the most expensive tulips, a single bulb sold for three times what Rembrandt was to be paid for his most famous painting, *The Night Watch*, and 20 times the annual salary of a carpenter. But in spring 1637, the bubble burst, and the bottom fell out of the market. Fortunes were ruined. It was not until the early nineteenth century that it was discovered that the broken color was caused by a virus. The virus affects the epidermis of the flower petals, causing colors to fade or intensify, with anthocyanins accumulating in vacuoles, creating the striking color patterns. The virus eventually kills the bulb. Today, far more modestly priced tulips are a major industry in the Netherlands, exported throughout the world, and virus-free tulips with an imitation of the broken pattern are popular—but they do not have the gorgeous and almost hypnotic appeal of the original "Rembrandt tulips" that caused speculators to go wild in the 1600s.

Courtesy of Rijksmuseum

Geel-rode tulp (yellow-red tulip), a "broken tulip." Anonymous, from a Dutch painting, 1700–1800. A virus creates the intricate color patterns.

a great increase in attention to both the ecology and evolutionary biology of plant-pathogen interactions.

We have very little understanding about why a pathogen that has been present in the environment suddenly becomes virulent. In some cases, the pathogen is introduced from elsewhere (as in the case of Dutch elm disease and chestnut blight in North America). In other cases, such as potato blight (Box 11B), the sudden spread of the disease may be a result of ecological factors, such as the response of the pathogen to weather, or of an evolutionary change in the pathogen, but we really do not know. Much more research is needed on the ecology of disease emergence.

### Plants are attacked by many different disease-causing organisms

Different pathogens attack different parts of plants and cause disease at different life history stages. Fungi are

## BOX 11B Effects of Plant Disease on Humans: Potato Blight and the Irish Potato Famine (the Great Famine)

The plant disease known as potato blight resulted in the horrific Irish potato famine in Ireland in 1845–1848. Potatoes were introduced to Europe from South America around 1600. *Phytophthora infestans*, the water mold that causes potato blight, can kill a field of potatoes within a few days under certain weather conditions. In one week during the summer of 1846, almost the entire potato crop of Ireland was killed. Due to political and economic factors, large parts of the population of Ireland had become dependent on potatoes as a basic food source. They almost certainly relied on a small number of susceptible genotypes, because potatoes are propagated asexually. Approximately a million people (about 12% of the population of the country at that time) died in the famine and disease that followed the blight, and more than another million emigrated to the United States and Canada rather than face starvation. Callous political decisions and economic inequality greatly exacerbated the deaths due to the famine. Those who emigrated from Ireland left an enormous cultural imprint on their adopted countries, particularly in cities such as New York and Boston. The population of Ireland never recovered; from over 8 million just before the famine, it was still under 5 million in 2018. The blight also affected continental Europe, starving about 100,000 people and leading to widespread political unrest and upheaval. Potato blight also killed large proportions of the potato crop in the United States and Canada at about the same time, but the diets and economies of these countries were far more diverse, and the disease had far less impact there.

the pathogens most likely to kill seeds, either in the soil or on the maternal plant. The tropical forest tree *Strychnos mitis* (Strychnaceae) in Uganda was found to suffer almost 90% mortality of fresh seeds, although consumption and passage through the guts of monkeys dramatically reduces this seed mortality (Lambert 2001). More commonly, fungi kill seeds in soil seed banks. While it seems reasonable to assume that the longer seeds remain in the soil, the more vulnerable they are to attack by fungi, this assumption has not been well investigated.

Plants are probably most vulnerable to pathogens as seedlings. A disease called "damping-off" is an important cause of seedling mortality in tropical forests (Augspurger 1984, 1988) and also occurs in temperate ecosystems (Packer and Clay 2000). Water molds such as *Pythium* and *Phytophthora* are major causes of damping-off. Damping-off may occur before or after the seedling emerges. Seedlings are attacked at or below the soil surface; the seedling wilts as the soft stem tissue collapses, and the seedling eventually topples over. The pathogen then spreads throughout the plant, killing it. Damping-off pathogens can also cause root rot in older plants.

Many different pathogens attack leaves and fruit. Fungi, bacteria, and viruses can destroy photosynthetic tissue, thereby reducing photosynthetic leaf area, reducing plant size or reproductive output, leading to leaf loss, and increasing the risk of plant mortality. For example, Asian soybean rust is caused by a fungus, *Phakopsora pachyrhizi*, that infects the leaves of many cultivated and wild plants in the Fabaceae. It has spread rapidly worldwide in recent years and was apparently carried to the southern United States from South America by Hurricane Ivan in September 2004. Soybean rust survives only on green tissue, so infestations in the Midwest are eliminated each fall. However, noncultivated plants such as the highly invasive vine *Pueraria montana* (kudzu, Fabaceae) in the southern United States are now a permanent reservoir for the fungus, from which crops and native plants may be reinfested every year through spores carried by the wind. Little is known about this kind of interaction between pathogens, wild plants, and crops, and it is an important topic for future research.

Fruits are also subject to diseases. The bacterium *Xanthomonas axonopodis* pathovar *citri* is an example of such a fruit disease. The pathogen causes citrus canker; infected *Citrus* (Rutaceae) trees drop their fruits and leaves prematurely, so this disease can be devastating to citrus crops. It is currently overrunning citrus trees in Florida. It has devastated citrus agriculture there and elsewhere and may completely eliminate citrus growing in that state, a major supplier of these fruits to the U.S. Undoubtedly similar pathogens affect wild plants, although these pathogens have been poorly studied. Fruits are generally most vulnerable to attack by various pathogens when the protective outer skin is damaged by insects or by bruising, although some pathogens can also enter the fruit at the vulnerable juncture where it is connected to the petiole. (You may notice this when fruits like apples and pears begin to rot at their cores, while the outside of the fruit appears unblemished.) Pathogens that attack fruits may cause them to drop before they are mature or may make them less attractive or palatable to seed dispersers.

Root diseases and butt rots (diseases that affect the base, or butt, of trees) can cause stunting and death in mature trees. Many of these diseases are caused by fungi and water molds. The water mold *Phytophthora cinnamomi*, which has been introduced to *Eucalyptus* forests in Western Australia, causes high levels of mortality in many tree species. Many fungi native to western North America, including laminated root rot, *Phellinus weirii*, and annosus root and butt rot, *Heterobasidion annosum*, can quickly kill large numbers of mature conifers, resulting in striking changes in forest structure. Root diseases and butt rots generally invade wounds created by other causes, such as wood-boring insects, and then travel through the tree via the cambium or heartwood at the tree's center.

Plants are also vulnerable to sexually transmitted diseases called smuts that affect flowers. Smuts are caused by several fungi in the Basidiomycota. They are important floral and fruit diseases of many grasses (including economically important grains such as wheat, corn, and sugarcane) and rushes. They also infect plants in other families, including the Caryophyllaceae, Dipsacaceae, Liliaceae, and Portulacaceae. The anthers of plants infected with anther smuts, for instance, produce fungal spores instead of pollen (Figure 11.23), and the spores are transmitted by pollinators. These plants can no longer reproduce sexually, and instead they become agents

Courtesy of Michael Hood

**Figure 11.23** Flowers of *Saponaria ocymoides* with anthers infected with a smut fungus (note their sooty black appearance). The anthers have been transformed and now produce smut spores rather than pollen.

for fungal reproduction (Jarosz and Davelos 1995; Alexander et al. 1996). For example, insects visiting infected flowers of *Silene latifolia* (bladder campion, Caryophyllaceae; see Figure 14.7B) spread the anther smut fungus *Ustilago violacea* (Thrall and Antonovics 1995).

Rusts may also attack flowers and other plant parts in addition to leaves. In some cases, they dramatically alter the morphology of flower parts. For instance, Barbara Roy (1996) found that infection of an *Arabis* species (rock cress, Brassicaceae) by the rust fungus *Puccinia monoica* prevents the plant from reproducing by causing the plants to produce showy pseudoflowers instead of normal flowers. The pseudoflowers attract pollinators that carry the fungus to other plants.

An important group of plant pathogens responsible for systemic infections (infections of the entire plant) is the mollicutes, tiny wall-less bacteria that include mycoplasmas and spiroplasmas. The mollicutes are important pathogens not only of plants, but also of humans, other vertebrates, and arthropods; those that affect plants are referred to collectively as phytoplasmas. These pathogens can cause a variety of diseases that infect both herbaceous plants, such as grasses, and woody plants, including both crops and wild species. Spiroplasmas (helical, motile mollicutes) can be transmitted between insects and plants (including major crops) and can cause disease in both. One of the most common types of plant disease caused by phytoplasmas are "yellows" diseases. Plants with yellows typically experience leaf yellowing and stunting. These diseases can also result in abnormal morphology, with many small, branched, axillary shoots growing from the nodes, giving them a bunchy appearance called a "witches' broom" and preventing flowering or further growth of the stem or branch. Infected flowers may change into leaflike structures that do not produce fruits or seeds. Eventually, the plant may be killed. (Witches' brooms can be caused by many different factors in addition to pathogens, including salt spray and parasitic plants, as well as genetic mutations.)

Other systemic pathogens are responsible for diseases known as cankers and wilts. Cankers begin with **necrosis** (localized tissue death) of the bark and cambium on stems or branches, while wilts cause disruption of the xylem, but both can ultimately be fatal. These pathogens can sometimes cause widespread death of mature trees, as we will see below.

### Plants have immediate defenses and long-term evolutionary responses to pathogens

Some of the chemical defenses against herbivory described above also, or even primarily, serve as defenses against fungal, bacterial, and viral pathogens. The hormonal events that signal herbivore damage appear to be quite distinct from the series of biochemical events signaling pathogen invasion, which is generally mediated by salicylic acid (rather than the jasmonates).

**Phytoalexins** are secondary compounds that act as specific defenses against pathogens. Phytoalexins are produced at the site of an infection to kill microbes. A host of different classes of chemical compounds are known to function as phytoalexins, including many of the chemical defenses mentioned above. Through mechanisms that are not yet fully understood, plants can also acquire localized and even systemic resistance to pathogens. The systemic resistance observed in plants is analogous to the acquired immunity humans develop to some diseases, although it is very different biochemically, physiologically, and evolutionarily.

A physical defense against infection is **phloem plugging**, in which the phloem clogs up in response to pathogen infestation, preventing the spread of the infectious agent through the vascular system of the plant. Similarly, in some plants, cells in and around the infected tissue die, sealing off the area so that the infection will not spread. This tissue death reduces the amount of photosynthetic surface area and can therefore reduce the plant's growth rate.

**Pathogenicity** is the ability of a microorganism to cause disease in a host. Infection by a pathogen depends on interactions between resistance genes in the host and avirulence genes in the pathogen. **Virulence** is used both to mean the ability of an infectious agent to produce disease, and a measure of the degree of damage inflicted by a pathogen on the host organism. Avirulence genes code for biochemical products that may be recognized by the products of resistance genes in the host, preventing infection. Pathogens lacking avirulence genes for a particular host are not recognized by that host and are able to infect it. Resistance to a pathogen is the ability of a potential host to reject infection by that pathogen, preventing colonization by the pathogen and the onset of disease. Resistance genes code for proteins that are believed to act as receptors that recognize particular biochemical products of avirulence genes in specific pathogens. Modification of an avirulence allele to a virulent form that avoids recognition by a host enables the pathogen to circumvent that host's defenses.

Host shifts—the ability to infect new taxa of hosts—may be the result of evolutionary changes in the pathogen genome or of ecological changes that are not determined by evolutionary genetic change (Parker and Gilbert 2004). Plant-pathogen interactions create ecological dynamics, as some plant genotypes and species succumb to disease, allowing resistant taxa to take their places. They also result in coevolutionary dynamics, as fluctuating selection acts on both pathogens and plants. Both processes may act to maintain genetic diversity among individuals within populations as well as species

diversity at the community level in both pathogens and plants. Because both resistance genes and virulence alleles of avirulence genes appear to have costs, polymorphisms are maintained in both hosts and pathogens. Without such costs, plants should logically evolve to become resistant to all pathogens; of course, that prediction makes the unrealistic assumption that genetic variation exists for resistance to all possible pathogens. For example, some avirulence genes are needed by the pathogen to promote growth or reproduction within the host, so it is not possible for the pathogen to evolve other forms of these genes that are not recognized by the plant.

In contrast to initial infection, the severity of disease is determined by many different traits in both host and pathogen. Pathogens depend on their interaction with the plant to colonize and reproduce in the plant. They produce toxins that break down cell walls and detoxify host defenses. Genetically based traits that confer tolerance to pathogens include many of the same physical defenses that function against herbivores, enzymes that break down fungal cell walls, and enzymes that detoxify pathogen toxins. **Tolerance** of a disease organism is the ability of a plant to maintain its fitness when infected with a pathogen.

## Pathogens can shape plant populations and communities

Plant disease incidence varies spatially. Physical factors, such as deep shade and moist soils, as well as biotic factors, such as density and proximity to related plants, can increase the incidence of attack by fungal pathogens on both seedlings and mature plants. This can potentially have many consequences. One suggested explanation for the high local diversity in tropical rainforests is the effect of predators (insects, mammals, and birds) and pathogens (primarily fungi and water molds) on seeds and seedlings. Daniel Janzen (1970) and Joseph Connell (1971, 1978) proposed a hypothesis that offers a conceptual framework for this idea. The Janzen-Connell hypothesis argues that if specialist natural enemies are concentrated around maternal parent plants, conspecific seeds and seedlings close to the parent are more likely to be killed. Consequently, there will be strong selection for dispersal away from the parent plant. If only heterospecific individuals survive near mature trees, this will result in high diversity as the mature trees are eventually replaced by saplings belonging to other species. Many observational and experimental studies have been conducted to test this hypothesis, but reviews and meta-analyses have come to different conclusions about its robustness, with some finding limited or no support and others suggesting that the patterns it predicts do occur. The Janzen-Connell hypothesis rests on the assumption that each tree species has a specialist pathogen that differentially kills its own seedlings. In reality, the pathogens that most commonly kill seeds and seedlings are typically generalists that do not differentiate between tree species. While this may be a fatal flaw of the hypothesis, the reality may be more nuanced and complex. For instance, there may be different genetic strains of the pathogens, and some pathogens may be fatal under some circumstances but harmless under other conditions, complicating whether seedlings really are more likely to be killed near the parent tree or not.

One of the best-known and most dramatic examples of a pathogen altering plant communities is the canker disease called chestnut blight. The American chestnut (*Castanea dentata*, Fagaceae; Figure 11.24) once extended over large parts of the eastern deciduous forests of North America but has been virtually wiped out by the chestnut blight. The fungus enters a tree through wounds caused by other factors and grows under the bark to the cambium, eventually killing the cambium around the entire circumference of the trunk. While the tree can resprout from its base, the sprouts also eventually become infected and die. The disease is caused by a fungus, *Cryphonectria parasitica*, which first appeared in 1904 in New York City, apparently introduced from Asia on nursery stock of infected Asian chestnut trees. By the mid-1920s, the pathogen had spread throughout the range of the American chestnut. As the infection spread, trees that had not yet been killed by the fungus were cut for their valuable lumber before the fungus could destroy the wood, eliminating the chance for selection to occur for potential disease resistance.

Courtesy of S. Scheiner

**Figure 11.24** *Castanea dentata* (American chestnut, Fagaceae) tree in full flower in northern Michigan U.S.A., far outside its normal range. This individual was undoubtedly planted, and only its great distance from natural populations has allowed it to avoid infection by the chestnut blight fungus and reach reproductive size.

The indirect effects of this pathogen greatly altered the composition, food webs, and ecosystem ecology of the communities in which the chestnut had once been abundant. Today, a number of oaks (*Quercus*, Fagaceae) and other tree species dominate many of these forests. Food webs were substantially changed in some forests as a species that produced large crops of nuts each year was replaced by oaks that typically have large acorn crops only once in several years, followed by years with minimal crops. Efforts are currently under way to breed resistance genes from Asian chestnut species into American stock, or to use genetic engineering techniques to transfer those resistence genes, with the ambitious goal of wide replanting of resistant American genotypes (Burnham 1988). The American Chestnut Foundation was established with the goal of restoring this species throughout its historical range.

Plant pathogens continue to alter plant populations and communities on a very large scale. Sudden oak death is a rapidly spreading disease that first appeared in the 1990s (Figure 11.25). It has particularly affected *Quercus* species (Fagaceae) and *Notholithocarpus densiflorus* (tanoak, Fagaceae) in forests and woodlands in coastal California and Oregon, but it has also spread to native shrub species, including *Arbutus* species (madrone, Ericaceae) and *Arctostaphylos manzanita* (manzanita, Ericaceae), and cultivated shrub species. These shrubs may harbor the pathogen without being killed themselves, which may lead to its increase and spread. Caused by the water mold *Phytophthora ramorum*, this disease has already killed tens of thousands of trees on the West Coast and has spread on nursery plants to the East Coast but not yet escaped to forest trees. Much of the eastern and Gulf regions of the U.S. are vulnerable. The disease has spread rapidly on cultivated species in the United Kingdom, Ireland, and some other European countries. The origins of the pathogen are not known, but the networks of spread have been documented using genomic analyses. There is great concern that it will continue to spread to other geographic areas, due to the widespread distribution of appropriate host trees and compatible climates.

In the grasslands of central California, in areas not dominated by intensive agriculture, invasive annual grasses have largely replaced the original extensive grassland vegetation of perennial grasses and herbaceous dicots (often called *forbs*). (Many factors are responsible for this change, including competition, grazing, and changes in fire regimes, but we will not explore that complex story here.) Carolyn Malmstrom and colleagues (2005) hypothesized that plant diseases also might play a role in the decimation of the native perennial grasses. *Elymus glaucus* (blue wildrye, Poaceae) is a native perennial grass, but it has largely been replaced with *Avena fatua* (wild oats, Poaceae), an invasive, widespread, and dominant annual. In a series of experiments, they found that the incidence of viral disease in the native *E. glaucus* (caused by barley and cereal dwarf viruses) was more than twice as high when it was grown with the invasive *A. fatua* than when grown alone. In contrast, the presence of the native grass

Courtesy of Carolyn Malmstrom

Carolyn Malmstrom

(A)

(B)

Both photos courtesy of S. Scheiner

**Figure 11.25** (A) Sudden oak death causes tree wounding, as shown here infecting oak trees in coastal California, U.S.A. (B) Dying oaks can cause shifts in forest species composition as sudden oak death selectively kills dominant trees.

had no effect on disease in *A. fatua*. The virus is carried by aphids, and aphids were 50 to 800 times more dense when *A. fatua* was present than when the native *E. glaucus* was grown by itself. The authors suggest that the increase in disease-carrying aphids facilitated by the presence of *A. fatua* may have contributed to the decline of the native grass species.

### Plant pathogens can interact in complex ways with other organisms

While insects are associated with the spread of many plant diseases, in some cases there is a closer symbiotic association between an insect and a pathogen. Bark beetles are often associated with fungal species that interact in many different ways with the beetles and the trees. The fungi may be present on the insects' surfaces, carried in specialized structures, or may be carried by commensal mites that are carried by the beetles. A tree is infected with the fungi when the beetles tunnel into the cambium. Some fungal species help feed the developing beetle larvae, while other fungal species are serious tree pathogens. The fungi which cause blue-stain disease can destroy the tree's vascular tissue, killing it, although whether these fungi are the primary cause of tree death is uncertain. Some fungal species can also cause the mortality of beetle larvae by out-competing other fungi that the larvae feed on, and these fungi can also kill the tree. Scientists are just beginning to understand the complex interactions between different species of fungi, bark beetles, mites, and conifers, which vary with environmental conditions and the nutritional and water status of the tree.

Bark beetles are also responsible for the spread of Dutch elm disease, a wilt disease caused by the introduced fungus *Ophiostoma ulmi* (closely related to the blue-stain fungus), which has killed millions of *Ulmus americana* (American elm, Ulmaceae) trees in North America. Both beetles that are native to North America and introduced beetles from Europe are involved in its spread (the fungus was probably introduced from Asia; the term *Dutch* originated because the disease was first identified in the Netherlands). The fungus kills the trees by blocking water flow in the xylem and by producing a toxin.

In a study of the interaction between soil-dwelling pathogens, herbivores (fungi and nematodes), and plant succession in the Netherlands, Wim VanderPutten and Bas Peters (1997) found that the pathogens facilitated species replacement within the plant community. On coastal sand dunes, *Ammophila arenaria* (marram grass, Poaceae) is replaced as the dunes become stabilized by another grass, *Festuca rubra* ssp. *arenaria* (sand fescue). In a series of experiments, the researchers found that the soil pathogens severely reduced the competitive ability of marram grass relative to sand fescue. In the absence of the pathogens, marram grass was not competitively inferior to sand fescue. Thus, the interaction between competition and the effects of pathogens may be responsible for the decline of the marram grass on stabilized dunes and its replacement by sand fescue. Other factors may also affect the progress of disease in natural populations. Janet Morrison (1996) found that plant genotype, microhabitat, and the diversity of the plant community determined the rate of infection of *Juncus dichotomus* (path rush, Juncaceae) by a smut fungus, with the presence of neighbors belonging to other species reducing infection rates.

Pathogens can also interact with the effects of grazers. Michael Bowers and Christopher Sacchi (1991) studied the effects of herbivore exclusion in an early successional old field in Virginia. A dominant plant species, *Trifolium pratense* (purple clover, Fabaceae), increased in response to herbivore exclusion. However, these high-density, ungrazed clover populations became severely infected with the fungus *Uromyces trifolii*, which killed many of the plants. In contrast, grazed areas had much less disease. As a result, in the year following the infection, the density of clover plants was much higher in the presence of grazing than in its absence.

### Endophytes are symbiotic organisms that live inside plant cells

**Endophytes** are intracellular, symbiotic microorganisms. These symbiotic associations of bacterial and fungal cells living inside plant cells may affect plant function in many ways (see Box 4C). Their function is poorly understood, but they may be helpful, harmful, or apparently neutral in effect, or their effects may vary with the plant's condition, the environment, or the plant's interactions with other organisms, such as mycorrhizae (discussed in the next section). Some fungal endophytes may behave as saprophytes when the plant tissue dies, contributing to the decomposition of those tissues. Many different species of endophytes are found in plant cells, constituting a plant microbiome community.

Fungal endophytes, fungi that live within aboveground plant cells, are the best studied, particularly in pasture grasses. In many cases these fungi are mutualists that protect plants from herbivory and pathogens (possibly by the production of alkaloids) and increase their competitive ability. Their effects can range from positive to harmful (Clay 1990). Keith Clay and Jenny Holah (1999) found that they could alter plant community structure by increasing the dominance of infected grass species at the expense of biodiversity. Kari Saikkonen and collaborators (Saikkonen et al. 1998; Huitu et al. 2014) have found that these endophytes can alter the chemical ecology of the grasses they inhabit and defend against grazing mammals, as well as reduce plant competitive ability or cause disease. Although fungal endophytes are best known in grasses, they have been found in all woody plants that have been

examined for their presence, and they are known in trees, shrubs, and ferns. As with other symbioses with microorganisms, the plant may provide nutrients as well as carbon and energy from photosynthesis to the fungal cells. The endophytic fungi may also benefit the grass by protecting it from drought and nutrient deficiencies.

### Mycorrhizae are essential for terrestrial life

Plants are often severely limited by the nutrients available in soils (see Chapter 4). An interaction exists between plants and certain fungi that is extremely common, widespread, and of overwhelming ecological importance for enabling plants to survive and thrive, particularly under conditions of phosphorus limitation and other low-nutrient conditions. It is fair to say that this interaction is essential for life in most terrestrial ecosystems. The structures in which this interaction occurs are called **mycorrhizae**. Mycorrhizal symbioses between various fungi and the roots of terrestrial plants can confer a number of benefits to the plants. (The singular form is *mycorrhiza*, and the word refers to the plant-fungal symbiosis, not strictly to the fungus.) These interactions depend on a fungal species successfully infecting and living in or on the roots of a plant. Mycorrhizal symbioses are often (but not always) mutualistic (see Box 4C). Nutrients taken up from the soil by the fungal mycelium are transferred from the fungal cells to the root cells of the host plant; in turn, the fungus gains carbon and energy in the form of carbohydrates from the plant. Mycorrhizae can have other functions as well, as we discuss below.

Mycorrhizae are critical components of ecosystem function in many environments and play a major role in nutrient uptake by plants and in nutrient recycling in soils. About 80% of angiosperms and all gymnosperms are involved in some form of mycorrhizal symbiosis (Wang and Qiu 2006; Strullu-Derrien et al. 2018). Mycorrhizae occur in all plant divisions, including some mosses and ferns. A given host species can have associations with one or many fungal species, including different types of mycorrhizae. There are also some species that never form symbioses with mycorrhizal fungi; mycorrhizae are almost never found in association with plants in the families Brassicaceae and Amaranthaceae and are rarely found with Proteaceae or other species with proteoid roots (see Chapter 4). The kinds of plants with which different groups of mycorrhizae are associated, and some of the fungal taxa involved, are listed in Table 11.1 and illustrated in Figure 11.26. The overwhelming importance of mycorrhizae in natural ecosystems and throughout the plant kingdom is often overlooked, although with the development of molecular genetic and other analytic tools in recent years, it has become easier to identify and study mycorrhizae. Because mycorrhizal fungi and nitrogen-fixing bacteria (see Chapter 4) both form symbioses with plants and both occur on roots, many people confuse these two very different kinds of interactions that differ in a number of important ways. We cut this Gordian knot of confusion here. *Fungi* are complex, multicellular eukaryotes, while nitrogen-fixing *bacteria* are single-celled prokaryotes without a nucleus. Mycorrhizae are found on most plants in most places, while nitrogen-fixing symbioses are found only on particular plant groups, and only in particular environments. Nitrogen fixation is the process of transforming elemental $N_2$ from the atmosphere into ammonia and thus involves a single nutrient, nitrogen. Mycorrhizal

**TABLE 11.1** The associations of the five major types of mycorrhizae[a]

| Mycorrhizal group | Plant taxa involved | Fungal taxa involved |
|---|---|---|
| Ectomycorrhizae (ECM) | Dipterocarpaceae (98%), Pinaceae (95%), Fagaceae (94%), Myrtaceae (90%), Salicaceae (83%), Betulaceae (70%), Fabaceae (16%), and some others | Basidiomycota (most), Ascomycota (less common), and Zygomycota (rare) |
| Arbuscular mycorrhizae (AM) | By far the most common mycorrhizae in angiosperms, except for nonmycorrhizal families (Brassicaceae, Portulaceae, Caryophyllaceae, Proteaceae, and some others); also common in many gymnosperms, except Pinaceae, and in some ferns and other taxa | Phylum Glomeromycota (including families Gigasporaceae, Acaulosporaceaea, and Glomeraceae and others) |
| Arbutoid ectendomycorrhizae | Many species in the order Ericales | Basidiomycetes |
| Ericoid mycorrhizae | Many species in the order Ericales | Basidiomycetes and some Ascomycetes |
| Orchidaceous mycorrhizae | Orchidaceae | Basidiomycetes and some Ascomycetes |

*Source*: M. A. Selosse and F. Le Tacon. 1998. *Trends Ecol Evol* 13: 15–20.
[a] See Figure 11.26.

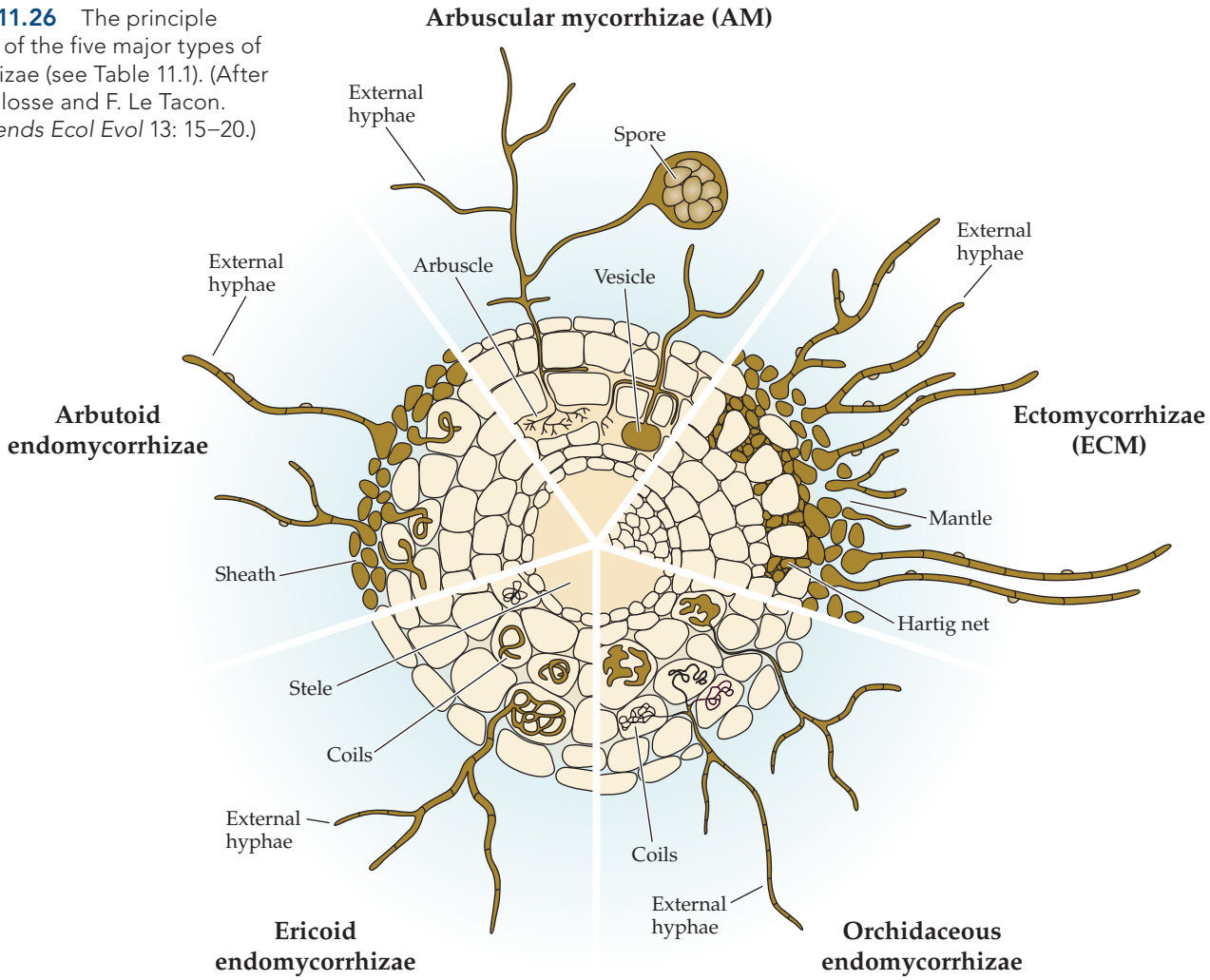



**Figure 11.26** The principle features of the five major types of mycorrhizae (see Table 11.1). (After M. A. Selosse and F. Le Tacon. 1998. *Trends Ecol Evol* 13: 15–20.)

symbioses, on the other hand, function in a wide variety of ways. They greatly facilitate the availability and absorption of many different nutrients as well as protecting plants from pathogens. Of course, you can always remember that mycorrhizae are the ones with the impossible spelling! A plant species can have both symbiotic nitrogen-fixing bacteria and mycorrhizae. For legumes growing in phosphorus-deficient soils, mycorrhizae can increase nodulation and nitrogen fixation by bacteria as well as growth of the plants on which they occur.

## Arbuscular mycorrhizae and ectomycorrhizae are the two most ecologically important groups

The most ecologically important and common groups of mycorrhizae are **arbuscular mycorrhizae** (**AM**) and **ectomycorrhizae** (**ECM**). There are also other more specialized groups of mycorrhizae (see Figure 11.26 and the discussion below). The most abundant and the most ancient mycorrhizae are the AM, which occur in both herbaceous and woody plant species (Redecker et al. 2000). The fungi in AM grow inside plant cells, with hyphae that extend outward into the soil. AM fungi occur on 75%–80% of terrestrial plant species. The AM are a monophyletic group that evolved over 400 million years ago. They were critical for plants' emergence from aquatic environments onto terrestrial habitats (Pirozynski and Malloch 1975) (see Chapter 17). Most grass species and individuals have AM, but AM are also important in trees, including both colonizing species and forest dominants. Most tropical trees form AM symbioses, although trees in some important tropical families are ectomycorrhizal (Brearley 2012; discussed below). In the eastern and midwestern U.S., common forest dominants that depend on AM include *Acer saccharum* (sugar maple, Sapindaceae), *A. rubrum* (red maple), and *Liriodendron tulipifer*a (tulip poplar, Magnoliaceae) (Phillips et al. 2013).

AM fungi assist their plant hosts to take up phosphorus by having hyphae (strands of living fungal filaments)

that extend beyond the zone around plant roots where phosphorus has been depleted by root uptake. AM are highly efficient at scavenging available phosphorus from the soil, due in part to the phenomenally large surface area of the many small hyphae that extend into the soil. The phosphorus taken up by the fungus is transferred to the plant after the fungus itself uses what it needs. More recently it was discovered that AM are also important in taking up both inorganic and organic nitrogen (Phillips et al. 2013). This ability is important in both temperate and tropical ecosystems, and especially important for many crops.

The body of the AM fungus grows as a delicately branched stucture, called an **arbuscule**, inside the cortical cells of the roots and in the intercellular spaces between root cells, with hyphae (fungal strands) extending out several millimeters or more into the soil (see Figure 11.26). AM associations are not highly specific; the same fungal species associate with many different host species, and an individual host plant or plant species may have a number of different AM fungal species that grow with it symbiotically. The community composition (and possibly function) of these fungi differs from one biome to another, and among local sites (Kivlin et al. 2011). The benefits derived by a plant from different mycorrhizal associations may vary greatly, and the growth rates of an AM fungal species can also vary depending on its plant host (Reynolds et al. 2003).

In constrast with AM fungi, ECM fungi have arisen multiple times in different phylogenetic groups of fungi and are highly diverse. Just under 8000 species of ECM fungi have been definitively identified, although it is possible there may be as many as 25,000 species in total, most as yet undiscovered (Rinaldi et al. 2008). They have arisen more recently than AM fungi (possibly about 200 million years ago) and are more specialized. Although ECM are found on only about 3% of plant species, they are ecologically very important because they are essential symbionts of the most common trees in many forests, including the dominant plants of the vast taiga (northern coniferous forest biome, see Figure 18.23) as well as some important temperate and tropical forest trees. This means that they dominate large areas on the Earth and that an enormous amount of biomass and carbon stores in soils consist of ECM tissues. Most ECM are found in association with trees rather than shrubs or herbaceous plants. In temperate and boreal forests they are found in association with tree species in the Pinaceae, Betulaceae, Fagaceae, and Salicaceae. In the eastern and midwestern U.S., ECM species include many gymnosperms as well as hardwoods such as *Quercus alba* (white oak, Fagaceae), *Q. rubra* (red oak), *Q. velutina* (black oak), and *Carya* (hickory, Fagaceae). They are symbionts of tropical trees in the family Dipterocarpaceae, which are major dominants in the tropical rainforests of Southeast Asia, and also of several other tropical and subtropical tree families. Dipterocarp trees in tropical Africa share an ECM fungal ancestor with those in Southeast Asia, implying that the symbiosis goes back before the separation of Gondwana (Brearley 2012) (see Figure 17.6). The relationship between ectomycorrhizal host plants and fungi is often highly specific, although there are some ECM fungi that associate with many plant species (such as *Amanita muscaria*, Figure 11.27) and some plant species that associate with many different ECM fungal species (e.g., some pines). ECM are particularly common in northern coniferous and temperate deciduous forests and in soils where nitrogen is particularly limiting. Most coniferous trees cannot grow, or grow very poorly, without these fungal symbionts.

ECM form a thick "woolly coat" (called a mantle) around root tips, with long, thick mycelial structures called rhizomorphs that can extend from millimeters to several meters into and across the soil. ECM are distinguished by two structures, one inside the root: the **Hartig net**, a complex of mycelia that grows in between the root cortical cells and enmeshes them, and one outside it: the **mantle**, a dense network of hyphae (called a mycelium) that partially or fully ensheathes the outside of the root (see Figure 11.26). Plant species with ECM grow short, stubby roots (Figure 11.28), a growth form that may be controlled hormonally by the fungus. ECM help plants to access otherwise inaccessible soil nutrients using a different mechanism than that used by AM. ECM fungi produce enzymes that are secreted into the soil and break down soil organic matter to retrieve both

Onderwijsgek/CC BY-SA 3.0

**Figure 11.27** *Amanita muscaria* is a common mycorrhizal mushroom in the Basidiomycota. It is both toxic and psychotropic. The "mushroom" is the reproductive structure of the largely underground mycorrhizal fungus and the one most associated with European fairy tales.

**Figure 11.28** (Top) *P. radiata* and *Amanita muscaria* ECM, showing highly branched short roots with many root tips coated in white fungal mycelia. (Bottom) Ectomycorrhizal association between *Pinus radiata* and *Suillus brevipes*. Notice the short roots ensheathed in the fungal mantle, and the external fungal hyphae extending outward. (From N. Malajczuk et al. 1982. *New Phytol* 91: 467–482.)

nitrogen and phosphorus, and sometimes other minerals. While AM roots are highly effective in scavenging nutrients from already decomposed organic matter in soils, the ECM fungi mine nutrients by decomposing organic matter themselves, similar to free-living saprophytic fungi.

Plants with mycorrhizae can take up several times the amount of phosphorus as nonmycorrhizal plants, and as a consequence, they can live and grow much more successfully on soils in which phosphorus is low. However, the phosphorus uptake of mycorrhizae is highly variable. Some species of AM fungi are highly effective at taking up phosphorus and transferring it to their hosts, while others provide very little phosphorus to their hosts. Environmental factors are also important in determining whether phosphorus uptake is enhanced by the fungus. At extremely low levels of phosphorus, the fungi are apparently not able to function well, and the degree to which roots become infected with the fungi is low. At slightly higher phosphorus levels, the fungi colonize roots more successfully and become highly effective at enhancing phosphorus uptake. At high levels of available soil phosphorus, fungal growth inside the root is apparently actively suppressed by the plant, and the roots take up phosphorus directly. AM fungi may be more active in phosphorus uptake when nitrogen availability is high and phosphorus is low.

In addition to enhancing phosphorus availability, AM increase plant access to some other minerals, such as copper and zinc. ECM tend to have much more diverse roles. If the fungal mantle of ECM totally encases the roots, all water and mineral uptake must take place through the fungus. Because the surface area of the mass of hyphae is far greater than that of the root system, the uptake of water, nitrogen, and other dissolved minerals may also be enhanced.

### Specialized mycorrhizal interactions include those associated with the Ericaceae and Orchidaceae

Several types of mycorrhizae are sometimes grouped as **endomycorrhizae**, including AM, Ericoid and Orchid mycorrhizae, because they all include structures that penetrate and live inside plant cells, in contrast to ECM, whose structures live between plant cells but not inside them. However, ericoid and **orchidaceous mycorrhizae** are quite distinct in important ways in terms of their ecology, morphology and evolutionary biology. **Ericoid mycorrhizae** are specialized for symbioses with plant species in the heath family, the Ericaceae (including heathers, blueberries, and rhododendrons; Figure 11.29, Table 11.1). Plants in this and some other families in the Ericales generally

Courtesy of G. Fox

**Figure 11.29** *Erica mammosa* (Ericaceae). Taken at the Ramskop botanical garden in Clanwilliam, Western Cape, South Africa.

live in habitats with acidic, extremely nutrient-poor soils. Ericoid mycorrhizae appear to be necessary for these plants to survive, grow, and reproduce in their natural habitats, and they clearly depend on these associations to obtain mineral nutrients from these infertile soils. Few, if any, nonmycorrhizal plants live in these habitats.

Orchids have a special, obligate relationship with orchidaceous mycorrhizal fungi, depending on them for ordinary growth and development. Orchid seeds are exceptionally tiny, with no stored food reserves. The presence and activity of mycorrhizal fungi are necessary for seed germination and early seedling development and growth. Particularly when the plant is young, the fungus may supply it with carbon and energy in the form of carbohydrates, rather than receiving them from the plant. Some orchids never become photosynthetic and depend on mycorrhizae throughout their life-span for all of their nutrients, including carbohydrates. It is not known where these fungi obtain their carbon, but they may get it by parasitizing other plants or by breaking down organic matter in the soil, functioning as saprophytes, like many nonmycorrhizal fungi. Not enough work has been done on mycorrhizae in orchids (particularly in the tropics, home to the majority of orchids) to fully understand the extent and nature of this extreme form of plant dependence. The plant family Orchidaceae has the largest number of species in the world, so orchidaceous mycorrhizae present an interesting and important topic for future ecological and evolutionary research.

Other groups of mycorrhizae have been characterized in various ways; one of these is the arbutoid **ectendomycorrhizae** (see Figure 11.26 and Table 11.1). This group has characteristics of both the ECM and ericoid groups but is not well understood or described.

### Mycorrhizae function in other ways in addition to nutrient uptake

Besides facilitating nutrient uptake, mycorrhizae protect plants from pathogens, reduce the negative effects of heavy metals in the soil, and help plants tolerate drought and other stresses. ECM fungi can enable their hosts to tolerate toxic heavy metals and other ions in the soil by sequestering and binding them in their cell walls. Some mycorrhizal fungi manipulate plant hormones to influence plant growth and form. AM fungi can alter the microbiome in the rhizosphere by supporting some bacteria and suppressing others. A surprising function of some AM fungi is the ability to create networks of hyphae in the soil that connect different individuals, even sometimes of different species. Nutrients and other materials may be transported through these networks from one plant to another.

Mycorrhizae can also exert a substantial effect on plant water relations. AM often decrease the negative effects of drought on the host plant. It is possible that this is due to indirect effects of alteration of plant morphology and nutrient status by the fungus. The fungal structures are capable of carrying water rapidly from the soil to the roots. In addition, the fine hyphae extending into the soil from AM may change the soil structure itself, increasing the movement of water toward the roots and helping to bind the soil particles together, decreasing erosion (which is especially important in sands and sandy soils, as in beach dunes). One of the most surprising effects of both ECM and AM is their protection of the host plant from bacterial and fungal pathogens and nematode attack. The protective mechanisms are diverse, particularly in ECM. The fungus may either secrete pathogen-killing chemicals directly into the rhizosphere or stimulate the root to produce such compounds. The mantle of ECM that sheaths the root can directly shield them from pathogens and root-consuming invertebrates.

Plant responses to infection by AM fungi may be another way in which plants gain protection against microbial pathogens, parasitic nematodes, and root-feeding insects. The fungal infection of the plant cells and initiation of the mycorrhizal symbiosis appears to trigger the immune responses of the plant by the jasmonate signaling pathway. This response then confers a state of heightened alertness to attack in which defenses can be more quickly and strongly mobilized when an attack occurs. Because the jasmonate signaling pathway functions throughout the plant, the protection stimulated by the AM fungus may also protect the aboveground shoots and leaves (Jung et al. 2012). Mycorrhizal fungi also have complex interactions with different bacterial species, which undoubtedly play a role in some of the ways they affect plants and ecosystem properties such as carbon cycling in soils (Nazir et al. 2009).

### Are mycorrhizal fungi mutualists or parasites?

You may wonder, why would any self-respecting fungus do so many nice things for a plant? Soil fungi are heterotrophic, like animals, and typically degrade organic matter or parasitize living things to make a living. (Others are predators, particularly on soil nematodes.) Mycorrhizal associations provide a way for fungi to obtain carbon and energy from a living host plant. They appear to often be genuinely mutualistic associations, with the plants supplying carbohydrates obtained by photosynthesis to the fungi and receiving mineral nutrients and other benefits from the fungi.

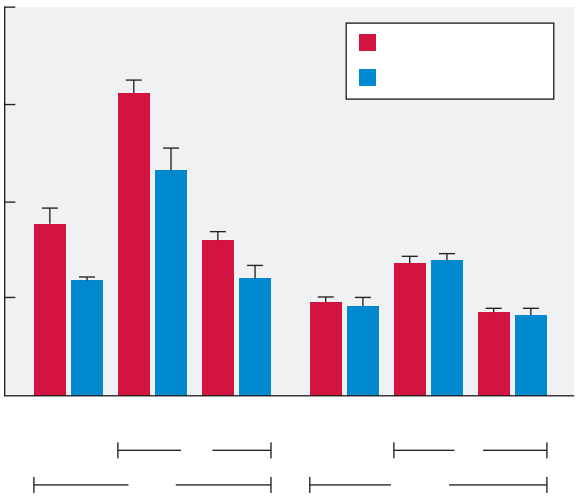
However, like any relationship, this one is not inevitably benign. One partner (either the plant or the fungus) can become parasitic on the other, receiving benefits without supplying them. Sometimes mycorrhizal fungi

do not provide any minerals or other detectable benefits to their host yet continue to receive nutrients from the plant. Sometimes the fungus provides only minor amounts of nutrients, the nutrients made available are not needed or do not enhance plant growth, or the benefits provided are vastly outweighed by the carbon and energy drain on the plant. But sometimes the tables are turned, and the plant provides little benefit to the fungus, while receiving benefits (see Box 4C).

High levels of phosphorus and possibly nitrogen in the soil may act to suppress infection by mycorrhizal fungi. Other factors, such as aridity, toxic minerals, and inadequate or excess concentrations of other minerals (such as, for example, too little boron), may inhibit the formation and function of mycorrhizae because these conditions either harm the fungal partner or function in some more complex way to inhibit the symbiosis. In some cases, the plant may be capable of discarding the fungus when phosphorus is available, accepting colonization by mycorrhizal fungi only when it is of benefit to the plant. The fungus may, on the other hand, be the controlling member of the pair, not only determining the association but controlling the plant's growth, development, and allocation patterns by producing hormones that override the plant's own hormonal signals.

### The influence of mycorrhizae can depend on plant-plant interactions as well as on soil nutrients

The presence of mycorrhizae can alter interactions between plants. Evelina Facelli and José Facelli (2002) studied how AM mycorrhizae and heterogeneity in soil phosphorus altered plant competitive interactions in experimental populations of *Trifolium subterraneum* (subterranean trefoil, Fabaceae). The effects of mycorrhizae depended on the amount and pattern of phosphorus availablility and the density of competitors (Figure 11.30). Plants with mycorrhizae had a higher average shoot biomass than those without mycorrhizae. Mycorrhizal root colonization was lower for plants grown at high density than at low density; greater intraspecific competition reduced the degree of the mycorrhizal symbiosis. Plant density also affected the benefits of mycorrhizae. Plants grown at low density had a higher biomass if they had mycorrhizae, while plants grown at high density had the same biomass whether or not they had mycorrhizae. Patchiness of the soil phosphorus intensified inequalities in size among seedlings, and mycorrhizae increased this size inequality for seedlings growing at low densities, but not for those at high densities. (No one said ecology is simple.) The researchers speculated that in heterogeneous soils, mycorrhizae may alter plant population structure because differences among plants in the symbiosis could lead to some plants securing a disproportionate share of soil resources.



Mycorrhizae and other interactions between plants and soil organisms may alter competitive interactions and influence plant species coexistence in nature (Francis and Read 1994; Bever 2003). James Bever (2002) found that mycorrhizae altered plant competitive relationships in a complex fashion. Two grassland species, *Panicum sphaerocarpon* (panic grass, Poaceae) and *Plantago lanceolata* (narrowleaf plantain, Plantaginaceae), were grown experimentally with a set of several different AM fungal species. Different AM fungi grew better on one or the other plant species. Most surprising was that *Plantago* benefited much more from the fungal species that grew better on *Panicum*, and the fungus that grew best on *Plantago* benefited *Panicum* more. Bever suggested that this phenomenon might be a mechanism leading to plant species coexistence. As each species became more numerous, its mycorrhizal fungi would increase in the soil, disfavoring its own further increase but helping its competitor. Ecologists have just begun to unravel how mycorrhizae alter competitive interactions among plants; this field is ripe for

research gains and new insights. Mycorrhizae also have important implications for applied ecology. For instance, ecological restoration of degraded ecosystems may depend on reintroduction of the fungi necessary for the establishment of mycorrhizae (Bever et al. 2003) (see Chapter 13), and mycorrhizae may help to ameliorate phosphorus limitations for crops when phosphate fertilizer is unavailable or disadvantageous.

Another phenomenon associated with mycorrhizae is the potential for interconnection between different individual plants via mycorrhizal hyphae. Two plants that are infected with the same mycorrhizal fungus can be connected by a hyphal strand running through the soil. We examine some of the evidence for these interconnections and their effects in Chapter 10. We do not yet know whether this phenomenon is a rare curiosity or a common occurrence, or how important it is in determining plant community composition. We also know little about the effects of such interconnections on competition, or how common or important such effects are in nature. Future research on mycorrhizae is certain to come up with many surprises and to expand our understanding of plant ecology.

## Summary

- Plants are the foundation of nearly all terrestrial food webs, and they are eaten by a wide range of heterotrophic organisms.
- Plants defend themselves against herbivores and pathogens using physical and chemical defenses: Leaves may be low in nutrients or difficult to digest. Toxic phenolics, alkaloids, terpenes, toxic proteins, protease inhibitors, and cyanogenic compounds are important chemical defenses.
- Although herbivory is ubiquitous, the natural world is (mostly) green. Herbivory certainly affects plant population dynamics, community structure, and plant abundance and distribution, but these effects vary greatly among different systems and under different conditions.
- Generalist herbivores affect individual plants, plant populations, and evolution differently than do specialist herbivores.
- Fungi, water molds, bacteria, and viruses can cause diseases in plants. Plants have complex physical and chemical defenses against these pathogens. In addition to their effects on individual plants, pathogens can have complex and far-reaching effects on populations and communities of plants and, indirectly, on people.
- Most plants depend on symbiotic relationships with fungi living on or in their roots, called mycorrhizae, to obtain phosphorus and other minerals from the soil. Most terrestrial plants have mycorrhizae and depend on them to survive, grow, and reproduce in their natural habitats.
- Plants may obtain other benefits from mycorrhizal associations as well, including protection from pathogens.
- Two of the most important types of mycorrhizae are ectomycorrhizae (ECM) and arbuscular mycorrhizae (AM).
- Mycorrhizae are ubiquitous and incredibly important ecologically, but their importance is often overlooked. They are found on most terrestrial plants, and play a critical role in the function of most ecosystems. The relationships between plants and mycorrhizal fungi are complex and variable.

Courtesy of G. Fox

12

# Community Diversity and Structure

On a hike up a mountain in the central Rocky Mountains in the U.S., we might start in a forest dominated by ponderosa pine (*Pinus ponderosa*), then soon find ourselves among lodgepole pine (*Pinus contorta*), then move on to spruce-fir forest (*Picea engelmannii* and *Abies lasiocarpa*), and end our hike high in the alpine tundra. Each area contains a very different collection of species, yet some species are found in many different areas, and it is not obvious whether there are meaningful boundaries between them. In other settings, abrupt changes in some environmental factors lead to large and abrupt differences in species composition between neighboring communities. For example, abrupt boundaries can be found between species-rich communities on unusual, serpentine soils and surrounding communities on more typical soils (serpentine soils are nutrient poor and often high in toxic elements; see Box 4A). Humans also create boundaries that define communities; there are communities in vacant urban lots and along roadsides.

In previous chapters, we looked at the interactions of plant species with their environments, including other species. In this chapter, we examine the properties and characteristics of communities as a whole and some of the factors causing those properties. We first consider community patterns, then the processes that create and maintain those patterns. What determines the boundaries of a community? Are communities real entities with their own properties, or are they just random collections of individuals and populations? Are boundaries between communities abrupt or gradual? How predictable are the patterns? Questions of pattern are critical because they set the stage for the rest of the debate. Once patterns are identified, theories can be constructed to explain them. In this chapter we examine ways of measuring patterns within and among communities. The issues of process focus on determining which mechanisms—such as the physiological tolerances of species, competition,

Above: *Aloe dichotoma* (quiver tree) near Nieuwoudtville, North Cape (Namaqualand) South Africa.

herbivory, evolution, biogeography, historical contingency, and random factors influencing colonization—are most important in determining the observed patterns.

As you will see, over the past century ecologists spent much time and energy considering whether communities are real and also related questions such as how to identify their boundaries. In recent decades there has been a shift in focus to which processes determine which species are found where, and how the traits of those species determine their distribution. While the species compositions of communities can be important in their own right, especially for making conservation decisions (see Chapter 19), ultimately we want to be able to link community patterns to the processes described in many of the other chapters. This last goal is still somewhat elusive, as many different processes can produce similar patterns, and as a single process can result in many different outcomes. It is just this challenge, though, that makes this topic so exciting.

## 12.1 There Are Many Ways of Thinking about Communities

What is a community? A **community** is a group of populations that coexist in space and time and that potentially interact with one another directly or indirectly. By *interact* we mean that the populations affect one another's dynamics. This definition of *community* includes all plants, animals, fungi, bacteria, and other organisms living in an area. It seems simple enough, but ecologists often use terminology inconsistently. For example, we speak of "plant communities" even though plants are only a subset of the entire community; if we focus only on plants, we are ignoring the decomposers, herbivores, pathogens, pollinators, and many other organisms. John Fauth and associates (1996) have discussed some ways through this terminological tangle (Box 12A). Sometimes we use the term *local community* to emphasize that we mean the plants growing in a single place, rather than a group of communities. A related term is **stand**, which comes from forestry and originally referred to a group of trees growing together, although now it is often used in reference to all of the plants in a location, not just the trees.

Two closely related older terms that were formerly in wider use in plant ecology have been incorporated recently into some modern vegetation classification schemes (Table 12.1). An **association** is a particular community type, found in many places and with a certain physiognomy and species composition (e.g., Table 12.1 refers to the *Andropogon gerardii–Heterostipa spartea–Sporobolus heterolepis* grassland association). The term **formation** was originally used to denote a regional climax community, a concept that is no longer generally accepted (see Chapter 14); modern usage is generally more specific, referring to a physiognomic subtype (such as the temperate shrub and herb vegetation formation in Table 12.1). A similar term is **community type**, which refers to a set of communities that have the same dominant species.

In practice, the boundaries of plant communities are usually defined by the identity and abundance of the dominant, or most common, species. Sampling is then confined within those boundaries. Typically, a number of local communities or stands are used to sample the presence and abundances of species as well

**TABLE 12.1** An example of the classification of a North American plant community

| | |
|---|---|
| **Upper levels** | |
| Formation class | Mesomorphic shrub and herb vegetation |
| Formation subclass | Temperate and boreal shrub and herb vegetation |
| Formation | Temperate shrub and herb vegetation |
| **Mid levels** | |
| Division | *Andropogon-Stipa-Bouteloua* grassland and shrubland |
| Macrogroup | *Andropogon gerardii-Schizachyrium-Sorghastrum nutans* grassland and shrubland |
| Group | *Andropogon gerardii-Heterostipa spartea-Muhlenbergia richardsonis* grassland |
| **Lower levels** | |
| Alliance | *Andropogon gerardii-Sporobolus heterolepis* grassland |
| Association | *Andropogon gerardii-Heterostipa spartea-Sporobolus heterolepis* grassland |

*Source*: After D. Faber-Langendoen et al. 2014. *Ecol Monogr* 84: 533–561.

*Note*: This classification follows the U.S. National Vegetation Classification system developed by U.S. federal agencies, the Ecological Society of America, and NatureServe. The classification uses a three-part system based on physiognomy and floristics in which the upper level is related to global-scale vegetation patterns, climate, hydrology, and substrate; the middle level is related to regional floristic-physiognomic types; and the lower levels are based on floristic composition.

## BOX 12A Communities, Taxa, Guilds, and Functional Groups

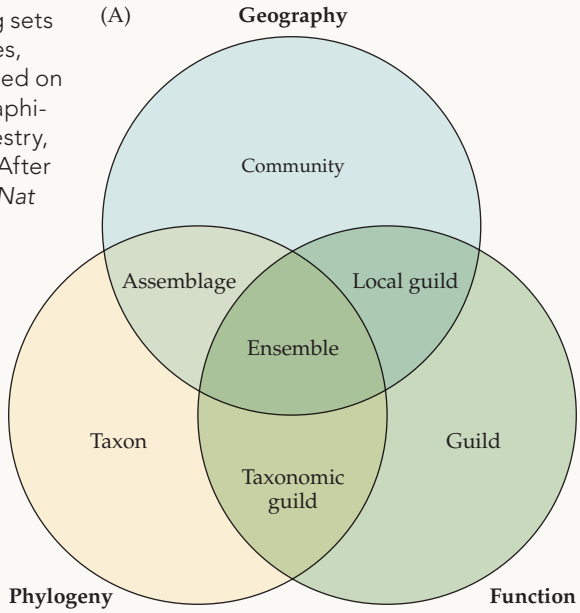
A very confusing array of terms has developed to describe communities. Sometimes the same term is used in different ways by different ecologists, while in other cases different terms are applied to the same concept. Here we describe a recently proposed scheme to define these terms.

This scheme defines groups of species based on three criteria: geography, phylogeny, and function, as shown in Figure A. Geography in this scheme defines **communities**, sets of organisms living in the same place at the same time. **Phylogeny** is the pattern of relationships among species (or higher taxa) based on evolutionary ancestry. Phylogeny defines **taxa**, sets of organisms that share a common ancestor. Function defines **guilds**, sets of organisms that use biotic or abiotic resources in a similar way. The term *guild* is taken from animal ecology but has been used by plant ecologists as well.

Plant ecologists often use the term *functional group* to describe a concept related to the guild. Functional groups can be defined in a variety of ways, depending on the application, but these definitions are all based on a set of traits that identify functionally similar species. For example, Tammy Foster and J. Renée Brooks (2005) defined five functional groups of plant species in Florida scrub habitat based on physiological traits. The traits used to identify functional groups can be chosen informally or based on formal mathematical algorithms. Ecologists have used the concept of functional groups in a variety of contexts, including studies of the relationship between productivity and diversity (see Chapter 13) and attempts to reduce the number of types of plants that must be taken into account in global climate modeling. Extensive recent reviews of the concept of plant functional groups and ecological applications of this concept are provided by Lavorel and Cramer (1999) and Woodward and Cramer (1996).

Intersections of the sets described by the terms above define narrower groups of species. The intersection of geography and phylogeny defines **assemblages**, groups of related organisms living in the same place.

(A) A scheme for grouping sets of species into communities, guilds, and ensembles based on the combination of geographical location, common ancestry, and shared resource use. (After J. E. Fauth et al. 1996. *Am Nat* 147: 282–286.)




The intersection of geography and resource use defines **local guilds**. The trees in a forest in Ontario are an example of a local guild: they include distantly related species such as sugar maple (*Acer saccharum*), an angiosperm, and eastern hemlock (*Tsuga canadensis*), a conifer. The intersection of phylogeny and function defines **taxonomic guilds**, groups of related organisms that use the same class of resources in a similar way. The genus of oak trees (*Quercus* spp.) is an example. The intersection of all three sets defines **ensembles**. The grass species living together in a prairie are an ensemble. The annual Asteraceae in Australia's Great Sandy Desert are another ensemble.

Plant communities as traditionally defined are made up of a combination of ensembles, local guilds, and assemblages. Typically, terrestrial plant communities are defined as all the vascular plants living in a given space. Most species in this group are primary producers with similar resource requirements. So, for example, all of the grass species in a forest understory are one ensemble of that community. The combination of all understory forbs and graminoids would constitute a local guild, as these species would include more distantly related species of monocots, dicots, and possibly ferns. Some flowering plants are not primary producers, but parasites or saprophytes (Figure B); however, these species are also included in the plant community. The plant community could therefore be considered an assemblage, as it includes species that use different resources. The true community, of course, includes all species (e.g., animals, fungi, bacteria), not just plants. Thus, the traditionally defined plant community is actually a subset of the full community and has properties of ensembles, local guilds, and assemblages.

(B)

© Dan Hanscom/Shutterstock.com

(B) *Monotropa uniflora* (Indian pipe, Ericaceae) is an example of a flowering plant that is a parasite rather than a primary producer. It obtains its carbon from mycorrhizal fungi, which in turn obtain it from other plants.

as associated environmental variables. Based on data from a number of stands, a community type, formation, or association can be characterized.

Only in special cases (e.g., small islands, ponds, forest preserves surrounded by suburban development, vacant lots) are the boundaries of communities defined easily. Even then, the movement of organisms and the transport of matter by wind and water may make their boundaries indistinct. Ecologists, therefore, are often of two minds when dealing with communities. On the one hand, we recognize that their boundaries are often fuzzy; on the other hand, we often need to define discrete entities for convenience of analysis. Typically, plant ecologists define a community based on the relative uniformity of the vegetation and use their knowledge of species biology to decide when they are moving from one community to another.

Ecologists based in different countries and educated in different historical traditions tend to view communities in somewhat different ways. In particular, ecologists in continental Europe were historically influenced by the floristic-sociological approach, most extensively developed by the Swiss ecologist Josias Braun-Blanquet and his students and colleagues. This approach emphasizes the discrete, distinctive nature of communities. In contrast, ecologists in English-speaking countries have been more strongly influenced by the history described in the following paragraphs; as a result, they tend to think of communities as blending continuously into one another. These distinct ways of thinking are becoming less prominent as a result of increased travel and communication among ecologists worldwide.

### The debate between Henry Gleason and Frederic Clements shaped modern ideas about plant communities

The passionate debate about the nature of communities that occurred over 100 years ago is a great example of how real science makes advances in understanding; the process often may zigzag more than moving full steam straight ahead, personalities play a role, and ultimately, evidence is what decides things. The debate also helps clarify what an ecological community is, and provides the context for our current views. Despite the fact that this debate was settled by ecologists many decades ago, it echoes popular misconceptions about the natural world. Many nonscientists believe that all living things are tightly connected to all other living things and that they all work together for the success of the whole community or the whole living world, the ideas of the ecologist Frederic Clements that a community is a "superorganism."

Within ecology, there is a range of views on the nature of communities. The extremes are sometimes labeled the Clementsian and Gleasonian views, named after Frederic Clements and Henry A. Gleason, their first major proponents in the English-speaking world. The debate between them helped shape modern ideas about plant communities. These two extreme viewpoints differ in the importance they ascribe to biotic versus abiotic factors, and to predictable versus random processes in shaping community structure. Today, while most ecologists borrow some ideas from each of these viewpoints, to a large extent we have moved beyond both of them to newer ways of thinking about communities.

Courtesy of The American Heritage Center, University of Wyoming

Frederic and Edith Clements

The Clementsian view was the majority view among English-speaking plant ecologists during the first half of the twentieth century. Clements saw plant communities as highly organized entities made up of mutually interdependent species. In his view (Clements 1916), communities are **superorganisms**—the analogue of individual organisms—that are born, develop, grow, and eventually die. Clements believed that the different species in a community functioned together, much as the different tissues and organs of an individual organism function together in a real organism. Succession (see Chapter 13), in this view, is analogous to the processes of an organism developing and growing; the trajectory and end point of succession, like the development and growth of an animal, are highly predictable (Clements 1937). Two of the hallmarks of the superorganism concept are the presence of very tight linkages among species within communities and cooperation among the species in a community for the benefit of the function of the entire community.

Even at the height of Clements's influence, many ecologists held more moderate and nuanced views that did not really accept communities as being discrete superorganisms but still found it useful to treat them that way. This modified version of Clementsian ecology acknowledged that species interactions such as competition, mutualism, and predation may be important in determining community structure. Community structure is the relative composition of a community; for example, 83% of trees in a temperate deciduous forest in the northeastern U.S. might be white oaks (*Quercus alba*), 11% red maples (*Acer rubrum*), and the remainder composed of 20 other species. Even Clements himself acknowledged the effects of abiotic factors such as site history and soils in determining community composition. He focused on the idealized nature of communities and saw them as spatially distinct, with one superorganism complex

giving way to another with a very different collection of species. For example, he viewed certain community types as departing from what he saw as the normal state only because of some unusual feature such as frequent fires, or a soil deficient in some nutrient. His major focus was on the nature and development of the community as a superorganism, however, not on the boundaries between communities.

In striking contrast to Clements, Gleason held that communities are the result of interactions between individual species and their environment (biotic and abiotic factors) in combination with chance historical events. Each species has its own environmental tolerances and thus responds in its own way to environmental conditions (Gleason 1917, 1926). The implication of this view was that along an environmental gradient, different species would have their boundaries at different places. Not only were communities not tightly linked superorganisms, but defining the collection of species living together in a particular place as a cohesive entity was an arbitrary human construct.

The Lutesther T. Mertz Library of the New York Botanical Garden, Bronx

Henry A. Gleason

According to Gleason, chance events determine whether a species is actually found, as long as it is within the range of the environmental conditions a species can tolerate in a given location. At the local scale, chance dictates whether a seed happens to get to a particular spot. At larger scales, the chance events of history play a major role in community composition. For example, species in the family Cactaceae (the cacti) are found in many arid communities of the Americas because the family originated in this region, while arid regions elsewhere have no cacti (except where they have been introduced by humans). On a more local scale, one can find locations in arid regions of the Americas that have no cacti because the seeds did not get there. Furthermore, the mix of species changes from place to place as one moves across the landscape. According to this viewpoint, community composition changes gradually, rather than abruptly at boundaries between communities (unless there are abrupt environmental boundaries). In between the views of Clements and Gleason, some ecologists believed that identifiable community types exist but that these tend to blend into other community types.

Most ecologists (at least in the U.S.) were not receptive to Gleason's ideas. While Gleason's work was well known, it failed to influence many plant ecologists until after Clements's death in 1945. Why Clements's views held sway for so long is not fully understood. Certainly the fact that Clements put forward the first comprehensive theory in plant ecology (Clements 1916) was a reason for its initial great attraction. He wrote prolifically; in the 1920s he published seven books, including the first plant ecology textbook written initially in English. (Eugenius Warming, a Danish botanist, wrote the first plant ecology textbook—*Plantesamfund*—in Danish in 1895; it was translated into German, Russian, and then English in 1909.) Clements was said to have an extremely strong personality that could dominate scientific meetings. The very small number of plant ecologists active during the first half of the century may have been sufficient reason. Gleason, finding little interest in his ideas, abandoned his work in plant ecology by 1930 and spent the rest of his career as a taxonomist.

Not until the 1950s did the separate but almost simultaneous work of John T. Curtis and Robert H. Whittaker convince many ecologists that Gleason's views were largely correct. Curtis and his students mapped the vegetation of Wisconsin and looked at how species' optima and ranges were distributed along environmental gradients (Curtis 1959). Clements's theory predicted that many species would have the same environmental optimum and range, thus forming a distinct community. By contrast, Gleason's theory predicted that each species should have its own optimum and range. Curtis's study found just such independence; each species had a different set of environmental tolerances (Figure 12.1) and, consequently, a different range along environmental gradients. A key innovation that contributed to this study was the development of **ordination**, a set of techniques for describing patterns in complex vegetation; we discuss ordination in detail later in this chapter.

Courtesy of Thomas Wentworth

Robert H. Whittaker

Whittaker (1956) also demonstrated that Gleason was right about the nature of boundaries between communities. Going up a mountain, one of the most striking patterns one encounters is the dramatic turnover from one community type to another as elevation increases. Whittaker realized that if he could demonstrate that species **turnover** was gradual even along such a gradient in elevation, this would provide very powerful evidence in support of Gleason's ideas and in contradiction of Clements's superorganism model. Whittaker did just that. He showed that forest communities along an elevational gradient in the Great Smoky Mountains of Tennessee changed gradually in species composition without abrupt boundaries, with each species having an

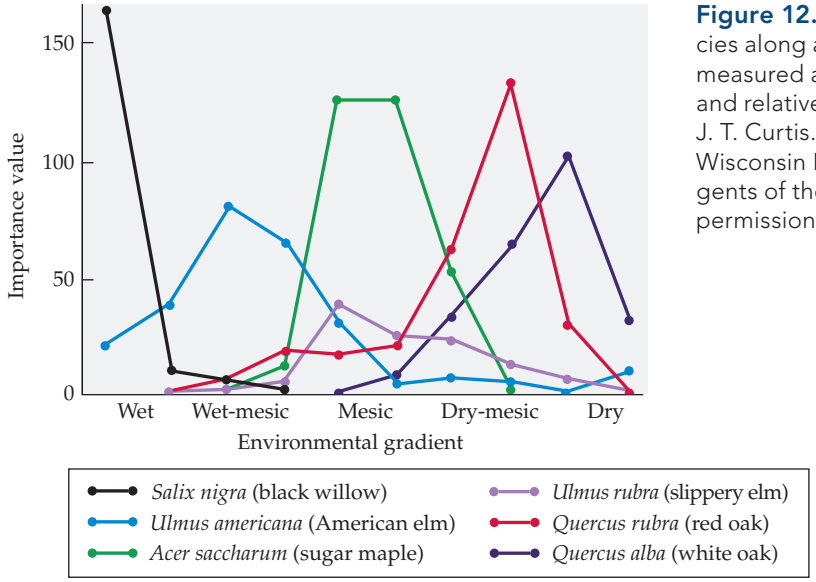



**Figure 12.1** Change in the importance of various tree species along a moisture gradient in Wisconsin. Importance was measured as the sum of the relative cover, relative density, and relative frequency of a species in a community. (After J. T. Curtis. 1959. *The Vegetation of Wisconsin*. University of Wisconsin Press: Madison, WI; © 1959 by the Board of Regents of the University of Wisconsin System. Reprinted by permission of The University of Wisconsin Press.)

independent distribution (Figure 12.2). He then repeated this work in other areas, including the Siskiyou Mountains of Oregon and California, and the Santa Catalina Mountains in southern Arizona.

Another important line of evidence that strongly affected many ecologists' views of communities was a series of studies begun in the 1970s on the distributions of plant species during and after the most recent glaciation (see Chapter 17). Many of these studies looked at fossil pollen that had accumulated over time in the sediments on lake bottoms. At any one point in time, the pollen of the most common species was deposited and fossilized together, revealing which species were found together in that community. Some of the earliest and most influential of these studies were carried out by Margaret Davis. She showed that many species that co-occur today were not necessarily found together during the most recent glacial period; rather, species were distributed among communities in the past in very different combinations (Davis 1981; see Figure 17.14).

Today, most plant ecologists take a position that incorporates some of Clements's and Gleason's ideas but in many ways has diverged from both. There is wide agreement that species are distributed individualistically and that community composition typically changes gradually along environmental gradients. Abrupt changes are most likely to be found where there are abrupt changes in the environment. However, abiotic boundaries and community boundaries do not always match. Because of processes such as dispersal from one habitat into another (see Chapter 15), populations of some of the species may extend partway into an unfavorable environment. Abrupt changes may also reflect past events, such as the edge of a fire or a part of a forest that was plowed at some time in the past. Thus, current environmental boundaries do not always match past boundaries. Ecologists still disagree about the relative importance of biotic and abiotic processes and chance events in determining community structure, but there is little if any support for Clementsian views.

Echoes of the controversy between the Clementsian and Gleasonian views of community among English-speaking ecologists of North America and the United Kingdom continue to influence ecologists schooled in those traditions. Related issues were debated among European and Russian ecologists, but not to the extent that they were argued in American and British journals; the ecologists of continental Europe and Russia had different concerns. Heavily influenced by Braun-Blanquet, they focused primarily on systems of community classification. The key difference is that the European tradition was more concerned with describing patterns than with analyzing processes, mainly sidestepping these arguments. We will return to their traditions later in this chapter. Somewhat related issues were debated among animal ecologists, although it was not until the last quarter of the twentieth century that animal and plant ecologists began to talk to one another regularly about conceptual issues.

## Today's ecologists have a different perspective on the issues in contention

The primary issues surrounding the nature of plant communities divide roughly into those of pattern and those of process (Stroud et al. 2015). Underlying these issues is theory: the effort to explain the patterns, which includes a search for the processes responsible (see Chapter 1). Our understanding of community patterns and processes is tied together by theories regarding the nature of communities—such as Clements's superorganism theory—which seek to explain the patterns and processes that we find in nature.

In Clements's original superorganism theory, the community was an organic entity: a distinct unit with strong emergent properties. **Emergent properties** are those of a complex whole (e.g., a community) that are more than just the sum of the properties of its components (e.g., the species that make up a community). Emergent

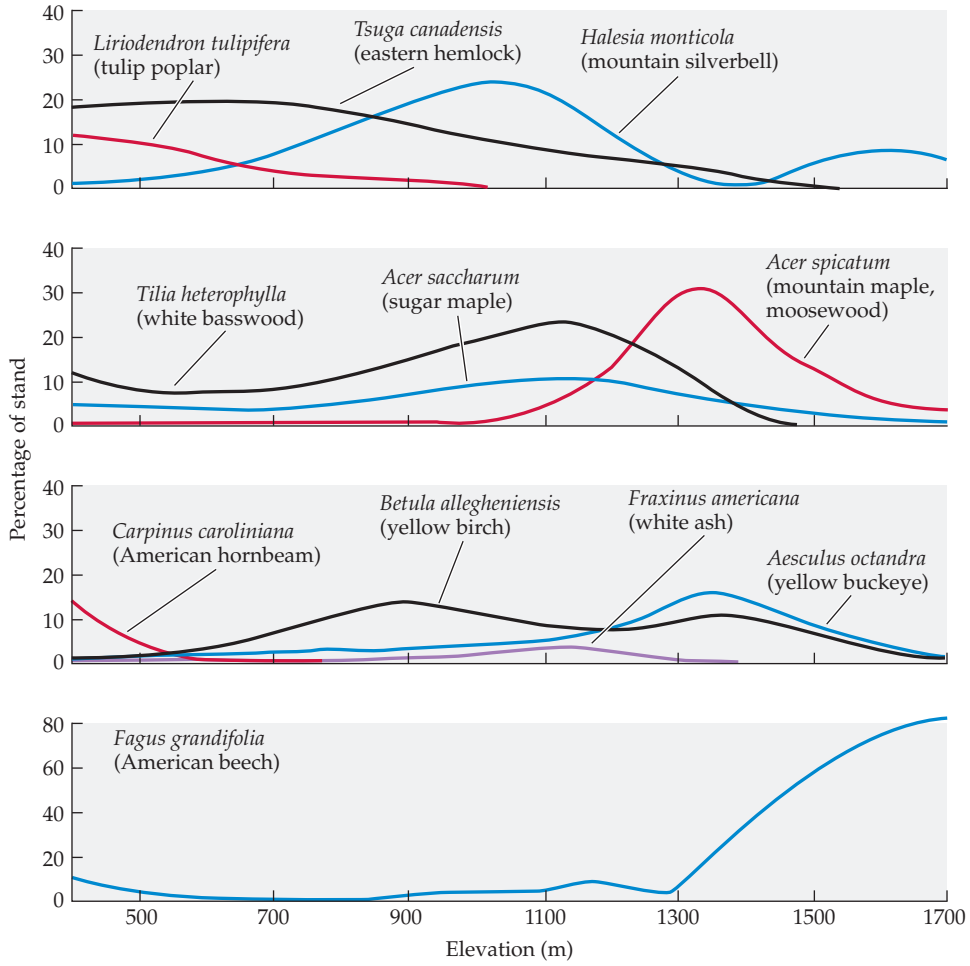



**Figure 12.2** Changes in plant species frequencies along an altitudinal gradient in the Great Smoky Mountains of Tennessee. (After R. H. Whittaker. 1956. *Ecol Monogr* 26: 1–80.)

properties of communities come about through interactions, such as competition, predation, and mutualism that occur among the species in a community (Box 12B). Species within the community in Clements's view were tightly linked and interdependent. In contrast, in Gleason's view, any community-level properties were simply the sum of the properties of individual species. Determining the truth of the matter requires documenting patterns (as Curtis, Whittaker, Davis, and others did), understanding the processes responsible for creating those patterns, and posing plausible explanations for those patterns and processes. In truth, neither view is correct. Species within communities interact, but only loosely and partially. However, those interactions still create emergent properties of the community.

Whether communities can be considered to have emergent properties depends in part on the relative contributions of biotic and abiotic processes in shaping community structure, including how strongly species interact with one another within communities. (We view this issue in somewhat different terms today than either Clements or Gleason did, as we will see in Chapter 14.) If these processes play a major role in shaping communities, then communities are likely to have emergent properties. But if communities are mainly structured by the tolerances of individual species for abiotic factors (such as minimum temperatures) in the environment, then community properties are largely aggregates of the individual species' properties. Emergent properties can be characteristic of communities even if Clements's superorganism view is incorrect. Emergent properties do not mean that communities consist of species adapted for one another's benefit and that all or most species in a community are tightly interlinked. Later ecological thinkers, for example, Eugene P. Odum, Howard T. Odum, Robert V. O'Neill, and their coworkers, have strongly emphasized the idea that communities and ecosystems have emergent properties, but none of them embraces the superorganism view (Odum 1971; Odum 1983; O'Neill et al. 1986; Odum 1988).

Beginning in the 1970s, but especially in the past two decades, greater attention has been paid to the role of evolution in structuring plant communities. This attention further eroded any notion of a community as

## BOX 12B A Deeper Look at Some Definitions: Abiotic Factors and Emergent Properties

Ecologists have distinguished between the effects of abiotic (nonliving) and biotic (living) factors in the environment for a very long time. Typical biotic factors include competition and predation; typical abiotic factors include soil nutrients, microclimate, weather, and general climatic influences. The problem with this terminology is that the distinction between abiotic and biotic factors may be far less clear than we might think. As we emphasize in Chapter 4 and Chapter 5, soils are a product of organisms and their interactions with their environment. The nitrogen available to the roots of a plant, for instance, depends on the actions of many different kinds of soil organisms and the interactions of the plant with those organisms. Likewise, the microclimate—and even the global climate (see Chapter 16) is affected by living things. We do not have a good substitute to suggest for the term *abiotic*, so while we use it here to mean "things like climate and soils," we recognize that these things may have major biotic components. On short time scales—say, years to decades—it is reasonable to treat climate and soils as things that are effectively abiotic, but on longer time scales they clearly have biotic components.

Another term that bears closer examination is *emergent properties*. A central issue in the argument over the nature of communities is the question of whether emergent properties exist. An emergent property is one that is found at a certain level of organization as a result of properties, structures, and processes that are unique to that level of organization. Emergent properties can be contrasted with properties that are merely aggregates of properties at a lower hierarchical level. The properties of a water molecule, for example, are not simply the properties of hydrogen and oxygen atoms taken together; the molecule has emergent properties that result from the way the hydrogen and oxygen atoms interact. As an ecological example, consider canopy photosynthesis. We cannot measure the photosynthetic rate of an entire forest canopy just by measuring the photosynthetic rates of the individual plants under average conditions. Canopy photosynthetic rates depend on how individuals interact in several ways, including shading one another, interfering with wind, and competing to taking up $CO_2$. The canopy photosynthetic rate is an emergent property of the community. In practice, this means that the information an ecologist gains from using a single-leaf or single-plant measurement of photosynthesis (such as those made with a modern IRGA-based system; see Box 3A) is informative about physiological functioning and capabilities only at that scale. If the ecologist wants an estimate of canopy photosynthesis, she must use a method such as eddy covariance (Chapter 5), rather than trying to scale up from individual estimates.

a superorganism—a single, adapted entity—while also highlighting that pairs or sets of species within a community might exist as parts of coevolved mutualisms or other interactions. In particular, recent advances in genomic technologies allow for explorations of coevolution between plants and associated soil microbes (see Chapter 4).

This shift from the study of emergent properties to the study of processes is emblematic of the seesawing between reductionist (mechanistic) and holistic (emergent) approaches that has characterized ecology for the past century. Ecologists in the 1980s and 1990s tended toward reductionist approaches. Since then there has been some movement back in the direction of holistic studies under the mantle of macroecology (Brown 1995, 1999; Lawton 1999). Interestingly, much of the modern effort in mapping species' distributions at large spatial scales in the field of macroecology currently and in predicting responses to climate change (see Chapter 16 and Chapter 19) takes a rather strict Gleasonian approach, where aspects of current or predicted climate (that is, abiotic factors modeled individually for each species) are the basis for projecting where species are likely to be found currently or in the future.

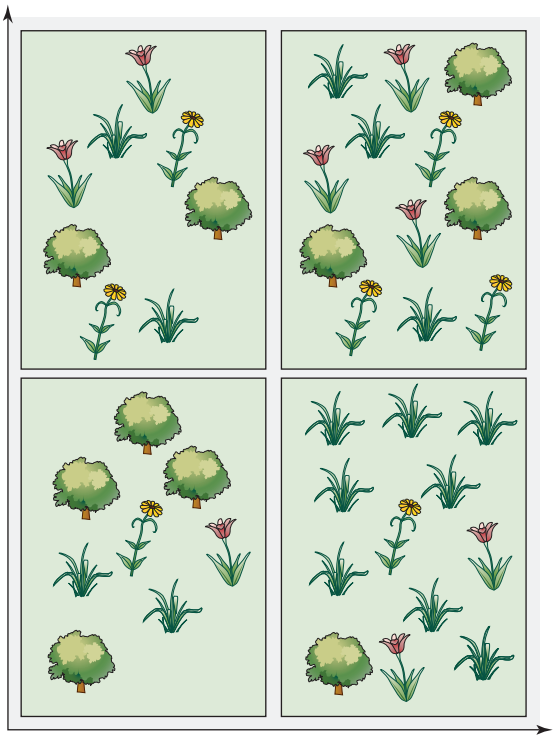
### The concept of communities is useful but has often been debated

As we have seen, the extent to which communities are "real" has been a contentious issue among plant ecologists for much of the past century. The heart of the debate has concerned what types of entities are real, and what types are just mental constructs. It is clear that populations and species are real entities, but are communities real entities as well, or are they merely convenient but arbitrary human inventions? Many additional studies have reinforced the conclusion of Curtis and Whittaker that, except where there are abrupt physical discontinuities, plant communities tend not to have discrete boundaries. You might start to think that this means that there are no community-level processes worth studying.

This conclusion would be wrong; in fact, that debate is miscast. Instead of focusing on patterns, let's shift the focus to processes and rephrase the question by asking whether community-level processes are important in structuring the living world. We have already discussed several processes responsible for interactions among species: competition (see Chapter 10), and herbivory and mutualisms (see Chapter 11). These are all

community-level processes that occur among the component parts of communities (e.g., populations of species). If such processes are important in structuring a particular system, then we can regard that system as a community with unique properties. Thus instead of asking whether we observe patterns that would be expected if communities are real entities, we focus on processes that can cause communities to be real entities. Such an outlook eliminates the need to worry as much about the existence of clear boundaries between communities. We can recognize the existence of communities for those questions and purposes for which it is useful to do so, and ignore them when it is not.

The debate about whether communities are real entities or imaginary constructs is more than just a philosophical or academic exercise. The U.S. National Vegetation Classification (USNVC) system is a unified framework for documenting and studying vegetation in the United States. Related efforts in other parts of the world include the European Vegetation Survey and the National Vegetation Information System of Australia. For example, the private organization The Nature Conservancy in the United States makes decisions about land acquisition and management based in part on the USNVC classification of the communities present (Table 12.1, Faber-Langendoen et al. 2014). Classifications of communities are incorporated into law. There are many legal restrictions on development and land use in communities categorized as "wetlands" in the U.S. In southern California, land development is regulated quite differently for "coastal sage scrub" than for "chaparral" communities, although the two include many of the same species, and scientists differ in defining a particular place as one or the other.



## 12.2 Biodiversity Describes Variation in Biological Organisms and Systems

**Biodiversity** is a concept fundamental to the science of ecology, and much of ecology is ultimately about explaining patterns of biodiversity. At its broadest scope, biodiversity refers to variability in biological systems and organisms at all levels, from molecules within cells to biomes across the globe. Most often in ecology we focus on the middle of this range: individuals within species, species within communities, and communities within landscapes. Because biodiversity is such a fundamental concept, unsurprisingly, it has engendered much debate among ecologists about what it is and how to measure it.

To begin to see why biodiversity involves a number of different issues, consider the communities in Figure 12.3. The four communities are identical in one respect—they all contain the same four species. But they are quite different in their relative compositions: communities C and D are each dominated by a single species, while in communities A and B, the numbers of individuals of each species are the same. A more complex figure might show other communities with more species. Each of these quantities—the number of species, the total number of individuals of all species, and the relative abundances of the different species—contributes to species diversity. The species in different communities may also be closely or distantly related to each other (phylogenetic relatedness). For example, one community might have species that all belong to a single family or closely related families, while another might have species from many different plant families that are distantly related to one another (phylogenetic diversity). The species in the community may have high or low trait diversity. For example, a community with low trait diversity may have only trees. Alternatively, a community may have grasses, herbs, trees, and shrubs; wind-pollinated and also insect-pollinated flowers; or both $C_3$ and $C_4$ grasses

active at different times of the year (functional or trait diversity). Measuring all of these aspects of biodiversity is important for understanding the processes responsible for forming and maintaining the community, which in turn can inform decisions on how to manage that community or conserve its biodiversity.

The concept of biodiversity encompasses many different ideas and metrics. Many of our current concepts and measures of biodiversity trace to the work of Robert H. Whittaker (1960). Unfortunately, in the process of codifying concepts of biodiversity, he used the same words for multiple metrics, thus sowing confusion for the next half century. Work over the past decade has succeeded in uniting and simplifying these concepts and taming that confusion, while also considering factors not addressed by Whittaker, such as the phylogenetic relatedness of species and the similarity of their traits. We start by dividing biodiversity based on the property being measured (species identity, individual abundance, phylogenetic relatedness, or ecological function) and on how variation in those properties is distributed within and among units—what has been called inventory diversity and differentiation diversity.

## Biodiversity metrics can be built from different types of information

Biodiversity metrics are built from four basic types of information—identity, abundance, phylogeny, and trait function—each with particular properties that can be quantified in more than one fashion and combined in a variety of ways to produce many different kinds of metrics. Consider the species shown in Figure 12.3. Each species has a name: its identity. Each has an abundance: two of each species in community A. Each is embedded in a phylogeny: the grass and the pink-flowered plant are both monocots and therefore are more closely related to one another than to the two dicots, the tree and the yellow-flowered aster. Each species has different functional traits: one is a tree and three are herbaceous, two have conspicuous flowers and are insect-pollinated, and two (the grass and the tree) are wind-pollinated. These different types of information about each individual and species are different aspects of the biodiversity of each of the communities.

The identity, abundance, phylogeny, and trait function of the members of a community depend on different kinds of information (Table 12.2). Species identities depend only on themselves and not on other species in the community. The same is true for the abundance of a species: it depends only on the number of individuals, or the biomass, of that species. So, the number of species and the abundance of those species just depend on counting up the species and their abundances in the community individually. (Besides numbers of individuals, abundance can be measured in a variety of other ways, such as frequency of occurrence, biomass, or geographic range.) However, the phylogenetic diversity and functional trait diversity both depend on the other species present. Phylogenetic diversity depends on how different the species are evolutionarily in the community, and functional trait diversity depends on how different the trait values are between the different species.

In the next section we look at how identity and abundance information can be used as metrics of biodiversity within a single community, and then we consider how biodiversity can differ between communities. After that, we look at within- and among-community diversity based on phylogenetic and functional trait information.

## Inventory diversity is the variation of types of objects

We begin with measures based on identity and abundance because historically that is how these ideas were developed and have been commonly used in plant ecology. Until the 2000s, most evaluations of biodiversity focused on species occurrences and abundances. The past decade, in particular, has seen extensive interest in and development of metrics of phylogenetic and functional diversity, and this continues to be a very active area of research (Magurran and McGill 2011; Tucker et al. 2017). The idea of inventory diversity unifies many separate ideas about how to think about and quantify diversity. **Inventory diversity** is the diversity of the individual things in a unit; for example, it might be the diversity of genotypes (the individual things) within a population of a species (the unit), or the number of species (the individual things) within a forest preserve (a unit). We might also be interested in the inventory diversity of all of a group of units, for example, all of the forest preserves in a region.

One way to describe the inventory diversity of a community is by a list of the species in it. **Species richness** ($R$) is the number of species on such a list (information about species identities) and represents the simplest and most intuitive measure of inventory diversity. For example, the four communities in Figure 12.3 all contain four species and so have the same species richness.

**TABLE 12.2** The properties of the three types of biodiversity information

| Information type | Magnitude property | Variability property |
|---|---|---|
| Abundance | Numbers | Evenness |
| Phylogeny | Divergence | Regularity |
| Function | Dispersion | Equability |

*Source:* After S. M. Scheiner et al. 2017. *Ecol Evol* 7: 6444–6454/CC BY 4.0.

How would one gather such a list? A simple and widely used method is to establish the boundaries of the community and then walk through it, identifying all of the plants and listing them. Such a survey should be done several times during the year because some species may be visible only during a single season. Spring ephemerals, for example, are perennial plants common in many temperate deciduous forests whose leaves, stems, and flowers are present for only 1 to 2 months in the spring; during the rest of the year, they exist only underground as dormant bulbs, corms, or rhizomes.

Species identity is only one aspect of biodiversity. Not all species exist in equal numbers: some are rare, some are common, and others are very numerous; this is information about abundance. Other common measures of abundance are biomass and cover, which we discuss later in this chapter. Abundance information has two properties, its magnitude and its variability. In Figure 12.3, both communities A and C have 8 individuals, and communities B and D both have 12 individuals. The latter have greater numbers, or abundance magnitude. While within each pair of communities (A and C; B and D) the abundance magnitudes are the same, communities A and B differ in another property from communities C and D. Community A has two individuals of each species, and community C has four individues of the tree species, two individuals of the grass species, and one of each of the species with flowers. These two communities differ in a property called **evenness** (**E**), or abundance variability. The community in which all the species have the same number of individuals is more even, and thus it has one of the essential elements of being more diverse than the second. The species diversity of a community depends on both its richness and its evenness: higher species numbers, with the individuals more evenly distributed among them, contribute to higher community diversity. A way to think about evenness as a contributor to diversity is to consider the following thought experiment: Pick two plants at random from a community. Are they members of the same species or different species? If all species have about the same abundances, it is more likely that they will belong to different species, and such a community is highly diverse.

In the example in Figure 12.3, the species with the greatest number of individuals is called the **dominant species** in that community. A community with maximum evenness (all species being equally abundant) does not have a dominant species. The greater the numerical preponderance of one or a few species, the lower the diversity of the community tends to be. Of course, a community may have one dominant species (with most of the individuals) and a large total number of species (each with a small number of individuals). Such a community would have a high diversity value, but this situation is actually fairly unusual. Other ways of characterizing variation among species in abundance are considered in Chapter 14. It is important to recognize that *dominance* does not mean quite the same thing in this context as in popular speech: the dominant species is merely the most abundant. It may not be the most competitive, and indeed the word *dominant* does not tell us anything about how individuals of this species interact with other species. Nor is it related to the idea of "dominance" in genetics (see Box 9A).

Species richness and relative abundances are aspects of inventory diversity. Often we wish to combine identity and abundance information into a single metric that accounts for both the list of species and their relative numbers, using an abundance-weighted measure of species richness (Box 12C). Such a measure is convenient for comparisons among communities, comparing the same community over time, or making comparisons among measures of different properties, such as phylogenetic and functional diversity (discussed below).

Many different indices of diversity have been proposed. Two of the most common are the Shannon-Wiener index and Simpson's index, both of which are special cases of the Hill metric (Box 12C). Each index has limitations, and each provides ecologists with somewhat different information. One way of thinking of the Shannon-Wiener index is that it quantifies how likely it is that you can guess the species of an individual chosen from the sample at random. If all of the species are equally represented, the chance is $1/R$, but as a single species becomes the most abundant, it becomes more and more likely that you can guess the individual's species (because it probably belongs to the dominant species). Simpson's index is a measure of the chance that two individuals chosen at random from the same community belong to the same species; as the abundance of a single species increases, that probability increases toward 1.

Which of these indices is more useful—species richness, the exponent of the Shannon-Wiener index, or the inverse of Simpson's index? Another way of asking this question is, Which value of q should be used in calculating inventory diversity? That depends on the question and on the data available. If one had hypothesized, for example, that some environmental factors were having a negative effect on community diversity over time by disproportionately affecting rare species, reducing their abundance, or eliminating them, then ${}^{1}D(A)$ might pick up such effects more readily. In contrast, if the hypothesis was that common species were more affected, then ${}^{2}D(A)$ might be more appropriate.

Now that we have our measure of diversity, we are ready to partition it into the two components described above, richness and evenness: ${}^{q}D(A) = R \times {}^{q}E(A)$. That is, diversity is the product of species richness ($R$) and evenness (E), which is the extent to which the species within

## BOX 12C A Unified Measure of Diversity

Over 45 years ago, Mark Hill (1973) proposed a measure of diversity that combined information on species identity and abundance that could unify commonly used measures into a single equation. However, it was ignored for 30 years, until Lou Jost (2006) showed that it gave us an intuitive measure of diversity, numbers equivalents, which is the effective number of species in a unit if all had the same mean proportional abundance. For example, in Figure 12.3 community B has four species, and each has exactly the same proportional abundance (3 of the 12 individuals); so for that community the effective number of species is four. In contrast, in community D more than half of the individuals are members of the grass species, and so the effective number of species is less than four. Hill's metric is

$$^{q}D(A) = \left(\sum_{i=1}^{s} p_i^q\right)^{1/(1-q)}$$

where qD(A) is the effective number of species, A is abundance, $S$ is the number of species, $pi$ is the proportional abundance of the $i$th species ($ni/N$), and q is a parameter that determines how species frequencies are weighted. The effect of changing q is shown in Table 12.3. qD(A) has the property that its value is close to 1 when a single species is the most abundant and that it equals species richness ($R$) when all species are equally abundant.

When q = 0, all species are given equal weight irrespective of their abundance so $^{0}D(A)$ is species richness. Increasing the value of q gives progressively more weight to the most abundant species. When q = 1, each species is weighted exactly by its proportional abundance so $^{1}D(A)$ is approximately equal to the number of common species. The value $\log^{1}D(A)$ is typically referred to as the Shannon-Wiener index. The quantity $^{2}D(A)$ is approximately equal to the number of highly abundant species and $1/^{2}D(A)$ is known as Simpson's index. Nor is an ecologist confined to using a single value of q. Comparisons of changes in $^{q}D(A)$ for different values of q can provide a broader picture of the effect of variation in abundances on diversity. Actually, neither the Shannon-Wiener index nor Simpson's index is a measure of diversity, as they are not species equivalents; but their Hill diversity versions—the exponent of the Shannon-Wiener index and the inverse of Simpson's index—are species equivalents.

Besides unifying different measures into a single equation, the same equation can be used for phylogenetic and functional diversity. To measure the effective number of phylogenetically distinct species, we need to partition the total branch lengths among the amounts accounted for by each species ($L_i$, see Figure 12.5). Each species accounts for a proportion of the total divergence, $l_i = L_i/L$, and the Hill metric is

$$^{q}D(P) = \left(\sum_{i=1}^{s} l_i^q\right)^{1/(1-q)}$$

This metric measures the number of phylogenetically distinct species and, as it is of the same form as the metrics for species diversity used above, has all of the same properties as $^{q}D(A)$. For example, $^{0}D(P)$ equals species richness ($R$). If all species are equally divergent, then $^{q}D(P)$ equals species richness for all values of q. Conversely, $^{q}D(P)$ converges toward 1 if the cladogram consists of several very closely related species and one very divergent species. We can do the same with trait information. For each species, we calculate its mean proportional dispersion from the other species, $d_i$, and the Hill metric is

$$^{q}D(T) = \left(\sum_{i=1}^{s} d_i^q\right)^{1/(1-q)}$$

This metric is the effective number of equally distinct species and has all of the same properties as $^{q}D(A)$ and $^{q}D(P)$.

a sample have equal abundances. The q emphasizes that the measure of evenness depends on how abundances are weighted (Table 12.3).

One confusing aspect of the terminology surrounding inventory diversity was that Whittaker proposed a hierarchy that depended on the spatial extent of the sample: point diversity (within a community), **alpha (α) diversity** (community), **gamma (γ) diversity** (landscape), and epsilon (ε) diversity (region). The problem is that those extents were never precisely defined. Hanna Tuomisto (2010a) clarified them by collapsing these indices into just two concepts: γ-diversity, $D(A)_\gamma$; and α-diversity, $D(A)_\alpha$. If you have a set of communities, the inventory diversity of the entire set is γ-diversity, while α-diversity is the average diversity of the individual communities. Besides communities, the units could be small quadrats or entire landscapes. The two measures—γ-diversity and α-diversity—are independent of any particular length, area, volume, or duration.

Courtesy of Kalle Ruokolainen

Hanna Toumisto

Just as for many other measurements, a diversity index has some uncertainty associated with it, so we also need an estimate of the precision of the diversity parameter before comparisons can be made. Methods for estimating the precision of these measures are described by Lande (1996) and Dixon (2001). One major problem with any metric of inventory diversity is that it collapses a great deal of information into a single number. While

**TABLE 12.3** A comparison of changes in how abundances are weighted (q) on measures of diversity and evenness as applied to six communities, each containing five species

| | Community | | | | | |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| Species A | 20 | 30 | 40 | 50 | 60 | 960 |
| Species B | 20 | 30 | 30 | 20 | 10 | 10 |
| Species C | 20 | 20 | 10 | 10 | 10 | 10 |
| Species D | 20 | 10 | 10 | 10 | 10 | 10 |
| Species E | 20 | 10 | 10 | 10 | 10 | 10 |
| Sample size | 100 | 100 | 100 | 100 | 100 | 1000 |
| $^0D$ (= *R*) | 5 | 5 | 5 | 5 | 5 | 5 |
| $^1D$ | 5 | 4.50 | 4.13 | 3.89 | 3.41 | 1.25 |
| $^2D$ | 5 | 4.17 | 3.57 | 3.13 | 2.50 | 1.08 |
| $^0E$ | 1 | 1 | 1 | 1 | 1 | 1 |
| $^1E$ | 1 | 0.90 | 0.83 | 0.78 | 0.68 | 0.25 |
| $^2E$ | 1 | 0.83 | 0.71 | 0.63 | 0.50 | 0.22 |

that number might be convenient for making comparisons among communities or for the same community over time, any time you squash a lot of information into a single point (the metric), you lose a lot of information. Another problem is that this loses any information about where things are found: this approach is not spatially explicit, so a community in which species are evenly dispersed throughout is not distinguished from one in which species are not intermingled. One solution to these problems is to use other methods besides these indices for quantifying diversity and comparing diversity among communities. These methods include species-area curves (discussed below) and graphical approaches based on comparisons of abundance curves (discussed more fully in Chapter 14).

## Differentiation diversity is the variation among units

Imagine yourself walking through a forest. You stop and identify the species around you. You walk a bit farther and stop again. Again you identify the species you find. Beyond measuring inventory diversity, now you can ask an additional question: How different are the communities you passed through? Did you find mainly the same species as you walked along, or different ones? Those differences are quantified by **differentiation diversity**, how the composition of named objects (in this case, species) differs among groups (in this case, communities). It is typically referred to as **beta (β) diversity**, following the terminology of Robert Whittaker (1960). Of course, if you were doing a formal analysis of patterns of diversity, you would collect your data in a structured fashion, as discussed later in this chapter.

How we should measure differentiation diversity has had a controversial history in ecology because Whittaker originally defined β-diversity several different ways—the ratio of γ-diversity to α-diversity, the difference between γ-diversity and α-diversity, similarity in species composition among sites, and turnover in species composition along a gradient—while using the same term and symbol for all of them. Which of these versions had ascendancy has waxed and waned over the years. During the 1970s and 1980s, the primary measure was species turnover. More recently the argument has been over whether to quantify the relationship between α-diversity and γ-diversity using a multiplicative (a ratio) or additive (a difference) approach.

In an extensive review, the Finnish ecologist Hanna Tuomisto (2010b) showed that all of the measures could be mathematically related to each other and that the multiplicative relationship had many useful properties. Using her approach, the inconsistencies of Whittaker's definitions can be resolved mathematically to provide a better way of thinking about differentiation diversity. Consider a set of three communities—one has six species, one has three species, and one has two species (Figure 12.4); they share some species, so there are nine species total in the entire set. How different are they? For these communities, γ-diversity = 9.0 (the total number of species) and α-diversity = 3.7, the average of 6, 3, and 2. Finally, β-diversity = 9.0/3.7 = 2.4. So the effective number of communities is 2.4 because if each of those communities contained 3.7 unique species, the total number of species would be 9.0 (= 2.4 communities × 3.7 species/community), the number of species in the entire set. While it might seem odd to be speaking of fractional species and subunits, these are theoretical quantities, rather than actual entities.

## Phylogenetic diversity is variation in evolutionary relationships

One important aspect of diversity that is not captured by those previously discussed is the degree of evolutionary variation among the species in a community.

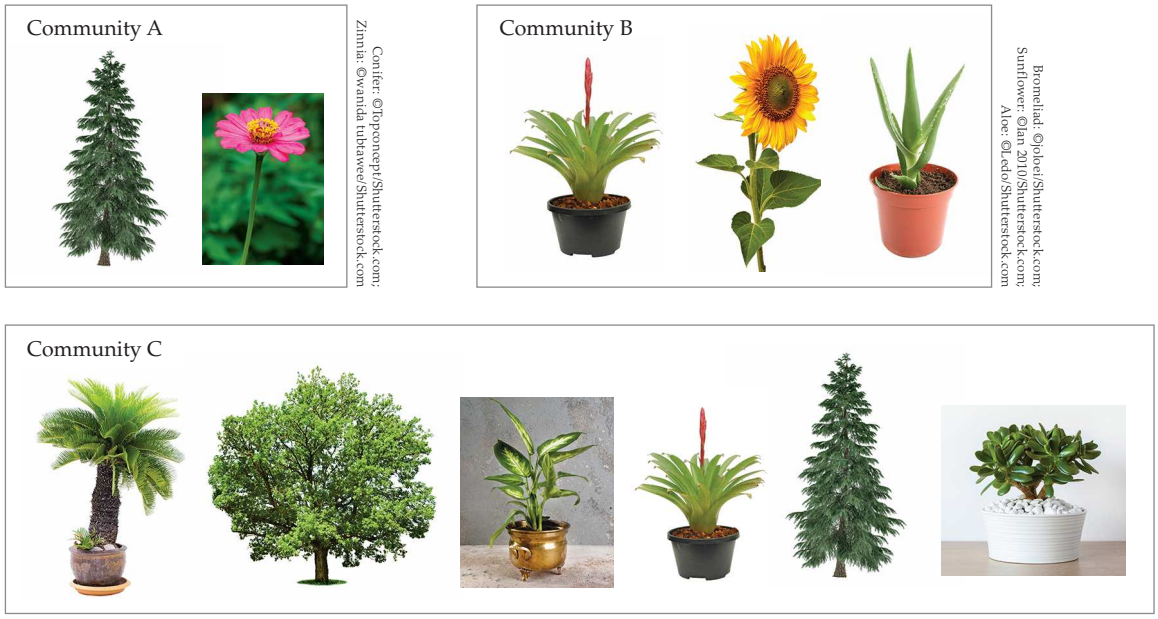



**Figure 12.4** The concept of differentiation diversity as shown by three communities that contain a total of nine species. The communities share some species, but each has at least one species not shared by the other communities. The communities differ in their phylogenetic diversity: community A contains a conifer and an angiosperm, while community B contains only angiosperms. The communities also differ in their functional diversity: community B contains only herbaceous species, while communities A and C both contain trees.

This variation can be measured by examining the phylogenetic relationships among species (Faith 1992; Cadotte et al. 2010). **Phylogenetic diversity** is the extent to which species within a set (such as in a community) are closely or distantly related. Because closely related species tend to have similar traits and ecosystem functions, phylogenetic distances may provide a measure of trait or functional similarity. For example, if only some evolutionary lineages are likely to adapt to more stressful environments (e.g., colder or drier habitats), phylogenetic diversity will be lower in extreme conditions. Communities composed of species spanning a greater proportion of a **cladogram** (a phylogenetic tree) may be more diverse because those species are likely to have different ecological functions, or if they have similar ecological functions, those functions are likely based on different physiological or morphological traits. We could use these measures in a variety of ways: to test hypotheses about the limits of adaptation, or to determine how the amount of diversity might affect ecosystem processes (see Chapter 14).

Phylogenetic diversity can be characterized with respect to the magnitude (divergence) and variability (regularity) in the amount of evolutionary distance among a set of species. The total lineage divergence represented by the cladogram in Figure 12.5 is simply the sum of the branch lengths; it is the oldest metric and known as

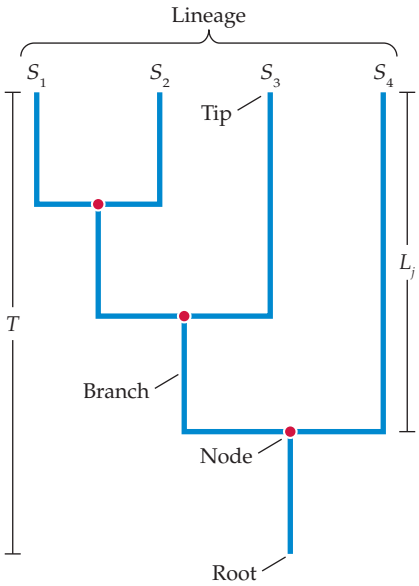



**Figure 12.5** The components of a cladogram, a diagram of the evolutionary relationships among a set of species. Each species consists of a lineage measured over a total duration $T$, the time from the tip to the root. Each branch of the cladogram has divergence $L_j$ that may be shared by multiple lineages that coalesce at a given node. Assume that for species 1 the length of the branch from the tip to its first node is 2, so the total length of the branches of the entire cladogram is 20. First divide the divergence of a given branch, $L_j$, equally among lineages that share that branch ($S_j$), and $b(S_i)$ is the set of branches in the path from the root to the tip of the $i$th species, then the proportional share for the $i$th species of a given branch is $L_{ij} = L_j/S_j$. The amount of divergence of the $i$th species ($L_i$) is then the sum of those proportional branches. Species 1 has a total divergence of (2/1 + 2/2 + 2/3 + 2/4) = 4.17. Similarly, species 4 has a total divergence of (6/1 + 2/4) = 6.50. The total lineage divergence is the sum of the branch lengths.

$$L = \sum_i^s \sum_{j \in b(s_i)} L_{ij}$$

Faith's phylogenetic diversity (Faith 1992). This value, $L$, is analogous to the total number of individuals, $N$, for abundance information. If we want to compare the divergence of two communities, we need to standardize the cladograms to the same time depth ($T$) by dividing the total branch length by $T$. The result is M(PR) = $L/T$, which measures the effective number of phylogenetically independent species. For example, in Figure 12.5, there are four species, but the effective number of phylogenetically independent species is 2.5 because some species are closely related. When all species are phylogenetically independent so that they diverge at the root of the cladogram, M(PR) equals the total number of species. Divergence of all species at the root cannot happen for actual relationships, as divergences always occur as pairs of splitting species, but we may be unable to measure those pairwise splits because they occurred closely spaced in time; thus the data may show only multiple splits occurring together. We can also calculate the effective number of phylogenetically distinct species, ${}^{q}D(P)$ (see Box 12C). As with abundance diversity, our measure of variability, regularity, can be computed as ${}^{q}E(P) = {}^{q}D(P)/R$ (Table 12.4, Figure 12.6). As with species richness and abundance diversity, phylogenetic diversity can be analyzed with respect to α-, β-, and γ-diversity (see Figure 12.4). For example, we might want to know if the species within each community in a landscape are more or less closely related than those of the entire set of species in that landscape, which might indicate that processes such as competition or abiotic factors are responsible for community composition (see Chapter 14).

**TABLE 12.4** Lineage divergences (in units of time) and phylogenetic diversities (in units of numbers of equivalent species) for the cladograms shown in Figure 12.6

| | A | B | C | D |
|---|---|---|---|---|
| $L_1$ | 4.00 | 5.00 | 3.25 | 4.17 |
| $L_2$ | 4.00 | 5.00 | 3.25 | 4.17 |
| $L_3$ | 4.00 | 5.00 | 3.75 | 5.17 |
| $L_4$ | 4.00 | 5.00 | 5.75 | 6.50 |
| $l_1$ | 0.25 | 0.25 | 0.21 | 0.21 |
| $l_2$ | 0.25 | 0.25 | 0.21 | 0.21 |
| $l_3$ | 0.25 | 0.25 | 0.23 | 0.26 |
| $l_4$ | 0.25 | 0.25 | 0.34 | 0.33 |
| L | 16 | 20 | 16 | 20 |
| M(PR) | 2.00 | 2.50 | 2.00 | 2.50 |
| ${}^{0}D(P)$ | 4.00 | 4.00 | 4.00 | 4.00 |
| ${}^{1}D(P)$ | 4.00 | 4.00 | 3.87 | 3.93 |
| ${}^{2}D(P)$ | 4.00 | 4.00 | 3.75 | 3.86 |
| ${}^{0}E(P)$ | 1.00 | 1.00 | 1.00 | 1.00 |
| ${}^{1}E(P)$ | 1.00 | 1.00 | 0.97 | 0.98 |
| ${}^{2}E(P)$ | 1.00 | 1.00 | 0.94 | 0.96 |

## Functional diversity is variation in traits

Another property of a set of species is their **functional diversity**. The name is somewhat misleading. What is actually measured is how those species differ in their traits. The assumption is that those traits affect the functional roles of the species in the community, for example, how those species affect the rates of carbon captured in photosynthesis (see Chapter 2), such as leaf mass per unit area,

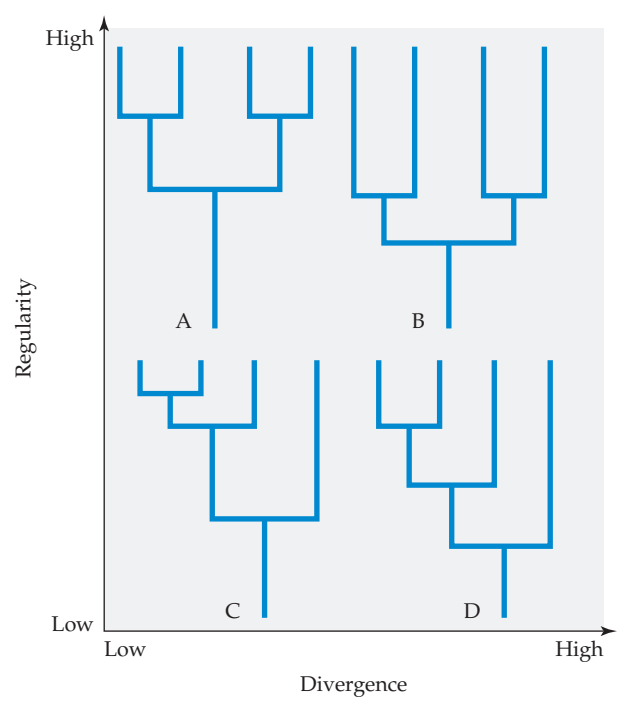



**Figure 12.6** Four cladograms that all have four species but differ in the magnitude (divergence) and variability (regularity) among the species in lineage divergences. Cladograms B and D have longer branches than cladograms A and C, and so they have a greater divergence. Cladograms A and B each consist of four equally divergent species and have a regularity of 1.0 for all values of q (see Table 12.4). Cladograms C and D each have three species that are more closely related than the fourth and so have a lower regularity. (From S. M. Scheiner. 2019. *bioRxiv* 530782. CC0. doi: https://doi.org/10.1101/530782/)

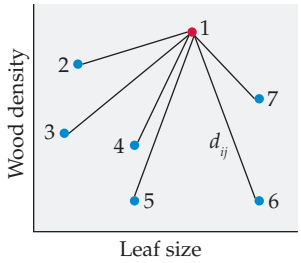



**Figure 12.7** The functional diversity of a set of species can be measured as the distance among the species arrayed in a space defined by their trait values, in this case by the traits of leaf size and wood density. The distance between two species is given by

$$d_{ij} = \sqrt{\sum_k \left(s_{ik} - s_{jk}\right)^2}$$

where $d_{ij}$ is the functional distance between the *i*th and *j*th species, and $s_{ik}$ is the value of the *k*th trait of the *i*th species. The proportional dispersion of the *i*th species is

$$\sum_j^s d_{ij} \Big/ \sum_i^s \sum_j^s d_{ij}$$

**TABLE 12.5** Trait values, distances, and functional diversities (in units of numbers of equivalent species) for the community shown in Figure 12.7

A. The seven species that have been measured for two traits, such as leaf size and wood density

| | Leaf size | Wood density |
|---|---|---|
| Species 1 | 2.1 | 2.8 |
| Species 2 | 0.6 | 2.3 |
| Species 3 | 0.4 | 1.4 |
| Species 4 | 1.4 | 1.2 |
| Species 5 | 1.4 | 0.4 |
| Species 6 | 3.1 | 0.4 |
| Species 7 | 3.1 | 1.8 |

or whether they are wind-pollinated or insect-pollinated (see Chapter 6). One simple way to measure functional diversity is to group the species (e.g., into grasses, forbs, or trees), which are referred to as **functional groups** (see Box 12A). The number of functional groups is a simple measure of functional diversity. Such a measure, however, provides limited information, as it fails to take into account variation among species within each group. For that, we want other measures of functional diversity.

Functional diversity measures some sense of each species' distinctiveness from the other species. For example, leaf size can affect photosynthetic rate and temperature balance (see Chapter 2 and Chapter 3). If all species have very similar leaf sizes, we would not consider them to be very distinctive. In contrast, if one of the species was *Gunnera insignia* (see Figure 3.22), that species would be very distinctive. We begin by determining species characteristics or traits—in this example, leaf size and wood density (Figure 12.7). Each trait defines an axis, and together those traits define a space within which the species are arrayed as a set of points. The magnitude of functional diversity (**dispersion**) is the mean distance among all of those points (Table 12.5). As with abundance and phylogenetic diversity, this metric can be standardized to have a maximal value equal to the number of species so that it measures the effective number of equally distant species, M(TR).

To measure variability of functional diversity (**equability**), we can use the same method that we used for abundance and phylogenetic diversity (see Box 12C). For example, in Figure 12.8, communities A and B have the species arrayed at the points of a pentagram, so all species have the same average distance to all other species and ${}^{q}D(T)$ equals species richness (*R*) for all q, just as with the other metrics. As with the other types of information, functional evenness can be measured as ${}^{q}E(T) = {}^{q}D(T)/R$,

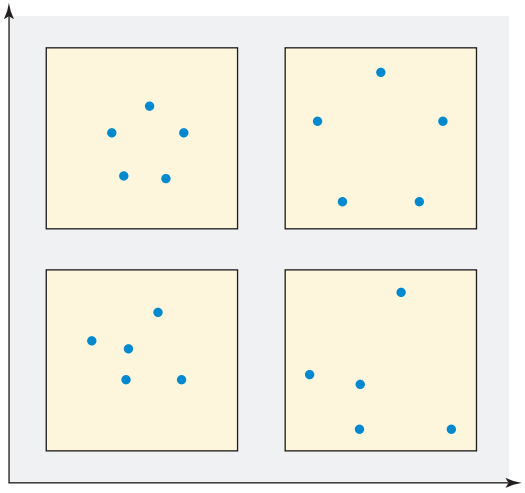


**TABLE 12.5** (*continued*)

B. The distances among species, which have been standardized to have a maximum value of 1.0, and the mean proportional dispersion for each species

| Community A | Species 1 | Species 2 | Species 3 | Species 4 | Species 5 | Species 6 | Species 7 |
|---|---|---|---|---|---|---|---|
| Species 1 | 0.00 | 0.63 | 0.88 | 0.69 | 0.97 | 1.00 | 0.53 |
| Species 2 | 0.63 | 0.00 | 0.38 | 0.53 | 0.80 | 1.23 | 1.00 |
| Species 3 | 0.88 | 0.38 | 0.00 | 0.39 | 0.53 | 1.13 | 1.08 |
| Species 4 | 0.69 | 0.53 | 0.39 | 0.00 | 0.30 | 0.75 | 0.73 |
| Species 5 | 0.97 | 0.80 | 0.53 | 0.30 | 0.00 | 0.68 | 0.88 |
| Species 6 | 1.00 | 1.23 | 1.13 | 0.75 | 0.68 | 0.00 | 0.56 |
| Species 7 | 0.53 | 1.00 | 1.08 | 0.73 | 0.88 | 0.56 | 0.00 |
| Mean proportional dispersion ($d_i$) | 0.2 | 0.1 | 0.1 | 0.1 | 0.1 | 0.2 | 0.2 |

C. Three measures of functional diversity—the effective number of equally distant species, M(TR); the effective number of equally distinct species, qD(T), for different values of q; and functional equability

| | |
|---|---|
| M(TR) | 5.22 |
| $^0$D(T) | 7.00 |
| $^1$D(T) | 6.94 |
| $^2$D(T) | 6.89 |
| $^0$E(T) | 1.00 |
| $^1$E(T) | 0.99 |
| $^2$E(T) | 0.98 |

and functional diversity can be divided into components within and among communities (see Figure 12.4).

### Different types of biodiversity information can be combined

The previous discussion of biodiversity metrics focused on species identities or the combination of species identities with information about abundance, phylogenetic relationships, or functional traits to create a set of metrics that are all in units of the effective number of species. Such an approach puts all of these metrics into units of the effective number of species, permitting direct comparisons. The biodiversity metrics presented here are not the only ones available. Other combinations of types of information can also be useful. For example, we might expect the effect each species has on an ecosystem function to be proportional to its abundance. In that case, we would want a metric that combined functional and abundance information. Many other approaches have been proposed for measuring and combining these various biodiversity properties, which are described in detail by Magurran and McGill (2011), Tucker et al. (2017), and Scheiner (2019).

## 12.3 Communities Can Be Measured in Many Ways

Assessing diversity requires collecting information on species distributions and their abundances, phylogenetic relationships, and trait characteristics. We might also be interested in the physical structure of the vegetation in the community: its **physiognomy**. In this section we explore the many ways that plant ecologists have developed to measure these properties. These methods are now well established within plant ecology with respect to distributions and abundances, but they have not yet been applied to phylogenetic and functional data.

### Measuring species richness can involve simple sampling procedures or complex mathematical estimates

While simply finding and listing all species is useful, this method has limitations. Most importantly, if we wish to compare communities, we need comparable samples—otherwise we might find that two communities are "different" simply because we sampled one more intensively than the other. The area sampled and method of sampling can have strong effects on the number of species found. To deal with the problem that the number of species found is dependent on sampling effort, we can use plot-based methods, in which sample plots, or **quadrats**, are marked out in a community and a list of species is collected for each. Ecologists use quadrats of different shapes, such as square, rectangular, or round, and they may be nested, contiguous, spaced along a line, placed in a grid, or placed randomly. These different arrangements can be used to ask different kinds of questions or to control for variation that occurs on different spatial scales (Krebs 1989). How large an area

should be sampled? We want to sample a sufficient area to include most of the species in the community and to minimize differences due to random sampling effects, such as missing a species because it is found in just one small patch. On the other hand, we do not want the area to be bigger than necessary, for practical reasons of time and effort and also because you might start increasing differentiation diversity if the area sampled spills over into very different vegetation.

By looking at how the total number of identified species increases as the data from different quadrats are combined, one can examine the effects of area on species richness. This issue was recognized over 150 years ago (de Candolle 1855). The **species-area curve** (Arrhenius 1921; Gleason 1922; Cain 1934) describes the increase in the number of species found as the area sampled increases. For terrestrial plants, this relationship is most commonly measured as a function of area, but you can also examine the change in species numbers with repeated sampling of the same quadrat over longer periods of time (e.g., White et al. 2006; Shurin 2007) or over both space and time (e.g., Adler et al. 2005; Fridley et al. 2006). For aquatic or soil systems, volume might be a more appropriate measure than area. So, although the term *species-area curves* is most commonly used, more generally we refer to these as **species richness relationships** (**SRRs**). These relationships can be used in two ways: to estimate **species density** (species richness at a specified area), and to make inferences about the processes responsible for the observed species richness.

Five processes are responsible for determining the shape of the species richness relationship. First, as the area sampled increases, more individuals are found and the chance of encountering a new species increases. Second, a larger area is more environmentally heterogeneous. For a given community, if it has a relatively uniform environment, the number of new species found for each increment in sampling area decreases. Eventually, very few or no new species are found, and the relationship reaches a plateau. Of course, if the area becomes large enough to encompass new environments, the curve begins to rise again. Third, individuals belonging to the same species may be clumped in their distibution because of limits to dispersal distances. For example, if seeds are wind dispersed, most end up near the parental plant, while many of those that are animal dispersed might often end up together in a seed cache (see Chapter 7). Fourth, competition among individuals (see Chapter 10) could create overdispersion among individuals that belong to either the same or different species. Fifth, if the sampling extends over a great physical distance such as an entire continent, or over very different environmental conditions such as a wetland and a dry upland, it might include multiple species pools. For example, in North America the eastern forests contain a very different set of species than the midcontinent grasslands. Together, these processes encompass a theory of some types of species richness relationships (Table 12.6). A separate theory can be built from the equilibrium theory of island biogeography (see Chapter 15).

There are two general approaches to building SRRs. One, a **spatially explicit approach**, depends on determining the particular spatial arrangement of species and landscape elements (e.g., habitat patches, farms, roads). The other uses a nonspatially explicit approach to measure average parameter values.

To see this distinction, imagine that we do the following: On the meadow shown in Figure 12.9, we overlay a

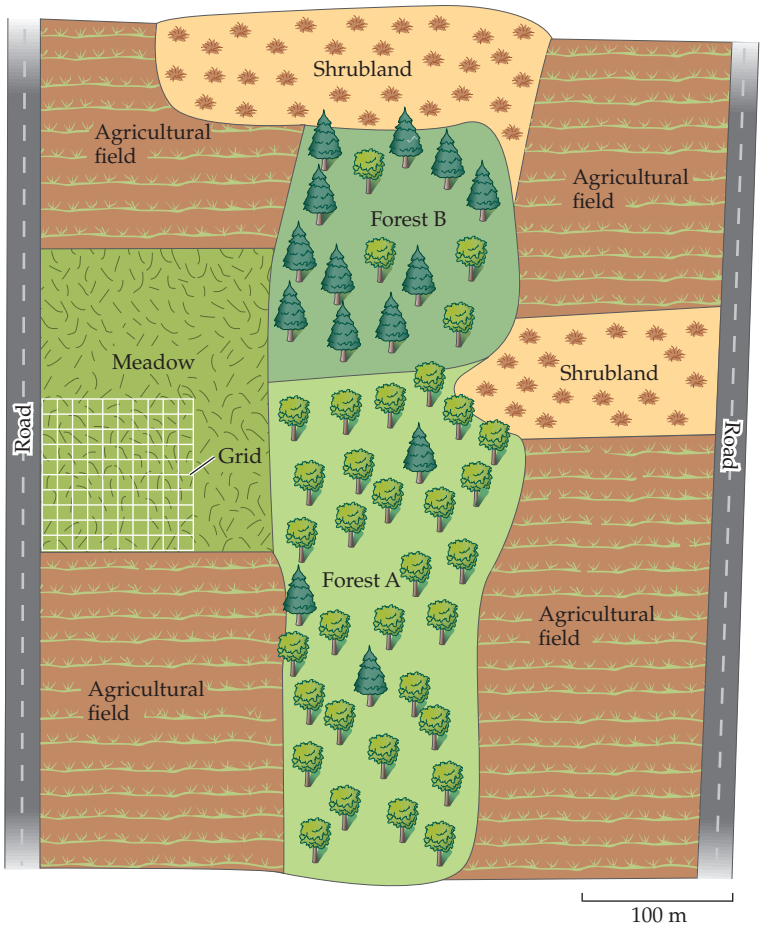



**Figure 12.9** Diagram of a landscape showing a patchwork of meadows, forests, and shrublands. For sampling purposes we can lay down a 100 × 100 m grid that arbitrarily divides the area into smaller plots.

**TABLE 12.6** The domain and propositions of a constitutive theory of type II and III species richness relationships (SRRs)

| Domain: Species richness relationships built from the continual aggregation of sampling units. The sampling units can be over space or time. The units can be contiguous or noncontiguous. The process of aggregation may or may not account for the spatial or temporal relationships of the sampling units. |
|---|
| Propositions |
| As sampling increases to include more area or more time, the number of individuals sampled increases. |
| 1. More individuals leads to more species. |
| 2. As sampling increases to include more area or more time, more environmental variation may be encountered. |
| 3. If all individuals of all species are uniformly distributed, and if the sampling unit is smaller than the mean distance between individuals, the SRR will have a positive relationship between space or time and species richness. |
| 4. Clustering in space or time of individuals of the same species will increase the rate at which species richness increases over space or time. |
| 5. Competitive exclusion of conspecifics or different species can create overdispersed or uniform distributions of species. Such distributions will decrease the rate at which species richness changes over space or time. |
| 6. If the sampling regime of a particular SRR is broad enough to encompass long time scales or large areas, those samples may include multiple species pools. |
| 7. The inclusion of multiple species pools will increase the rate at which species richness increases over space or time for spatially explicit or temporally explicit SRRs. |

*Source*: S. M. Scheiner et al. 2011. *Ecol Monogr* 81: 195–213.

grid of 10 × 10 m quadrats, and we compile a list of every vascular plant species growing in each square. We can now build a species-area curve in two ways (Figure 12.10). Using the spatially explicit approach, we start in one spot. Imagine that this spot is the bottom left corner of the meadow. Our first data point is the number of species in the 10 × 10 m square or quadrat in that corner. Next, we expand our square to 20 × 20 m and again count the total number of species. This number will include all of the species in the original square plus any new species in the larger square. We repeat this operation for a 30 × 30 m square. Initially, the number of species will rise rapidly. But as more and more of the meadow is captured in our ever-growing plot, the number of new species found with each expansion will be smaller because we already found most of them in the quadrats sampled before. At some point, it is likely that we will stop finding new species entirely, and the curve will level off; it will reach an asymptote. If we keep going, however, our plot will eventually run into the forest abutting the meadow. Now, suddenly, we will find a whole new suite of species, such as trees and forest floor herbs. The species-area curve will again rise rapidly with increasing area until, once again, it levels off as most of the forest gets sampled. This stair-step pattern will be repeated each time we cross a new community boundary (assuming that there are clear boundaries between communities).

In contrast, we can build our SRR using the **mean field approach**, which ignores the spatial arrangement of the squares. First, we calculate the average number of species in each of the 10 × 10 m squares. Next, we take all possible combinations of two squares, determine the total number of species in each pair of squares, and again calculate the average. (While we would not want to do

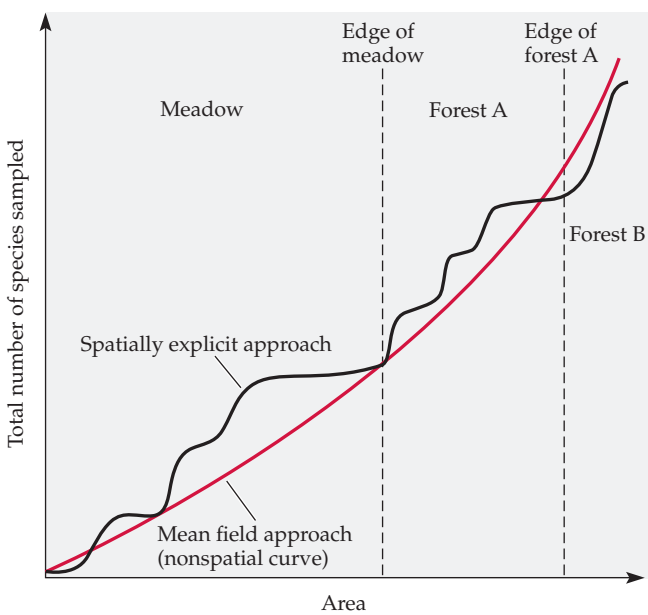



**Figure 12.10** Two species-area curves. One curve is from a spatially explicit tally, which forms a stair-step pattern. Within a community, the number of species found as area increases will level off. When a community boundary is crossed, however, the number of species found will once again rise rapidly. The other curve, from a mass field analysis that is not spatially explicit, forms a smoothly rising pattern.

this by hand, it can be done in a few microseconds on a typical personal computer.) We continue this process with all sets of three squares, four squares, and so forth. The result is a smoothly rising curve. Any discontinuities in the environment that would have produced a stair-step effect with the spatially explicit approach are, instead, averaged out. Which of these approaches is preferable? As each approach provides different information, the answer depends on the questions one wishes to address.

The work of Sam Scheiner and others demonstrates that the two analytic approaches (spatially explicit and mean field) can be combined with four different sampling schemes (Figure 12.11) to yield six different types of species-area curves (Scheiner 2003; Table 12.7). The stair-step curve shown in Figure 12.10 is an example of a type I curve. Such a curve is always based on a spatially explicit approach. The type I curve was the original way that ecologists built species-area curves. However, because it consists of a single quadrat at each size, it is not an accurate method for measuring species richness. It is much better to use type II or type III curves.

Courtesy of S. Scheiner

Sam Scheiner

A type IV curve, like a type I curve, is built from single data points. The difference is that each data point is from a ple of a unique area. It is typically constructed from samples of islands or island-like habitats (e.g., lakes, mountaintops, isolated forest stands, continents), so it is also known as an island species-area relationship (ISAR). The type IV curve is associated most often with oceanic islands due to its use in conjunction with the theory of island biogeography (see Table 15.1).

These curves yield different estimates of species richness. Spatially explicit curves (types IIA and IIIA) estimate mean α-diversity and its rate of change with sampling scale; put differently, the slope of a type A curve is an estimate of the rate of change of mean α-diversity as plot size increases (Tuomisto 2010b). Alternatively, curves built with nonspatially explicit curves (types IIB and IIIB) estimate mean γ-diversity and its rate of change as the size of the total area sampled increases (Figure 12.12). The conceptual difference between α- and γ-diversity is thus tied to the relationship between the area of the plot and the area of the entire sample. We return to these effects when we examine the components of scale and their effects on patterns of diversity in Chapter 15.

Once a curve for the SRR is constructed, the next step is to determine its mathematical form (Tjørve 2009). Many different mathematical models for the shape of the SRR have been proposed (Triantis et al. 2012), which all have one of two basic shapes: some are concave (the number of species always increases with area, but at a decreasing rate), and some are sigmoid (sometimes called S-shaped; these have the number of species initially increasing with area at an accelerated rate, but then switching to a decreasing rate up to some asymptotic limit). This difference in shape is important if one wishes to answer the question, How many species are there in the community? If the species-area curve rises indefinitely with increasing area, then the answer depends entirely on the area of the community. However,

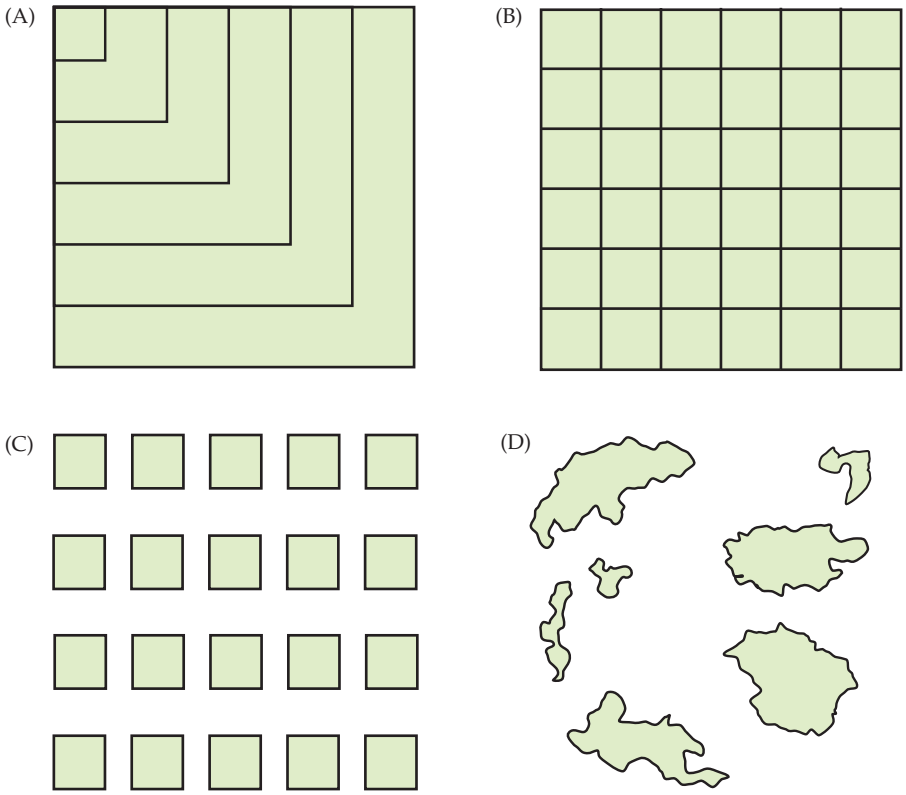



**Figure 12.11** Species-area curves can be built from four general sampling schemes: (A) strictly nested quadrats (type I curves), (B) quadrats arrayed in a contiguous grid (type II curves), (C) quadrats arrayed in a regular but noncontiguous grid (type III curves), and (D) areas of varying size, often islands (type IV curves). (After S. M. Scheiner. 2003. *Global Ecol Biogeogr* 12: 441–447.)

**TABLE 12.7** Examples of the six types of SRRs with a description of their features and scale parameters

| Type | Sampling scheme | Species density | Construction spatially explicit | Grain | Focus | Extent |
|---|---|---|---|---|---|---|
| I | Nested | The number of species in a contiguous sample unit of specified size | Yes | A sample unit nested within the larger or longer one | Same as the grain | The largest or longest sampling unit |
| IIA | Contiguous | The number of species in a contiguous sample unit of specified size | Yes | One or more adjacent sampling units | The cumulative area or time of all sampling units | Same as the focus |
| IIB | Contiguous | The number of species in an aggregated sample unit of specified size | No | One or more aggregated sampling units | The cumulative area or time of all sampling units | Same as the focus |
| IIIA | Noncontiguous | The number of species in an aggregated sample unit of specified size | Yes | One or more neighboring sampling units | The cumulative area or time of all sampling units | The maximum distance or time among sampling units |
| IIIB | Noncontiguous | The number of species in an aggregated sample unit of specified size | No | One or more aggregated sampling units | The cumulative area or time of all sampling units | The maximum distance or time among sampling units |
| IV | Independent units | The estimated number of species in a sample of a specified size | No | Independent space or time units | The cumulative area or time of all sampling units | The maximum distance or time among sampling units |

*Source*: S. M. Scheiner et al. 2011. *Ecol Monogr* 81: 195–213.

the "area" of a community is often arbitrary, making the question meaningless. Instead, we can rephrase the question as, How many species are contained in an area of size X? This value is known as **species density**. If the species richness relationship is best described with a sigmoid function, then this question has meaning, because the function has a definite plateau or asymptote.

The issue of the shape of the relationship and the sampling and analysis methods used is important for comparisons of diversity among communities, as such comparisons are only meaningful if they are comparing the same things. So if the SRR is concave, it is only meaningful to compare the species densities of two communities at the same area. If the areas sampled in the communities differed, the comparison would need to use the smaller area. An alternative to standardizing to a common area is to standardize to a common number of individuals or samples, a process called **rarefaction**.

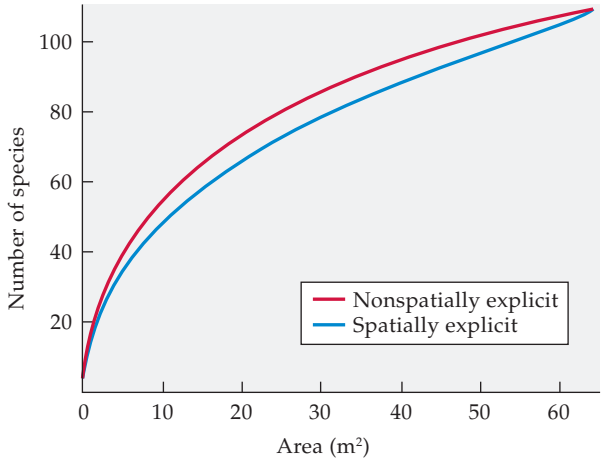



**Figure 12.12** Type IIIA (spatially explicit) and type IIIB (nonspatially explicit) SRRs from a vegetation survey of the Oosting Natural Area of the Duke Forest, North Carolina (Reed et al. 1993; Palmer and White 1994; Palmer et al. 2007; Chiarucci et al. 2009). The vegetation was sampled in a 16 × 16 grid of 256 contiguous modules, each module being16 m × 16 m. Six nested quadrats (with sides of 0.125 m, 0.25 m, 0.50 m, 1 m, 2 m, and 4 m) were located in the southwestern corner of each module, and in each the presence of all vascular plant species was recorded (Reed et al. 1993). These data are for the 0.25 $m^2$ quadrats. The points along the spatially explicit curve are measures of mean α-diversity; those of the nonspatially explicit curve are measures of mean γ-diversity. (From S. M. Scheiner et al. 2011. *Ecol. Monogr.* 81: 195–213.)

The dynamic nature of communities affects patterns of species richness in time as well as in space (Adler et al. 2005; Fridley et al. 2006; White et al. 2006; Soininen 2010). We might want to know if over many years a particular place has a limit to the number of species that are found there, or if the list of species continues to grow indefinitely. Imagine a set of quadrats in which the species present are recorded once a year. Over time, migration will add new species to the quadrats, while local extinction will eliminate others. In this instance, the number of species present in any single year may not change, although the total number of species seen over all years will keep increasing, reflecting the species that exist in the larger landscape around that community. If you graph the total number of species observed as a function of time, you will have a time-based SRR.

What is the equivalency of time and space? For example, how many years would it take for a 100 $m^2$ plot to include as many species as a 1000 $m^2$ plot? Peter Adler and William Lauenroth (2003) asked this question for quadrats at the Konza Prairie Research Natural Area in Kansas. They found that, on average, in a single year an area of 1000 $m^2$ contained about 100 species, and that it would take about 20 years to observe that many species in a 100 $m^2$ plot (Figure 12.13). This equivalency depended on the scale of the observation, however. In a single year, a 100 $m^2$ quadrat contained about 50 species, and it would take only about 8 years to observe that many in a 10 $m^2$ quadrat. In contrast, at large scales, it would take about 80 years to observe as many species in a 10,000 $m^2$ area as found in a 100,000 $m^2$ area in a single year. The reason is that at these larger scales, a 10,000 $m^2$ area contains nearly all of the species in the landscape, so while species are moving around within that landscape, on average the landscape has very few new species migrating into it. Such understanding of the dynamics of species over both time and space is critical in making decisions about how much land it is necessary to conserve to preserve species diversity (see Chapter 15).

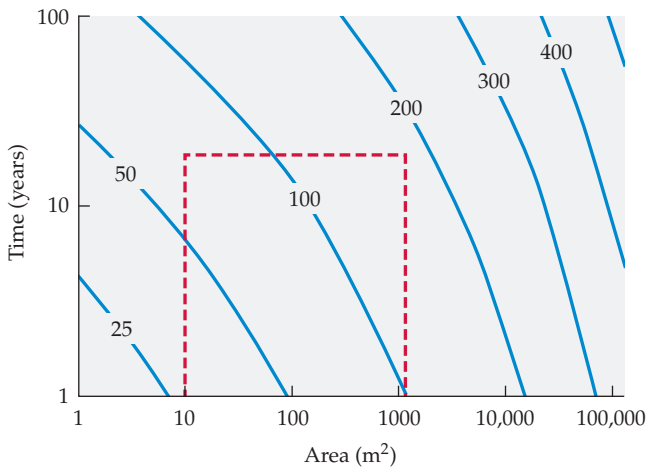



**Figure 12.13** Species richness in the Konza Prairie Research Natural Area, Kansas, as a function of area and time. Both area and time are shown on a logarithmic scale. The curved lines are isopleths indicating the number of species that would be observed for a particular combination of area and time. The lines are based on a model using data from 10 $m^2$ quadrats. The dashed box outlines the time and space extents of the data; extrapolation outside that extent assumes that the model parameters remain the same. (After P. B. Adler and W. K. Lauenroth. 2003. *Ecol Lett* 6: 749–756.)

## There are many ways to sample communities

Several techniques can be used to sample a community. Which to use in a given case will depend on the type of vegetation being sampled and the goal of the survey. For estimating species numbers using SRRs, type II and III curves constructed from the same number of equal-size plots will differ in the total area encompassed by the plots (see Figure 12.11). This difference leads to an important reason for using a type III (noncontiguous) rather than type II (contiguous) design. If you want to make inferences about a very large area or long time period, and if the environment differs substantially in distant locations or at distant times, noncontiguous plots are more efficient at capturing that heterogeneity.

Plant ecologists influenced by the Zurich-Montpellier school (the Braun-Blanquet approach) typically sample and compare multiple communities by placing a single large sample plot—called a **relevé**—in each stand. The relevé is located subjectively, with an attempt made to place it within a uniform patch of vegetation that is representative of the community. How large should the relevé be? Often, an initial community is chosen and a type I SRR constructed; the total area is increased until the curve reaches a plateau. That size relevé is then used for the remainder of the communities sampled (if they have similar enough characteristics and structure). Because botanists influenced by this school carried out many of the early vegetation surveys in the tropics, the only available quantitative descriptions for many of these areas have been done using the relevé method.

Ecologists in English-speaking countries have long used the approach of sampling a given stand or community using some number of smaller, non-nested quadrats, making the size and number of quadrats the same across stands or communities. These quadrats can be placed

either randomly within the community or regularly along a grid. An advantage of using a number of smaller quadrats, rather than a large relevé, is that any local patchiness will even out across the entire sample. This method also provides a measure of local heterogeneity. The issue of regular (systematic) versus random placement of the quadrats has been heatedly debated and involves considerations beyond the scope of this book (interested readers should consult Mueller-Dombois and Ellenberg 1974, Magurran 1988, and Krebs 1989).

The size and shape of the quadrats is also an issue. Forests in North America are often sampled using 0.1 ha quadrats. The total area sampled should generally be determined by the SRR, as we saw above. As for quadrat shape, square quadrats are typically used, especially for small areas (<4 $m^2$). Usually the quadrat is marked with a rigid sampling frame, formerly made of wood or aluminum, now typically made of PVC pipe. Larger quadrats can also be set up using posts and measuring tapes in the field, particularly in forests. For the largest areas, circular quadrats may be more efficient. In this case, a post is placed at the center of the sampling area with a rope or tape measure attached to a swivel point at the top. The desired radius is marked on the rope, and a circular area is swept out. This method has the advantages of avoiding the problem of precisely determining 90° corners for a square or rectangular quadrat, and of having a lower area-to-perimeter ratio than a square quadrat; this minimizes the problem of miscounting plants near the edge of the quadrat because it is difficult to decide if a plant at the edge is inside or outside the quadrat.

Ecological conditions may also dictate the shape of the quadrat. For example, on a steep hillside, if one wanted a given sample to be within a particular elevational contour, a narrow rectangular plot perpendicular to the direction of the slope would be most appropriate. A narrow rectangle might also be preferable in a patchy habitat if one wanted to include more patches in each sample. With a narrow rectangle, however, there is a greater chance of miscounting because more plants are near the edge of the quadrat. The shape and placement of quadrats are known to affect the number of species found. Depending on the pattern of environmental heterogeneity, a rectangular quadrat may contain more species than would square or circular quadrats of the same size, because it is more likely to span very local environmental differences.

A related approach to sampling vegetation is the use of **transects**: very long, narrow sampling areas along which species abundances are determined. Often transects are designed to run either across or parallel to gradients of environmental change, depending on the purpose of the sample. Transects are particularly useful in sampling very large areas. One kind of transect is just a line. In the line-intercept method, a line of a set length (e.g., 100 m) is laid through the community, and the proportion of the length of the line occupied by, touching, or under the canopy of each intercepted species is determined. In a forest, a tape might be suspended at some height, and both the overstory species directly above the tape and the understory species directly below the tape might be recorded along the length of the line.

A similar method is the belt transect. As with the line-intercept method, a line of a set length is laid through a community. Then all individuals lying within a belt of some set width on either side of the line (e.g., 0.5 m) are counted (belt transects can also be used to measure cover; below). A belt transect, in other words, is a very long, thin quadrat. A long belt transect may encompass more topographic or soil variation, and thus may include more species, than a square grid for a given total area sampled.

Quadrats and transects can be used to estimate several quantitative measures based on species' presence or abundance in the samples. **Frequency** is the percentage of quadrats in which a species appears. Such a measure has the advantage that it can be done quickly, even by a single person, because it is not necessary to determine how many individuals of a given species are in a quadrat. However, it is inaccurate for very rare species, or species with clumped distributions, and uninformative for very common species (those that appear in every quadrat).

**Density** is the number of individuals of a species per unit area. Density can be measured only for species with distinct individuals (e.g., trees) or where ramets are treated as individuals. For trees, it is common to count only those larger than a certain size. For small plants, quadrats are used and all of the individuals are counted. For larger individuals, particularly trees, plotless methods, such as the point-centered quarter method, can be used. In this method, a series of random points are located in the community, often along a transect. The area around each of the points is divided into 90° quarters, usually along compass headings. Then, in each quarter, the distance to the nearest individual from the central point is measured. These distance measures are then converted into densities. (Greig-Smith [1964] and Krebs [1989] discuss the details of these and related methods.)

Because different species may be represented more heavily by one or another of the above measures of abundance, it is sometimes useful to combine them in a weighted measure called an **importance value** (**IV**). John Curtis and Robert McIntosh (1951) first defined the

importance value of a species as the sum of the relative cover (that is, cover of that species divided by total cover of all species present), relative density, and relative frequency of the species in a community. Importance values can also be defined by combining other measurements of relative abundance, depending on what is most relevant in a given study.

Two measures of species abundance are often used instead of counting the number of individuals: cover and biomass. **Cover** is the percentage of the ground covered by a given species, and it may be measured as **basal cover** or **basal area**—the area occupied by the base of a plant, such as a tussock grass or a tree with a definable base—or as **canopy cover**—the amount of ground covered by the canopy of the plant. This measure has the advantage of being relatively easy to do, and it avoids the problem of how to count individuals for species that have multiple ramets (see Chapter 6). Cover for a given spot can tally to more than 100% if there are several layers of vegetation (e.g., in a forest, trees may cover 70% of the ground, herbs 50%, and shrubs 20%). Cover can be estimated in several ways.

The most precise, but most laborious, method of estimating cover is to map the vegetation. Efficient mapping techniques exist for large individuals—shrubs and trees. The line-intercept method is widely used to estimate cover. Alternatively, for herbaceous vegetation, if it is not too dense and the species are distinct, a photograph can be taken from above, which now can be done by drones. Cover is then determined using image analysis software. The point-intercept or point-quadrat method for quantifying cover uses a sampling frame with a grid of strings or a frame that is used to drop pins at precise locations. The identity of the plant at each point where the strings cross or a pin is dropped is then determined. The total number of "hits" (points occupied) by a species is an estimate of its cover, given that enough points are sampled. In other words, if a species were found at 12% of the points, its estimated cover would be 12%.

Photographic and point-intercept techniques are labor intensive and very time-consuming and are best for small sampling areas. For larger areas or less intensive efforts, visual estimates are typically used. Humans are able to discern differences in plant cover of about 10% by eye, so cover can be measured in this way. Nonequal cover classes can also be used (Table 12.8). The use of cover classes is an established part of relevé sampling. For large areas, if just trees are measured, remote sensing techniques are frequently used (see Box 15A).

An alternative measure of abundance is **dry weight** (the weight of a sample after it has been dried to remove its water content and only the actual plant tissue remains), which is a measure of **biomass**. Many ecosystem processes, such as the rate of photosynthesis, are likely to be related to species biomass, rather than numbers of individuals, so this measure is useful for answering questions about such processes. For many species, biomass can be a more appropriate measure than number of individuals because individuals within a population of the same species can differ in size by several orders of magnitude (consider a seedling and a full-grown oak tree).

A downside of measuring biomass is that it is usually destructive, disturbing the community being measured. In addition, such meaures are often of just the aboveground biomass, ignoring the root system, which can be difficult or impossible to excavate with any accuracy.

**TABLE 12.8** Three different systems for estimating the percentage cover by categorizing estimates into a limited number of classes

| Braun-Blanquet | | Domin-Krajina | | Daubenmire | |
|---|---|---|---|---|---|
| Class | Cover range (%) | Class | Cover range (%) | Class | Cover range (%) |
| 5 | 75–100 | 10 | 100 | 6 | 95–100 |
| 4 | 50–75 | 9 | 75–99 | 5 | 75–95 |
| 3 | 25–50 | 8 | 50–75 | 4 | 50–75 |
| 2 | 5–25 | 7 | 33–50 | 3 | 25–50 |
| 1 | 1–5 | 6 | 25–33 | 2 | 5–25 |
| + | <1 | 5 | 10–25 | 1 | 0–5 |
| r | <<1 | 4 | 5–10 | | |
| | | 3 | 1–5 | | |
| | | 2 | <1 | | |
| | | 1 | <<1 | | |
| | | + | <<<1 | | |

*Source*: D. Mueller-Dombois and H. Ellenberg. 1974. *Aims and Methods of Vegetation Ecology*. John Wiley and Sons: New York.

**TABLE 12.9** Distribution (percentage) of growth forms for various representative communities based on the Raunkiaer system

| | TR[a] | S | W | C | T | MTM | O | DG | SD | D |
|---|---|---|---|---|---|---|---|---|---|---|
| Phanerophytes | 96 | 65 | 54 | 10 | 1 | 34 | 30 | 1 | 0 | 0 |
| Chamaephytes | 2 | 17 | 9 | 17 | 22 | 8 | 23 | 12 | 59 | 4 |
| Hemicryptophytes | 0 | 2 | 24 | 54 | 60 | 33 | 36 | 63 | 14 | 17 |
| Cryptophytes | 2 | 5 | 9 | 12 | 15 | 23 | 5 | 10 | 0 | 6 |
| Annuals | 0 | 10 | 4 | 7 | 2 | 2 | 6 | 14 | 27 | 73 |

*Source:* R. H. Whittaker. 1975. *Communities and Ecosystems*, 2nd ed. Macmillan: New York, NY.
*Note:* Raunkiaer's growth forms are defined in the legend of Figure 6.3
[a] TR = Tropical rain forest; S = Subtropical forest; W = Warm-temperate forest; C = Cold-temperate forest; T = Tundra; MTM = Mid-temperate mesophytic forest; O = Oak woodland; DG = Dry grassland; SD = Semidesert; D = Desert.

For trees, a nondistructive alternative is to measure the **diameter at breast height** (**DBH**). DBH is measured at a standard height above the ground (curiously, 1.3 m in some countries and 1.4 m in others), often by measuring circumference using a special tape that converts circumference to diameter units. For some tree species, especially those that are commercially important, the relationship between DBH and total biomass is known. If the relationship is not known, the investigator can determine it by measuring and harvesting a small sample of individuals.

## One measure of a plant community is its physiognomy

**Physiognomy** is the form, structure, or appearance of a plant community. Knowledge of the physiognomy of a community may tell us about the adaptations of its dominant species to environmental conditions. The Danish ecologist Christen Raunkiaer developed one widely used system for describing community physiognomy. He classified species based on their growth form and on the location of their organs (such as meristems and underground bulbs) that can survive an unfavorable season or conditions, such as winter or a dry season (see Figure 6.3). According to this system, a tropical forest is made up almost exclusively of **phanerophytes** (Table 12.9). A desert community is typically dominated by shrubs—**chamaephytes**—and annual herbs. A prairie is dominated by grasses and perennial forbs—**hemicryptophytes** and **cryptophytes**. We can use this system both to classify general vegetation types (Raunkiaer's original intent) and to make inferences about the environmental factors that shape communities. Prairies, for example, are dominated by species whose meristems are at or below ground level, suggesting that in these systems, meristems above ground level are vulnerable to damage. In this case, the critical environmental factors are fire and grazing (see Chapter 13 and Chapter 19).

## Long-term studies are important for measuring communities

The measurements described above are like snapshots of a community at a single point in time. But ecological processes, such as succession (see Chapter 13) or response to global change (see Chapter 16), can take many years, so these short-term measurements may be inadequate to describe a community. Changes in a community may be gradual, as the environment changes from year to year, or dramatic, as after a hurricane or major fire. Therefore, a sample taken in a single year may provide only a part of the picture; in any case, its value is greatly enhanced by other samples taken at other times. If we want to study the effects of environmental variation, we need to sample that variation adequately; even if we are interested only in knowing what happens on average, a single year's data is not likely to be "average."

Some of our most important ecological questions can be answered only with long-term studies (Leigh and Johnston 1994). A good example of such a question is the effect of global warming. The changes occurring in temperature and precipitation may take decades to be manifested (see Chapter 16). From one year to the next, these changes are expected to be small relative to normal year-to-year variation. By monitoring the same communities and individuals over long periods of time, we may be able to most precisely detect the effects of long-term climate change.

Because ecologists have recognized the value of long-term data, many long-term studies have been

## BOX 12D The Long-Term Ecological Research Network

The Long-Term Ecological Research (LTER) Network is a collaborative effort, involving more than 1800 scientists and students, with the goal of investigating ecological processes that operate at long time scales. The network promotes synthesis and comparative research across sites and ecosystems and among related national and international research programs. The U.S. National Science Foundation established the program in 1980 to support research on long-term ecological phenomena in the United States. The network now consists of 26 sites across the United States and the Antarctic, representing diverse research emphases and ecosystems, from Arctic tundra (Toolik Lake, Alaska), to southern deciduous forest (Coweeta, Georgia), to agricultural fields (Kellogg Biological Station, Michigan), to the desert-urban interface (Phoenix, Arizona). Besides promoting long-term studies, the network also encourages studies that are integrated across a range of disciplines: physiological ecology, population biology, community ecology, and ecosystem science.

Worldwide, there are few sites for which long-term ecological and environmental data exist, so our understanding of ecological phenomena is strongly based on North American and European ecosystems. For example, we know a lot about the impacts of hurricanes on subtropical and temperate forests, but we know much less about Pacific typhoons, even though they affect more forests more often. Global scientific interest in developing long-term ecological research programs is expanding rapidly, reflecting increased appreciation of their importance in assessing and resolving complex environmental issues. In 1993, the U.S. LTER Network hosted a meeting on international networking in long-term ecological research. Representatives of scientific programs and networks that focus on ecological research over long temporal and large spatial scales decided to form the International Long-Term Ecological Research (ILTER) Network. They recommended developing a worldwide program to facilitate communication and the sharing of data. As of December 2018, 40 countries, including Malawi, Hungary, Malaysia, and Mexico, had formed national LTER programs and joined the ILTER Network consisting of over 800 sites. Others are actively pursuing the establishment of national networks, and many others have expressed interest in the model.

There are other, recently created global networks. The Nutrient Network (NutNet) focuses on understanding the relationships between nutrient inputs, productivity, and community diversity (see Chapter 14). It was begun in 2010 and currently includes more than 130 grassland sites worldwide. The PhenoCam Network uses imagery from networked digital cameras to track vegetation phenology (see Chapter 7) in a diverse range of ecosystems. It was established in 2008 and currently includes over 500 sites around the world.

established in different parts of the world. Within long-term study areas, quadrats or individuals are permanently marked. Repeated measurements—perhaps of cover, productivity, diversity, phenology, or other population, community, or ecosystem features—can then be taken according to a set schedule, often yearly but possibly at shorter or longer intervals. In such situations, nondestructive sampling methods are almost always preferable to destructive ones. If destructive methods must be used, they need to be done in predetermined locations so that, for example, the destructive sampling does not affect other sampling areas in the future. Some long-term studies are also useful for studying demography (see Chapter 8).

Ecology is a relatively young science, so most studies that we call "long-term" are less than a few decades old. However, there are several studies that have been going on for much longer. The oldest experiment in ecology is the "Park Grass Experiment" on the effects of nutrient addition and mowing on pasture communities at the Rothamsted Experimental Station (now the Institute for Crops Research-Rothamsted) in Hertfordshire in the United Kingdom, which has been ongoing since the 1850s. Studies on the Swedish island of Öland, led by Eddy van der Maarel and conducted since the 1960s, have included both descriptions of vegetation dynamics and experimental studies of factors maintaining species diversity. Forest dynamics have been studied at Hubbard Brook Experimental Forest in the White Mountain National Forest, near Woodstock, New Hampshire, since 1963. Hubbard Brook is now part of the Long-Term Ecological Research Network that was initiated in the U.S. in 1980 (Box 12D).

## 12.4 Plant Communities Can Be Compared by Many Methods

Given the various ways of quantifying the characteristics of communities, the next question we might want to ask is how different the communities in a landscape are from one another. The techniques used for such comparisons

depend on the nature of the variables being compared. If each community is described by a single characteristic (such as species richness, total biomass, or mean canopy height), then **univariate** (single dependent variable) statistical procedures are used. In contrast, if each community is described by many characteristics—that is, by multiple parameters (such as a list of all of the species present in a community, or their abundances), then **multivariate** (multiple dependent variables) procedures are needed. The latter methods differ in their assumptions and mathematical properties, and the choice among them has been the subject of much debate among ecologists. It is important to distinguish between multiple measures of outcome—multiple dependent variables, for which we need multivariate statistics, and more than one independent variable being used to explain the outcome (the dependent variable). For example, if one forest has 10 tree species, another has 4 tree species, and a third has 12 species, in this framework, the dependent variables are the presences or abundances of each of the 10, 4, and 12 species, and analyzing them requires multivariate statistics. If we try to explain the differences in species composition of the different forests by environmental factors such as the depth of the soil or the amount of yearly rainfall, these are the independent factors (and would not require multivariate statistics, even if there are several of them, as long as there is only a single dependent factor for each forest, such as an index of diversity).

## Non-numerical techniques were the first methods for comparing communities

While the methods used in describing and comparing plant communities have changed considerably over the course of the past half century in response to the ever-increasing availability, speed, and power of computers and software, the questions asked have largely remained the same. One such basic question is, Which communities in a landscape are most similar to one another? Later in this chapter we show how we can quantify the answer to this question. We can also address this question in a more qualitative way, however, using a method favored by European ecologists during the mid-twentieth century.

We might begin, for instance, with a survey of five sites that contain a total of ten species (Table 12.10A). The order of the sites and species in Table 12.10A is arbitrary. (Because this is a hypothetical example, we use numbers for the species, rather than names.) We can then try to change the order of the sites and species so that we group the sites that share the most species and group the species that are most often found together. This reordering is shown in Table 12.10B. The reordering has produced some patterns, with a group of species that are shared by sites A and B appearing in the upper left-hand corner and a group of species that are shared by sites C and D in the lower right-hand corner, while site E shares species with both groups. Species 2, 9, 10, and 5 are shared by both groups, while some of the species

**TABLE 12.10** Using the Jaccard index to measure the similarity of sites based on presence/absence data

(A) Typical presence/absence data[a]

| | Sites | | | | |
|---|---|---|---|---|---|
| Species | A | B | C | D | E |
| 1 | 1 | 1 | 0 | 0 | 1 |
| 2 | 1 | 1 | 1 | 0 | 0 |
| 3 | 0 | 0 | 1 | 1 | 1 |
| 4 | 0 | 0 | 1 | 1 | 0 |
| 5 | 1 | 0 | 1 | 1 | 0 |
| 6 | 0 | 1 | 0 | 0 | 1 |
| 7 | 0 | 0 | 1 | 1 | 1 |
| 8 | 1 | 1 | 0 | 0 | 0 |
| 9 | 1 | 1 | 1 | 0 | 0 |
| 10 | 0 | 1 | 1 | 0 | 0 |

(B) Reordered matrix[b]

| | Sites | | | | |
|---|---|---|---|---|---|
| Species | A | B | C | D | E |
| 8 | 1 | 1 | 0 | 0 | 0 |
| 2 | 1 | 1 | 1 | 0 | 0 |
| 9 | 1 | 1 | 1 | 0 | 0 |
| 1 | 1 | 1 | 0 | 0 | 1 |
| 6 | 0 | 1 | 0 | 0 | 1 |
| 10 | 0 | 1 | 1 | 0 | 0 |
| 5 | 1 | 0 | 1 | 1 | 0 |
| 3 | 0 | 0 | 1 | 1 | 1 |
| 7 | 0 | 0 | 1 | 1 | 1 |
| 4 | 0 | 0 | 1 | 1 | 0 |

(C) Matrix of Jaccard similarity values for the data in part A or B

| | Sites | | | | |
|---|---|---|---|---|---|
| Site | A | B | C | D | E |
| A | 1.00 | 0.57 | 0.33 | 0.13 | 0.13 |
| B | 0.57 | 1.00 | 0.30 | 0.00 | 0.25 |
| C | 0.33 | 0.30 | 1.00 | 0.57 | 0.22 |
| D | 0.13 | 0.00 | 0.57 | 1.00 | 0.33 |
| E | 0.13 | 0.25 | 0.22 | 0.33 | 1.00 |

[a] The presence of a species in a site is indicated by a 1.

[b] A reordered matrix attempts to group sites that share species and species that share sites.

(8 and 4) might be considered indicators of each group, since they each appear in just one group. With a much larger table showing many more sites and species, such patterns usually become even more apparent.

Using this sort of reordering method, plant community ecologists tried to discern patterns of communities across landscapes. Many ecologists were dissatisfied with this method, however, because it is partially subjective. For example, site C shares more species with site A than site E does, so we might rearrange the columns in the table to place sites A and C next to each other. In addition, we decided that there were two groups of sites, (A, B) and (C, D), even though there is considerable overlap in the species shared by sites in the different groups. To avoid this kind of subjectivity, ecologists have developed a variety of more objective, quantitative techniques.

## Communities can be compared by single factors using univariate techniques

Univariate statistical techniques are used whenever a single type of measurement is made, such as biomass per unit area. We might, for example, determine biomass in ten quadrats within each of six communities. A typical question would be, Do these communities differ in their average biomass per square meter? There are many standard statistical techniques for analyzing such data, such as analysis of variance (ANOVA). A number of statistics textbooks describe these methods in detail (e.g., Sokal and Rohlf 1995; Zar 1999; Scheiner and Gurevitch 2001; Gotelli and Ellison 2004; Lindsey 2004; Fox et al. 2015.) Most often, though, we are interested in comparing communities for which more than one outcome variable has been measured (for instance, the abundance of each of the species present).

## Most community comparisons use multivariate techniques

All multivariate statistical techniques rely on the same basic principles and approach. Each is concerned with comparing how different the members of a group of objects (such as stands or communities) are, based on the values obtained for a set of characteristics measured for all of the objects. The mathematical techniques then arrange those objects in one or more dimensions, based on their differences in the entire set of characteristics taken together.

Univariate statistics compare objects (such as samples) in a single dimension, a line. Multivariate methods require more than one dimension to express differences among objects. Let's begin with a simple univariate (one-dimensional) example. Imagine that we wish to compare three plant communities (A, B, and C) with mean biomasses of 500 g/m$^2$, 725 g/m$^2$, and 625 g/m$^2$. We can put them in order based on their mean biomass (A C B). Another way of accomplishing the same goal is to first determine the difference between each pair of means ($d_{AB}$ = 225, $d_{AC}$ = 125, $d_{BC}$ = 100). Using this information, we determine that A and B are most different, or that the order is (A C B). Note that $d_{AC} + d_{BC} = d_{AB}$ because there is only one way to put these numbers in order from smallest to largest—they fall on a line.

With two dependent variables, the process becomes a bit more complicated, but the basic principles are the same. Let's assume that we also measured mean foliage height for the three communities and found that the values are A = 3.2 m, B = 3.5 m, and C = 5.6 m. Instead of a single number, or scalar, giving the difference between two local communities, we now have both biomass and foliage height, so we need an ordered list of numbers, or vector (see Chapter 8), to give the difference between communities. We now want to consider these two dependent variables for each forest, to determine how similar they are to each other. In this case, we can use that old standby, the Pythagorean theorem ($x^2 + y^2 = z^2$), where $x$ and $y$ are the lengths of the sides of a right triangle and $z$ is the length of the hypotenuse, to measure the distance (or difference) between each community and all of the others. The measure of the difference between each pair of communities is referred to as the **Euclidean distance**. If we let $x$ represent mean biomass and $y$ represent mean foliage height, then for communities A and B, $x = 725 - 500$, $y = 3.5 - 3.2$, and $z = 225.0002 = d_{AB}$ (Figure 12.14). Similarly, $d_{AC} = 125.0230$ and $d_{BC} = 100.0220$. Now the distances do not sum because the points no longer lie on a straight line. Typically, units of measure are not

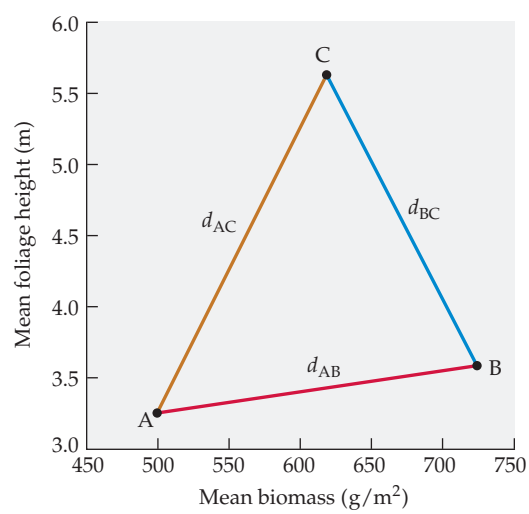



**Figure 12.14** Three plant communities (A, B, and C) measured for biomass and foliage height. Euclidean distances between the points are given by $d_{AB}$, $d_{AC}$, and $d_{BC}$.

used for Euclidean distances, because they are a complex combination of the units of measure for each descriptive variable. This is the same method used for measuring functional divergence (see Figure 12.7).

With only two dimensions, it is easy to graph and visualize our data. But what if we have many variables measured for each community? Information on species presences and abundances constitutes just such a large set of variables quantified for each local community studied. Instead of biomass and foliage height, think of the two axes in Figure 12.14 as representing the abundances of two forest trees, species 1 and species 2. Now, imagine that we add a third axis projecting out of the page, on which the abundance of species 3 is graphed. We might not be so readily able to imagine a fourth axis for the abundance of species 4. Typical data, however, include not just three or four, but tens to hundreds of species. Clearly, we cannot graph or even mentally picture such a multidimensional object. Instead, we resort to various techniques that reduce the problem from $n$ dimensions ($n$ = the number of species) to the two or three dimensions that we can visualize and that ideally summarize the majority of the variation among the communities.

The first step is to devise a measure to quantify how different each community is from all the others. If the abundances of species 1 and species 2 always increase together, for instance, or if species 1 always increases when species 3 decreases, then the variables are related monotonically. If these increases or decreases are also linear (when graphed against one another, the points fall on a straight line), we can use the multivariate equivalent of the Euclidean distance. This measure is the Pearson product-moment correlation, which is the familiar correlation coefficient, $r$. However, species abundances are usually not related in such a simple fashion.

Consider two species with different soil moisture requirements (where soil moisture is the independent variable): species 1 does best in wet soils, while species 2 does best in partly moist soils. Going from very dry soils to partly moist soils, both species would increase in abundance. But going from partly moist soils to wet soils, species 1 would continue to increase in abundance while species 2 would decline in abundance (Figure 12.15). The independent and dependent variables for some species would no longer be related in a monotonic fashion. In addition, many of the variables used to compare communities are not continuous. Data on the presence and absence of species, for example—the information contained in a species list—are dichotomous, consisting of just ones and zeros (see Table 12.10A).

For these types of data, ecologists have devised a number of distance metrics called **similarity measures**. Consider first those that are used for presence/absence data. For a pair of sites, we can define four types of species: those that are present in both sites ($a$), those that are present in the first site but absent in the second ($b$), those that are present in the second site and absent in the first ($c$), and those that are absent in both sites ($d$). This pattern can be represented as follows:

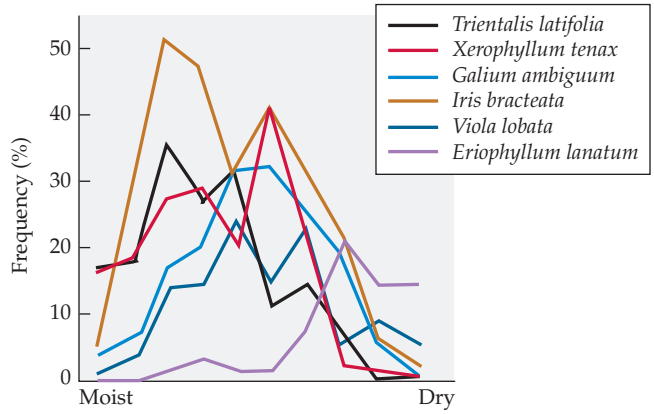



**Figure 12.15** Patterns of species abundance along a moisture gradient in the Siskiyou Mountains, Oregon. As seen in this case, changes in the abundances of species along an environmental gradient are unlikely to be related monotonically. Species abundance was measured as the frequency of quadrats within a site in which a species appeared. (Data from R. H. Whittaker. 1960. *Ecol Monogr* 30: 279–338; after J. H. Brown. 1984. *Am Nat* 124: 255–279.)

| | First site | |
|---|---|---|
| Second site | Present | Absent |
| Present | *a* | *c* |
| Absent | *b* | *d* |

For sites A and B in Table 12.10, $a = 4$, $b = 1$, $c = 2$, and $d = 3$.

The simplest, and oldest, measure of similarity, invented by the French ecologist P. Jaccard, is the Jaccard index (Jaccard 1901). It is the percentage of species contained in two sites that are shared by those sites:

$$S_J = a/(a + b + c)$$

For sites A and B, $S_J = 4/(4 + 1 + 2) = 0.57$.

Over the years, a number of similarity measures have been devised for both presence/absence and abundance data (Table 12.11). While various similarity measures work best in different situations and with different types of data, for presence/absence data the Jaccard index consistently performs best, or close to best, in the widest number of situations. Using the Jaccard index, we can now proceed to the second step, measuring the distance between each of our sites (Table 12.10C).

**TABLE 12.11** Some similarity measures used by plant ecologists

| Index | Formula | Index | Formula |
|---|---|---|---|
| Presence/absence indices | | Abundance indices | |
| Jaccard index | $S_j = \frac{a}{a+b+c}$ | Percentage of similarity | $S_{SD} = \frac{2a}{2a+b+c}$ |
| Sørensen-Dice index | $S_{SM} = \frac{a+d}{a+b+c+d}$ | Asymmetrical percentage of similarity | $S_O = \frac{a}{\sqrt{(a+b)} + \sqrt{(a+c)}}$ |
| Simple matching | $S_{AS} = \frac{b}{2a+b}$ | Minimum similarity | $S_{PS} = \sum \lvert p_{1i} - p_{2i} \rvert$ |
| Ochioi index | $S_{APS} = \frac{\sum(n_{1i} - n_{2i})}{\sum n_{1i} + \sum n_{2i}}$, for $n_{1i} \neq 0$ | Bray-Curtis index | $S_{MS} = \sum \min(p_{1i}, p_{2i})$ |
| Asymmetrical similarity | $S_{BC} = \frac{\sum \min(n_{1i}, n_{2i})}{\sum n_{1i} + \sum n_{2i}}$ | Morisita's index | $S_M = \frac{2\sum(n_{1i} n_{2i})}{(\lambda_1 + \lambda_2) N_1 N_2}$, $\lambda_j = \frac{\sum n_{ji}(n_{ji} - 1)}{N_j(N_j - 1)}$ |

*Source*: R. H. Whittaker. 1972. *Taxon* 21: 213–251; S. Janson and J. Vegelius. 1981. *Oecologia* 49: 371–376; H. Wolda. 1981. *Oecologia* 50: 296–302; M. P. Austin and L. Belbin. 1982. *Aust J Ecol* 7: 75–89; C. E. McCulloch. 1985. *Ecology* 66: 1676–1681.

Definitions of symbols:

$a$ = number of species in both sites
$b$ = number of species in second site only
$c$ = number of species in first site only
$d$ = number of species in neither site
$p_{1i}$ = the proportion of individuals in the $i$th species in sample 1 ($p_{1i} = n_{1i}/N$)
$n_{1i}$ = the number of individuals of species $i$ in sample 1
$N$ = the total number of individuals sampled
min($x$,$y$) = the smaller of the two values

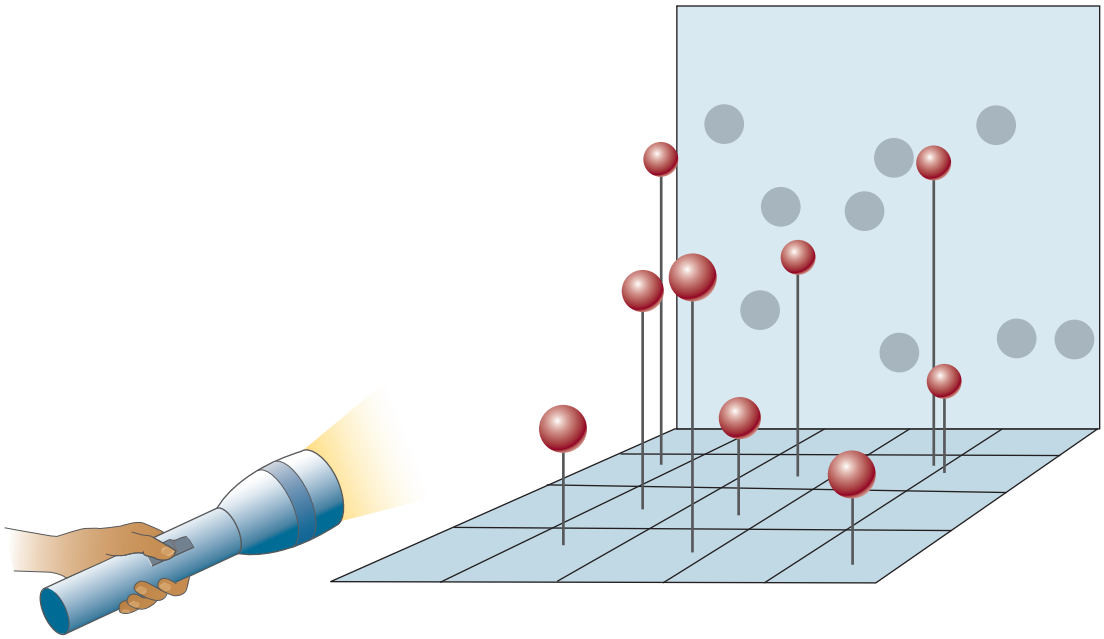
*Note:* These indices differ in the factors they emphasize (e.g., common versus rare species) and their robustness to deviations in assumptions.

## 12.5 Communities Are Distributed across Landscapes

### Ordination is a group of techniques for describing landscape patterns

Ordination is the process of taking information such as that contained in Table 12.10C—points arrayed in an $n$-dimensional space—and mathematically compressing it to fewer dimensions; in other words, ordination is a quantitative redescription of the data. Often the results are shown as a two- or three-dimensional graph. There are many ordination techniques, each of which has had its advocates. They are usually known by acronyms: PO (polar ordination), PCA (principal components analysis), PCoA (principal coordinates analysis), RA (reciprocal averaging), DCA (detrended correspondence analysis), NMDS (nonmetric multidimensional scaling). Each technique uses different mathematical approaches and has different assumptions, limitations, and advantages. While the techniques differ to varying extents in their mathematical details and approaches, they provide broadly similar information. The distinctions between and details of these methods are beyond the scope of this book, but those who wish to explore them further will find the topic covered in depth by Legendre and Legendre (1998).

The process of **dimension reduction**—taking highly multivariate data and collapsing them into a small number of dimensions—can be visualized as follows: Imagine a physical representation of a three-dimensional data set like that shown in Figure 12.16, in which each axis shows the abundance of a species and each ball represents a different site. Now, we shine a flashlight through this model so that its shadow falls on one wall. Each ball (site) is now represented by a point on that wall. We have reduced the data from three dimensions to two. In the process we have lost some information. Two points that were actually far apart in the three-dimensional model may now appear to be close to each other because of the way their shadows fall. We can minimize this problem, though, by carefully choosing the angle at which we shine the flashlight. Most likely the balls are not arrayed randomly throughout the model. Rather, they are probably arrayed in some more compact shape, such as an elongated ellipse. If we shine the flashlight



so that it is at right angles to the long axis of the ellipse, then points that were the farthest apart in the original model are still far apart in the shadow.

Of course, with actual data, the points are arrayed in a much more complicated *n*-dimensional space, and the "flashlight" is a set of mathematical steps that determine the coordinates of those points in the reduced space. Figure 12.17 shows two ordinations of the data from Table 12.10C using two different techniques; notice that both techniques produce the same overall picture of the relationship between the sites. Since ecologists are usually interested in broad patterns, rather than the exact numerical relationships between sites, the choice of distance metric and ordination technique is a matter both of convenience and of the nature of the data being analyzed.

Multivariate statistical methods began to permeate plant community ecology in the mid-1950s, sometimes by importation of techniques being used in other fields and sometimes by independent invention. Perhaps the most influential work was the development of polar ordination by J. Roger Bray in the mid-1950s, while he was a student of John Curtis (Bray 1955; Bray and Curtis 1957). Independently, David Goodall (1954) showed how factor analysis (now referred to as principal components analysis) could be applied to community surveys.

Robert Whittaker developed a general approach called **direct gradient analysis**, in which the ecologist chooses environmental axes and orders vegetation samples (stands) along those axes, examining the resulting patterns. Roger Bray and John Curtis adopted a different approach, called **indirect gradient analysis**, in which the stands are ordered by their similarity in species composition and the environmental factors responsible for the resulting patterns are inferred. The latter approach (a form of ordination) has the advantage of

n reduction. The
e-dimensional
on the back wall

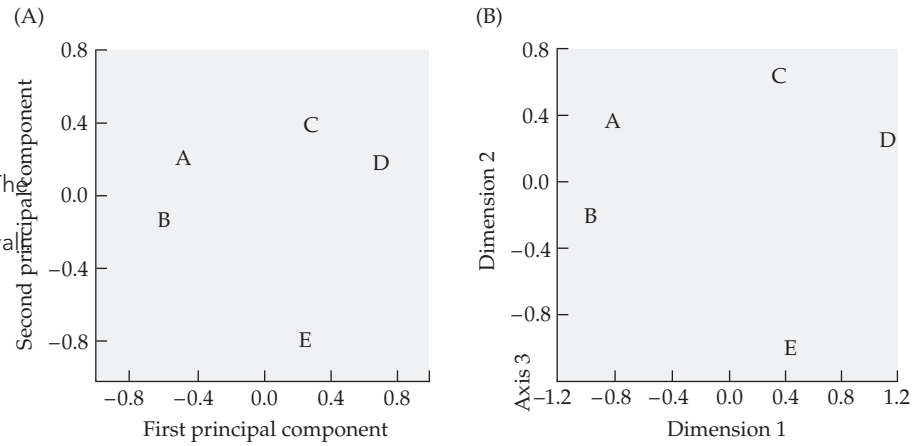



**Figure 12.17** Ordinations of the data in Table 12.1C using two different ordination techniques: (A) principal coordinates analysis (PCoA), and (B) nonmetric multidimensional scaling (NMDS). Notice that the two techniques produce the same broad pattern of relationship between the sites.

being more objective, but it cannot be used to examine patterns of vegetation along particular environmental axes that are known in advance to be of interest.

Ordination and related methods were further refined in the 1960s, 1970s, and 1980s by a number of ecologists, most notably Whittaker's students and collaborators in the United States (e.g., Gauch and Whittaker 1972), Michael Austin in Australia (e.g., Austin 1977), and Cajo ter Braak in the Netherlands (e.g., ter Braak 1986). Over the same period, the development of multivariate techniques in systematics (within the area known then as numerical taxonomy, Sneath and Sokal 1973) further influenced their use in ecology. Advances in computer hardware and software at this same time were particularly important in the spread of these techniques. During the 1960s and 1970s, using these methods required writing complex programs, making stacks of punch cards, and gaining access to a mainframe computer.

Today, these techniques are still commonly used, especially for pattern detection and hypothesis generation (Austin 2013). A series of communities or stands might be sampled to document relationships involving a particular ecological phenomenon—for example, nitrogen deposition, invasion by exotic species, or damage by a pathogen—with the object of relating the phenomenon to species compositions and environmental variables across communities. The patterns that are revealed can form the basis of experiments to test hypotheses about the mechanisms causing the patterns. These methods can also be used for purposes quite different from their original ones. For example, Rosemary Booth and Philip Grime (2003) used principal components analysis to describe the change in species composition of 36 experimental communities over 5 years in a field site in the United Kingdom. They used ordination to measure change in community composition over time, rather than for the more conventional purpose of describing natural communities in a landscape.

### Patterns of species difference among communities are caused by variation in the environment

Ordination can describe patterns of species distribution among communities in a landscape. But what are the causes of those patterns? For plant communities, the primary causes of differences in species composition are climatic, topographic, and edaphic (soil) factors—the physical factors that determine plant growing conditions (see Part I). A number of techniques exist for determining how important each of these factors is in a given set of communities. All of these techniques use the same basic strategy: looking for correlations between species distributions and environmental variables.

Consider a simple problem, that of determining the causes of variation in the abundance of a particular species across a landscape. One strategy for attacking this problem is to sample a number of sites across the landscape. At each site we measure the abundance of the species as well as a number of environmental variables, such as temperature, precipitation, elevation, slope, aspect, soil texture, and various soil nutrients. (Ideally, the climatic data would be based on long-term averages.) Then we calculate the correlation of each environmental variable with abundance. This procedure contains two unstated assumptions: (1) that the pattern of species distribution is based on a long-term equilibrium with the environment, and that the current pattern is close to that equilibrium, and (2) that we have measured the correct environmental variables.

Of course, correlation does not demonstrate causation, but a strong correlation provides one piece of supporting evidence for a possible cause of the observed pattern. The method just described has a problem, however: the environmental variables are themselves correlated with one another. Sites at higher elevations, for example, are likely to have both lower temperatures and greater wind speeds. To account for these correlations, ecologists use various statistical methods. One of these is a technique called **multiple regression**. This technique determines the relationship of the abundance of our species to each environmental variable while correcting for correlations among the environmental variables themselves. A limitation of this technique is that it cannot correct for very large correlations. For example, if the abundance of our species decreased with decreasing temperature while wind speed also increased with decreasing temperature, we could not determine whether the cause of decreasing abundance was temperature or wind speed. Our analysis would have narrowed the possible factors responsible for decreasing abundance, although the possibility would remain that the actual cause was some unmeasured third variable that is also correlated with temperature or wind speed, such as elevation.

Another, very different approach is performing manipulative experiments (such as growing the plant in growth chambers at different temperatures) to further narrow the possible causes and to test specific hypotheses (see Chapter 1). While such experiments are an important source of information, they have their own limitations. Because they are often done in highly artificial environments, such as a growth chamber or greenhouse, they cannot account for the myriad of abiotic and biotic factors and their interactions in nature. Competitive interactions, for example, might differ at high and low temperatures. Manipulative experiments done in a field setting are technically more challenging but can provide a more realistic test of the importance of an environmental factor. A very powerful approach is to combine several different kinds of experiments and lines of evidence; if the results of using many

different types of evidence agree, the conclusion is far stronger (Scheiner 2004).

The methods just described examine single species and one or a few environmental variables. How can we combine information about the species composition of whole communities and environmental variables? We do this using methods related to multiple regression. Instead of using species abundance as the dependent variable, however, we use position along an ordination axis. We can see how this works with data from a study of *Pinus contorta* (lodgepole pine, Pinaceae) forests in the Canadian Rockies by George La Roi and Roger Hnatiuk (1980). Lodgepole pine is the most widely distributed tree in western North America (Figure 12.18A). In Alberta, Canada, where this study took place, it is the dominant species following fire in midelevation forests (1000–2000 m). In this study, 63 stands were sampled in Banff and Jasper National Parks; measurements included the number and sizes of all trees, cover of all understory species, soil chemistry, soil temperature, soil moisture, elevation, slope, and aspect. Figure 12.18B shows these stands arrayed along two environmental axes, elevation and moisture. Because each axis consists of a single environmental variable, this graph is a form of direct gradient analysis.

We can go further in order to account for correlations among the environmental variables by using **canonical correspondence analysis (CCA)** (ter Braak 1986). In CCA, each axis consists of combinations of multiple environmental variables or the vegetation data. In effect, CCA is a hybrid between direct and indirect gradient analysis. Two ordinations are carried out, one on the species data and one on the environmental data, and then the correlations between the axes of the two ordination sets are determined. If we are lucky, the first one or two axes of the environmental data will be highly correlated with the first one or two axes of the vegetation data. An analysis of plants growing on dikes in the Netherlands provides an example of this type of analysis (Figure 12.19). As with ordination, various distance metrics, dimension reduction techniques, and correlation procedures can be used (Ludwig and Reynolds 1988).

## What types of data are used?

We have described analyses that can use either presence/absence data or abundance data. Which type of data is most appropriate? The answer depends on the questions being addressed relative to the scale of variation. Different kinds of variation are expected to occur at different scales. By *scale* we do not mean geographic extent; rather, we mean ecological scale—how different the communities are ecologically. A set of neighboring old fields all abandoned within the past 5 years would represent small-scale ecological differences. All of the fields would have similar soil conditions and species compositions. Differences among the fields would most likely show up as differences in species abundances rather than differences in species presences. In

(A)

Courtesy of John Worrall

(B)

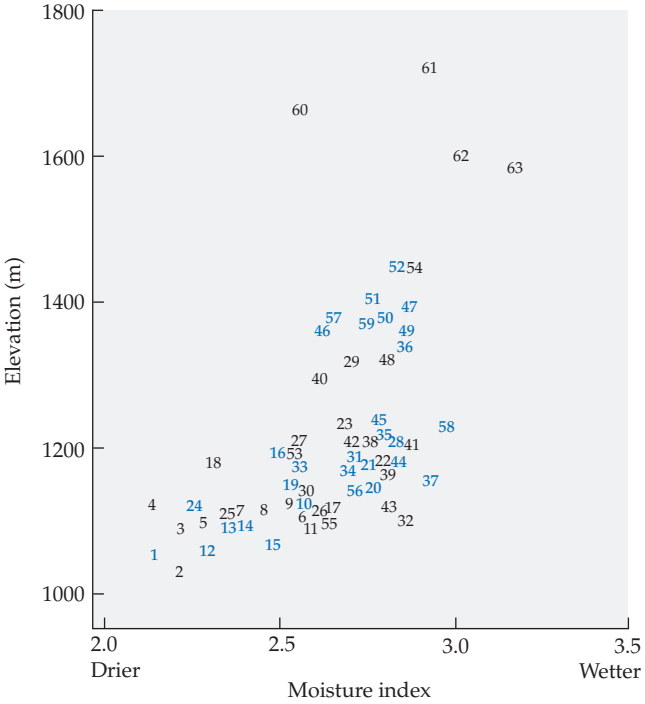



**Figure 12.18** (A) Stands of *Pinus contorta* (lodgepole pine, Pinaceae) dominate the landscape in the middle elevations of the Canadian Rockies. The cones open in response to fire, and the result is stands like this one near Allison Pass. (B) Direct gradient analysis of *P. contorta* stands in the Canadian Rockies. The stands are arrayed along two gradients: moisture and elevation. Blue numbers indicate stands in Banff National Park. Black numbers indicate stands in Jasper National Park. (After G. H. La Roi and R. J. Hnatiuk. 1980. *Ecol Monogr* 50: 1–29.)

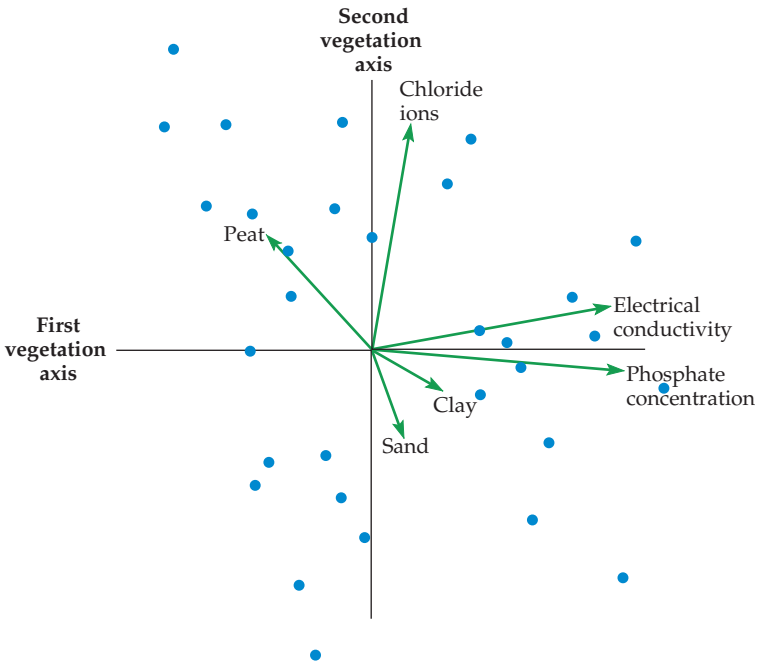



**Figure 12.19** Canonical correspondence analysis of plants growing on dikes in the Netherlands. The points show the average position of each species along the vegetation axes. The first vegetation axis is highly correlated with electrical conductivity ($r$ = 0.83) and phosphate concentration ($r$ = 0.86), while the second vegetation axis is correlated with chloride concentration ($r$ = 0.86) and to a lesser extent with the amount of peat ($r$ = 0.49) and sand ($r$ = –0.40) in the soil. These correlations are indicated by how closely the environmental variables (arrows) line up with the vegetation axes. (After C. J. F. ter Braak. 1986. *Ecology* 67: 1167–1179.)

contrast, if those fields were included in an analysis with surrounding forest patches, or with fields abandoned decades ago, then we would have a much larger ecological scale. Now species presences might capture the differences among communities. Geographic extent is not synonymous with ecological scale, but the larger the geographic extent, the more likely it is that a wider range of ecological conditions will be captured in the analysis.

You might expect an ecologist to want to collect all possible information on all species. However, measuring abundances for all species can run into practical limitations. In deserts, many annual herbaceous species are present as aboveground plants for only short periods after it rains, and in some years they might not appear at all. A study of such a landscape might be forced to focus only on woody species—trees, shrubs, and cacti. In tropical forests, where species diversity is very high and many epiphytes and lianas are concentrated very high in the canopies of the trees, most studies focus on just the tree species. In many species, distinguishing individual plants can be difficult, and it may be necessary instead to use cover or biomass estimates.

At larger scales, presence/absence data may be more informative than abundance data because of the signal-to-noise problem. At large scales, we are usually trying to find the one or few causes of differences among very different types of communities. But abundances can vary for a number of reasons having to do with random demographic factors (see Chapter 8 and Chapter 13). While this variation may be of interest in relation to other questions, in this instance it acts as "noise" (random variation) in the analysis, obscuring the main patterns. By using presences and absences instead of abundances, we reduce this noise. However, if we use presence/absence data, it is important that we include as many species as possible, especially rare species. The distribution of each species contains information about the environment and its effects on the plant community. The more species we include in the analysis, the more information we have. There is a redundancy to this information, so the more species we include, the more confident we can be about our conclusions. A good rule of thumb is that in any ordination, the number of species included should be at least three or four times the number of sites.

## Classification is an alternative approach to describing communities in a landscape

One use of landscape analysis is the classification of communities into larger groupings (see Table 12.1). This classification can then be used for landscape management purposes. For example, if we wished to preserve particular types of communities, our classification scheme could help us choose which places to preserve to meet this goal. Classification techniques are complementary to those of ordination. In effect, what we are doing is taking the pattern revealed by the ordination and drawing circles around subsets of the points. For example, in Figure 12.17, we might put sites A and B into one group, sites C and D into another group, and site E into a third group. We could even construct larger groups from these smaller ones. Just because we have created groupings does not mean that the communities form discrete types; our different groups might gradually blend into one another. Yet, for planning purposes, defining discrete community types would still be useful.

A variety of methods can be used to create groups. The two major types of classification techniques are monothetic-divisive and polythetic-agglomerative. A **monothetic-divisive analysis** begins by considering all of the sites and dividing them into two groups depending on the presence or absence of a single key species.

Each of the new groups is further divided based on new key species contained within each. The process is continued until we have reached the desired number of groups—a set of groups that seems natural and useful—or until division is no longer possible. Thus, *monothetic* refers to the use of a single species as the criterion at each step, and *divisive* refers to the dividing process.

A **polythetic-agglomerative analysis** works in the opposite direction. Using some distance metric (such as the Jaccard index), one groups together the two most similar sites. Then the next two most similar sites are grouped together, and so on. The grouping may involve two individual sites, a single site and a previously created group, or two groups. The process stops when all sites have been grouped together. Thus, *polythetic* refers to the use of multiple species in the distance metric, and *agglomerative* refers to the joining process. Various distance metrics and joining algorithms can be used. The result is a tree diagram (Figure 12.20). As with the monothetic-divisive approach, this diagram is used to devise some natural and useful classification scheme (Figure 12.21).

A useful adjunct to classification is the designation of **indicator species**. An ideal indicator species is found in all communities of a given type and not in any other community type. The use of indicator species makes classification of communities much easier. As with other procedures, there are various methods for choosing indicator species (Dufrêne and Legendre 1997). An analysis of a landscape might proceed along the following lines: First, one would choose a number of representative sites in the landscape and do a complete species survey of each. Second, one would classify the sites and identify indicator species. Third, one would survey the rest of the landscape, looking just for the indicator species. By looking for only a small number of species, one could survey many more sites with the same amount of time and effort that would be spent on complete species surveys. One note of caution: the term *indicator species* is sometimes used in ecology to refer to species that are thought to be characteristic of a particular kind of environment; for example, *Typha latifolia* (Typhaceae, broadleaf cattail) grow only in wetlands.

This variety of methods leads to three different approaches to classifying landscapes. The first approach relies on using all species. Polythetic-agglomerative methods are typical of this approach and are a hallmark of the floristic-sociological school of central Europe. The second approach relies on using dominant species or indicator species. Monothetic-divisive methods are typical of this approach and have been used primarily by Russian plant ecologists. The third approach relies on the physiognomy of the dominant species of the vegetation rather than species identity. In this approach, all broad-leaved deciduous forests, for example, would be classified together, even if they shared no species in common. It is also possible to combine these approaches. The current U.S. National Vegetation Classification system for classification of North American vegetation (see Table 12.1) combines aspects of the second and third approaches.

Which approach is most appropriate depends on the scale of classification and its purpose. At very broad continental or global scales, physiognomy-based classifications are the most useful because we are often looking for commonalities among communities that transcend species identities. Our discussion of global biomes in Chapter 18 uses this type of system. In contrast, at local scales, we are often

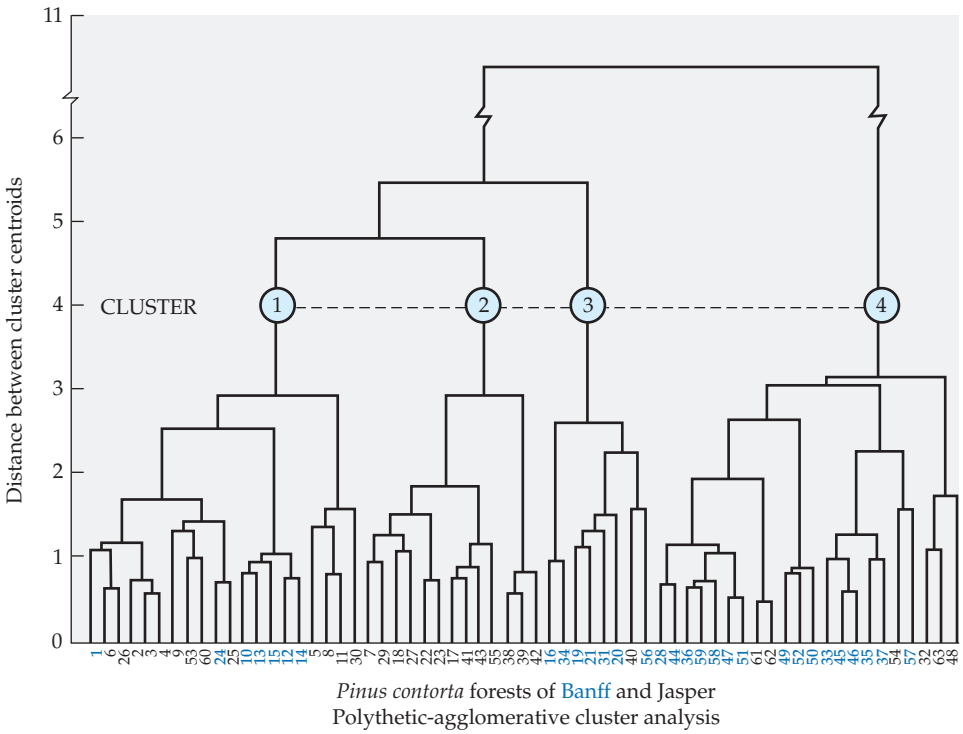



**Figure 12.20** A tree diagram developed by a cluster analysis using a polythetic-agglomerative method for the *Pinus contorta* stands in Figure 12.18. (After G. H. La Roi and R. J. Hnatiuk. 1980. *Ecol Monogr* 50: 1–29.)

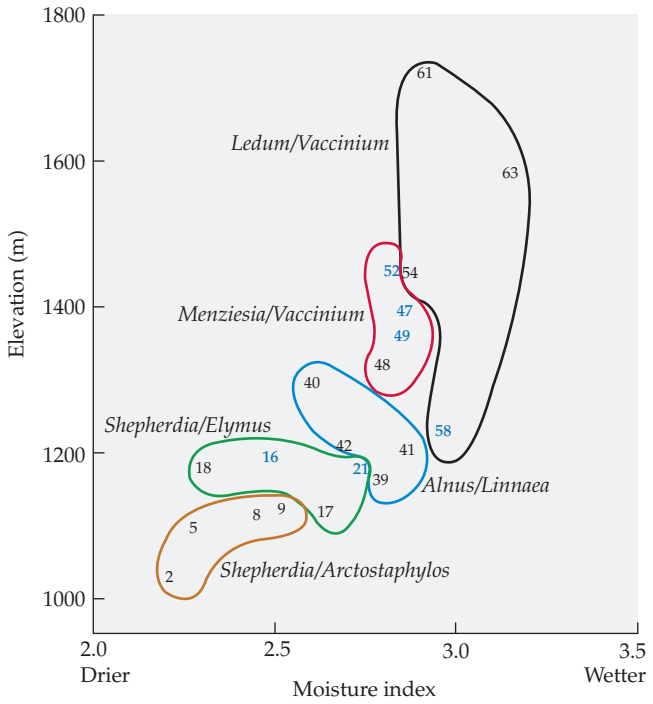



**Figure 12.21** Classification of the *Pinus contorta* stands in Jasper National Park graphed in Figure 12.18, using a combination of techniques, including the cluster analysis shown in Figure 12.20 and indicator species analysis. The clusters are named by the dominant species. The number of clusters differs from that in Figure 12.20 based on the additional analyses. (After G. H. La Roi and R. J. Hnatiuk. 1980. *Ecol Monogr* 50: 1–29.)

interested in how closely related species sort themselves into communities, so either of the other approaches would be useful. We may, on the other hand, wish to emphasize continuity among samples. If that is the case, ordination is most appropriate.

Classification using remote sensing proceeds by grouping pixels with similar spectral properties into a single category. There are two basic approaches: **Unsupervised classification** separates pixels into a user-defined number of classes based on statistical similarities between the pixels. This method is similar to the classification methods described previously. In this case, however, instead of each pixel (here, representing a spot on the Earth's surface) being characterized by the presence or abundance of a set of species, each pixel is characterized by the presence or intensity of a set of light waves. In contrast, **supervised classification** requires the classes to be defined explicitly by the user. Specific locations in the image (training regions) are characterized using ground-based knowledge. For instance, one begins by making species lists for a series of sites, which are then classified into groups using the techniques described above. Those sites are also mapped using the Global Positioning System (GPS) and located on the remotely sensed image. These sites then serve as the training regions. It is best to have multiple training regions per group in order to account for the inevitable variation between sites. Following either type of classification, accuracy is judged using **ground-truthing** operations, in which randomly selected sites are surveyed to see whether their actual species composition matches that predicted by the remote sensing classification.

## Summary

- A community is a group of organisms that coexist in space and time and potentially interact with one another.
- Ecologists have long debated the relative importance of the processes responsible for community structure. One extreme view is that communities are highly predictable entities in which species are tightly interlinked. An opposing view is that communities are chance groupings of species that are shaped by abiotic factors and historical contingency.
- The current consensus lies between these views. Deterministic and chance events and biotic and abiotic processes are all considered important.
- Biodiversity metrics are built from four basic types of information—identity, abundance, phylogeny, and function—each with particular properties that can be quantified in more than one fashion and combined in a variety of ways to produce many different kinds of metrics.
- The simplest community measure is species richness, the number of species in a community. One method for estimating species richness is the use of a species richness relationship.
- Other indices of species diversity take into account both the number of species and the relative abundance of each species, or diversity metrics can be based on phylogenetic relationships or ecological function.
- Species presence and relative abundances can be sampled using quadrats, transects, and other methods to estimate the numbers of individuals, frequency, cover, or biomass.

- Physiognomy, the general form of the vegetation, is another community property that can be measured. These measures are used to describe community patterns in time and space. Describing patterns is a first step in determining which processes are responsible for shaping the plant communities that we see around us.
- The information about similarities among communities can be further utilized to examine overall patterns among groups of communities. Ordination is a method for arranging these sets in continuous space. These arrangements can then be used to look for relationships between vegetation patterns and environmental variables.
- Another approach to looking for relationships in species composition among local communities is classification. A number of classification methods are available to sort sets of communities into groups; these methods are particularly useful for management and conservation purposes.

Courtesy of S. Scheiner

13

# Community Dynamics and Succession

The living world is in a constant state of flux. In previous chapters, we looked at processes of change at the level of the individual (Part I) and the population (Part II). As we have seen in Part III, change also occurs at the level of the community. Communities change as the species present change and as the populations change in number, age structure, or size structure, resulting in changes in physiognomy and ecosystem function. In this chapter, we explore patterns of community change and their causes. We also explore a variety of other aspects of communities: their continuity in time, their predictability, and whether they reach an equilibrium.

The process of succession has been central to plant ecologists' study of plant communities. Change can be of two types: cyclic change, such as that which occurs over the course of a year or in response to weather fluctuations, and longer-term community dynamics that may continue in a single direction. The latter is the focus of this chapter. Longer-term changes due to human and natural alterations in the climate are explored in Chapter 17 and Chapter 19. **Succession** is change in structure or composition of a group of organisms of different species at a site through time. The term denotes change over periods longer than a single season, although very long-term trends, such as those due to climate change, are not regarded as part of succession.

Succession begins when a disturbance—an event that removes part or all of a community—is followed by colonization of or regrowth in the disturbed site by plants. The use of the terms *disturbance* and *succession* can be misleading in that both encompass a wide variety of processes and mechanisms. We explore that variation in this chapter.

Above: The aftermath of a light surface fire in a Eucalyptus woodland in northern Australia.

As an example of one type of disturbance and resulting succession, consider the classic example of change in the plant community following the abandonment of an agricultural field in North Carolina, as worked out by Henry Oosting (1942) and his student Catherine Keever (1950). In the first year, a variety of species were likely to colonize the field. The dominant species were typically short-lived ones—annuals and biennials such as *Conyza canadensis* (horseweed, Asteraceae) and *Gnaphalium purpureum* (purple cudweed, Asteraceae). Over the next few years, these species tended to be replaced by herbaceous perennials such as *Andropogon virginicus* (broomsedge, Poaceae) and *Aster ericoides* (white aster, Asteraceae). By 10 years after abandonment, shrubs such as *Rubus* spp. (blackberries and raspberries, Rosaceae) and *Rhus* spp. (sumacs, Anacardiaceae) and trees such as *Pinus taeda* (loblolly pine, Pinaceae) began to dominate the field, although they may have been present as seedlings and saplings early in the process. Eventually, after 150 to 200 years, the pines were replaced by hardwood tree species—*Quercus rubra* (red oak, Fagaceae) and *Carya* spp. (hickory, Juglandaceae). If no subsequent large-scale disturbance occurred, these hardwood trees remained the dominant species. While the species that dominate late in the successional process are all natives, that is not true of the early colonizers, which can consist of a mix of native species, species that were introduced to North America centuries ago, and recent North American invaders (see Chapter 14). In some cases, these invasive species have changed the nature of the community as invasive woody vines have become much more common.

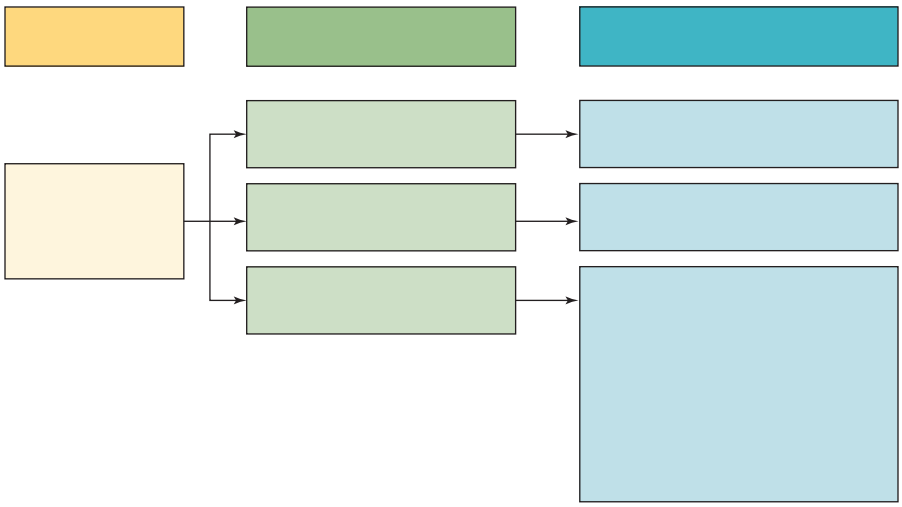
Discussions about the causes and nature of the successional process are part of the debate about the nature of communities (see Chapter 12). One extreme view is that succession is an orderly and predictable process that is a result of emergent community properties. An alternative view is that succession is a rather unpredictable series of events that result from interactions between individuals and the abiotic environment. As we will see, succession involves a number of different patterns, mechanisms, and causes. Ecologists have amassed a great deal of information about successional processes, consisting of a complex set of interlocking processes giving us a framework for understanding succession (Figure 13.1). This framework can be applied to particular systems, providing descriptions of the patterns of succession and the processes responsible.

## 13.1 Conflicting Theories Have Attempted to Explain the Mechanisms of Succession

### Are communities dynamic mosaics or regulated by predictable processes?

Several issues form the core of discussion about the processes responsible for succession. One group of questions flows directly from the sharply divergent views of communities advocated by Frederic Clements and Henry A. Gleason (see Chapter 12). First, what processes are responsible for successional changes? Do emergent properties and interactions among species play a major role? Or is succession the mere unfolding of the life history of each species independently of the others? If species interactions are important, are they mainly in the form of mutualisms, or is competition among species more important?

Second, how predictable is succession? If two adjacent areas, or the same area at two different times, were to undergo succession, would the same changes occur,



leading to the same end point? A **climax** is the hypothetical deterministic end point of a successional sequence. Do plant communities undergoing succession usually reach a stable climax state? To the extent that they do not, is this mainly because extrinsic disturbances (such as storms) prevent it, or because of internal dynamics (such as population cycles) inherent in the communities?

Third, are most plant communities static and unchanging, or are they in a state of dynamic equilibrium, or are they always in a fundamental state of flux? A stable equilibrium implies that they would return to the same structure and composition after small to moderate perturbations, remaining in that state until a new disturbance occurs. Clements's view was that communities are unchanging. Dynamic equilibrium implies very different processes but in some ways is a modern substitute for Clements's perspective.

The Clements-Gleason dichotomy, however, does not account for many of the other questions we consider relevant to understanding succession today. For example, it seems obvious that one might ask about the role of competition (see Chapter 10) in successional processes. This question, though, does not fit neatly into the Clements-Gleason dichotomy, as neither of these scientists ascribed an important role to competition in succession. This is especially odd, given that Clements wrote extensively about competition, but (given his superorganism viewpoint) he thought of it in a rather different sense than modern ecologists do. Similarly, neither Clements nor Gleason appears to have thought of herbivory as an important factor affecting plant community structure and function, though they were certainly well aware of it. Those original ideas have evolved (Box 13A). Ideas about successional change have expanded to the much broader context of tropical and arid ecosystems, moving beyond the processes in temperate and boreal forests and grasslands that were initially the focus of thinking about succession. Today we consider

## BOX 13A History of the Development of Modern Succession Theory

The initial ideas about succession as a predictable process leading to a static climax were developed in the 1890s and early 1900s, largely through the work of Henry C. Cowles (1899, 1911) and Clements (1916). These ideas were a major force in plant ecology in North America (and to some extent in the entire English-speaking world) for the next half century or more. The move away from these ideas took several directions.

R. T. Fisher studied old-growth forests on Pisgah Mountain in New Hampshire, including both hardwood and coniferous forests, over several decades, beginning in 1910. The dynamic and patch-like character of these ancient forests was cogently summed up by Fisher in a preface to a study by his successors (Cline and Spurr 1942, as cited by Dunwiddie et al. 1996):

> *The primeval forests, then, did not consist of stagnant stands of immense trees stretching with little change in composition over vast areas. Large trees were common, it is true, and limited areas did support climax stands (or preclimax stands), but the majority of the stands were in a state of flux resulting from the dynamic action of wind, fire and other forces of nature. The various successional stages brought about, coupled with the effects of elevation, aspect, and other factors of site, made the virgin forest highly variable in composition, density and form.*

This view was confirmed on September 22, 1938, when a major hurricane uprooted many of the trees in this forest, subsequently changing the vegetation, and ongoing studies suggest that such infrequent, large-scale disturbances are an important factor in forest community dynamics (Foster 1988). John Weaver, a contemporary and sometimes collaborator of Clements, worked in prairies of the central United States. While he accepted Clements's views of climax in prairie vegetation, his studies during the Dust Bowl of the 1930s led him to emphasize also the importance of competition among plants (Weaver 1968).

An influential perspective on community dynamics was introduced by Alexander S. Watt (1947). In his view, a plant community is composed of a mosaic of patches, and the patches are dynamically related to one another (see Chapter 10). Community structure results from a balance between factors that are predictable and those that are unpredictable (what we now might view as factors that drive the community toward equilibrium and disturbances that disrupt that tendency). Watt was careful to distinguish this view—that the community is a dynamic mosaic of patches—from that of a successional sequence in which a homogeneous community tends toward a final, determinate end point. He emphasized the ubiquitous patchiness of natural plant communities and stated that the persistence of patches is a fundamental characteristic of plant communities.

By the 1980s, influenced in particular by work in tropical forests and by broader trends in ecological thinking, many ecologists had begun to view plant communities as mosaics of patches rather than as homogeneous entities. In other words, they ceased trying to average over the entire community and refocused

(*Continued*)

**BOX 13A** *(continued)*

on the smaller scale of gaps and patches (Pickett and White 1985; Forman and Godron 1986). At the same time, plant ecologists abandoned the search for idealized, static types—such as climax communities—to embrace concepts of variability and dynamic processes as essential to the true nature of communities. In the 1980s, this changing perspective intersected with a very different tradition from Europe. There, the floristic-sociological school was part of the development of landscape ecology. The European floristic-sociological approach focused on measuring static patterns. When wed to the thinking of North American ecologists interested in patch dynamics, the result was a dynamic outlook that helped put successional studies into a wider landscape framework. We discuss these issues in more depth in Chapter 15.

both change in species composition and changes in the three-dimensional structure of the community (Pickett et al. 2009). Ecologists in some research areas use the term *vegetation dynamics* instead of *succession*; the term is frequently used by ecologists working in rangelands (grazed grasslands) and often emphasizes the role of human management and disturbance by people and their domestic grazing animals.

This century-long accumulation of knowledge has been synthesized into a current well-articulated theory of succession (Table 13.1). This theory can be contrasted with other theories about the processes that structure communities that focus on non-directional dynamics (see Chapter 15). Successional change is driven by factors that are external to a community, especially disturbances, and by the interactions among the species within a community (Figure 13.2). Those species interactions create a feedback loop such that changes in the number, size, and condition of a population further alter the physical environment and change the pattern of interactions. The end result is a successional trajectory (change over time) that is somewhat, but not completely, predictable. In the rest of this chapter we explore this process in detail.

## Scientific understanding can be influenced by methodology

The ability of scientists to understand nature is affected by the methods they use to study it. This is nowhere more apparent than in the study of succession. Changes in methods for studying succession have affected our understanding of the process; conversely, changes in ecologists' understanding of succession have also fostered changes in methods.

A central difficulty in studying succession is that it is often a very slow process—it can take decades or centuries. Ideally, one would like to be able to observe the entire successional process in a single community, but this is generally not possible. One traditional way to get around this limitation is by studying a **chronosequence**, a set of communities of different ages since disturbance. Users of the chronosequence approach assume that the differences

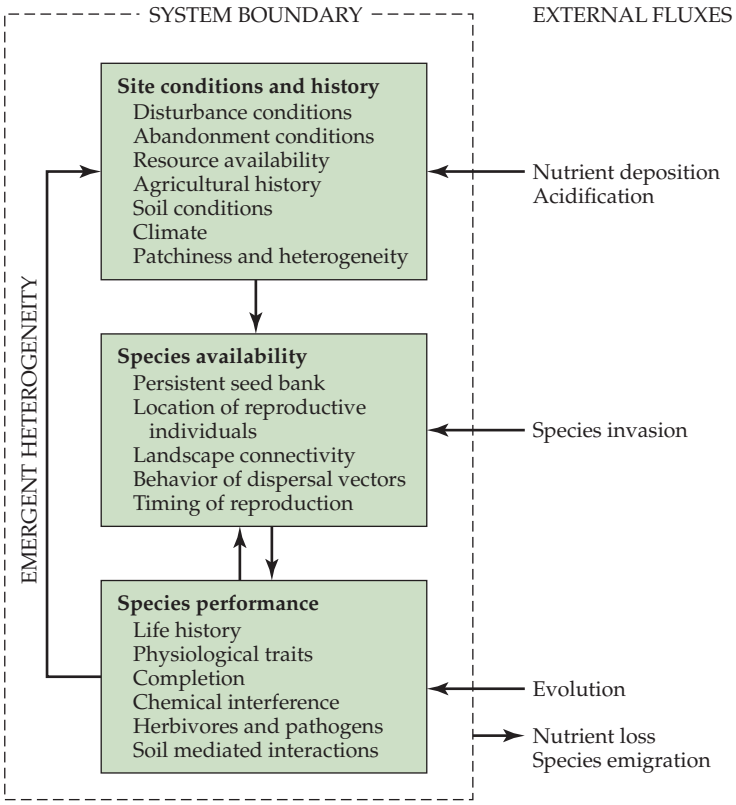



**Figure 13.2** An idealized filter model of successional processes, with reference to components of the causal hierarchy in Figure 13.1. The model focuses on how individual organisms come to occupy a site, and how interactions affect community composition and structure. Because this model is an explanation of interactions among organisms in a local environmental context, it mainly ignores the movement of species or nutrients from one community to another. (Modified from S. T. A. Pickett et al. 2011. In *The Theory of Ecology*, S. M. Scheiner and M. R. Willig [Eds.], pp. 185–218. University of Chicago Press: Chicago.)

**TABLE 13.1** The Propositions of the Theory of Succession

1. Succession is driven by the interactions of organisms with each other and the physical environment.
2. Successional patterns in communities result from the interactions of individuals.
3. Multiple trophic levels participate in the driving interactions.
4. Succession results from processes of disturbance, differential availability of species to a site, and differential performance of species within a site.
5. Successional causes can operate on any time scale.
6. The possible outcomes of interaction between individuals are tolerance (no effective interaction), inhibition, or facilitation.
7. The species composition of a site tends to equilibrate with the environment of that site.
8. The specific form of a successional trajectory is contingent on starting conditions, and the stochasticity of invasion and controls on species interactions.
9. Succession produces temporal gradients of the physical environment, biotic communities, and the interaction of the two.

*Source*: S. T. A. Pickett et al. 2011. In *The Theory of Ecology*, S. M. Scheiner and M. R. Willig (Eds.), pp. 185–218. University of Chicago Press: Chicago.

among these communities represent what would occur within a single community over time. Henry C. Cowles (1899) pioneered and popularized this approach, which was based on his studies of patterns of primary succession in the dunes and forests along the southern shore of Lake Michigan (Figure 13.3). Cowles hypothesized that the dunes were initially colonized and stabilized by plants with extensive root systems, such as the grasses *Ammophila breviligulata* (beach grass) and *Elymus canadensis* (wild rye) and small woody plants such as *Prunus pumila* (sand cherry, Rosaceae). Once the dune is stabilized by these plants, Cowles argued, larger shrubs, such as *Salix glaucophylloides* (dune willow, Salicaceae) and *Cornus stolonifera* (red osier, Cornaceae), can colonize it. Eventually, trees such as *Tilia americana* (basswood, Tiliaceae) enter the community, converting it from a thicket to a forest. Finally, under Cowles's scenario, the forest becomes dominated by *Acer saccharum* (sugar maple, Sapindaceae) and *Fagus grandifolia* (American beech, Fagaceae).

Courtesy of the University of Chicago Archive

Henry Chandler Cowles

The chronosequence approach rests on three important assumptions. First, it assumes that successional processes are highly predictable, that is, that similar communities will all undergo the same successional sequence. Second, it assumes that the climate and other aspects of the environment, as well as the **species pool** (the species available for colonization), will remain the same during the entire chronosequence. We know, however, that this is not always true; for example, new species may migrate into a region, species combinations may change during major periods of climate change (see Chapter 17), or species may go locally extinct. Especially in recent years, invasive species have changed successional trajectories (see Chapter 14). Finally, the chronosequence approach assumes that communities in the same general region are similar enough to be considered part of the same successional sequence. This last assumption was called into question during the 1950s, undermining the use of chronosequences—an example of how conceptual shifts drive changes in methods. A reexamination of several classic chronosequence studies by Edward Johnson and Kiyoko Miyanishi (2008), including those of

American Environmental Photographs Collection, AEP-MIS136, Department of Special Collections, University of Chicago Library

**Figure 13.3** A foredune along the shore of Lake Michigan at Big Point Sable, Ludington, Michigan, taken on September 13, 1916. On the left is *Ammophila breviligulata* (beach grass, Poaceae), and on the right is *Prunus pumila* (sand cherry, Rosaceae). The root systems of these species stabilize the open dunes.

Cowles and Oosting described above, showed that the actual successional trajectories do not follow the orderly sequences that were originally hypothesized.

Two new research approaches were used to directly test the first assumption, that the outcome of succession is highly predictable—an example of how methods sometimes drive conceptual advances. The first of these approaches was the use of long-term (or longitudinal) studies, which provided data that were not available to the earliest ecologists. Beginning in the 1920s and 1930s, ecologists established permanent study sites that were monitored for decades or longer. Ecologists have now directly observed long successional sequences and compared them with those predicted by the chronosequence method. In a famous example from North Carolina, Oosting (1942), and Keever (1950), used the chronosequence method to predict that pine forests would eventually yield to hardwood-dominated communities. This hypothesis was tested using long-term study plots established in the Duke Forest in 1934 and continuously monitored ever since. The methods of observation included tagging every tree with a trunk greater than 1 cm DBH. Based on these studies, Norman Christensen and Robert Peet (1981) determined that Oosting's hypothesized successional sequence was correct. A drawback to long-term studies is that time is still a limiting factor. A single researcher can examine succession over a period of decades, but not centuries. In addition, this method can be used only in situations in which sites can be established and monitored without fear of human disturbance, and in which there is a dedicated effort to maintain the study over a long period as investigators, funding, and even institutions change.

The second shift in methodology was an increase in the 1970s in the use of manipulative experiments to study community processes. Until then, plant community ecology was mainly an observational discipline, involving the collection and analysis of extensive quantitative data. Now ecologists began to test the explanations proposed for successional processes by applying different experimental treatments to communities and studying their responses. Three of the most extensive of these studies in the United States have been conducted at sites that are part of the Long-Term Ecological Research (LTER) Network (see Box 12C). In Chapter 1 we described one of those sites: the Konza Prairie Research Natural Area in Kansas. A second is Hubbard Brook, New Hampshire, where Frank Bormann and Gene Likens (1979) used whole-watershed manipulations to study succession. The third is the Cedar Creek Natural History Area in Minnesota, home to an ongoing study of prairie succession established by David Tilman. The experimental manipulations used at these sites include fertilizing plots with different levels of nitrogen, creating plots of different sizes, and establishing plots with different sets of species (see Chapter 5). Time is again an important limiting factor in these kinds of experiments. They are most useful for addressing questions over time spans from a few years to over a decade in unusual circumstances. That is why they have mainly been used in communities such as prairies and old fields, where successional change initially proceeds at a relatively fast pace.

Another method that has been used extensively over the years to reconstruct the history of particular communities is **dendrochronology**, the study of tree rings. This method is particularly useful for picking up the study of succession where the other two methods leave off: over a period of centuries. Many trees (especially in the temperate zone) produce annual growth rings (Figure 13.4). A tree's age can be determined by taking a core from it and counting the rings. Knowing the ages of the trees in a forest allows one to determine when various species became established at that site. Other characteristics of tree rings, such as their size, appearance, chemical composition, and isotopic composition, can provide additional information about the climate, soil, and growing conditions in times past. We discuss other methods of determining long-term, nonsuccessional change in plant communities in Chapter 17.

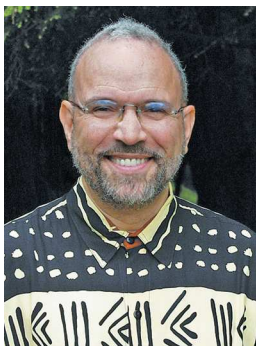
Courtesy of D. McIntyre

**Figure 13.4** Annual growth rings show clearly in this cross section of the trunk of a white pine (*Pinus strobus*).

## 13.2 Successional Change Has Three General Causes

The theory of succession (see Table 13.1, Figure 13.1) identifies three general causes of successional change: site availability, differential species availability, and differential species performance (see Table 13.1, proposition 4). Site availability is, in turn, the result of two factors: disturbance and site characteristics such as soil nutrients (see Chapter 4) and climatic regime (see Chapter 16). In this chapter we will mostly explore thre role of disturbance. Different kinds of disturbances over a range of scales play a crucial role in community dynamics.

Peter White and Steward Pickett (1985) define a **disturbance** as a relatively discrete event in time that causes abrupt change in ecosystem, community, or population structure and changes resource availability, substrate availability, or the physical environment. Disturbances can result in the removal of a substantial portion of the existing vegetation in an area. They can range from damage to a single plant to the destruction of forests or prairies over thousands of hectares. Sources of disturbance include fire, storms (windstorms, ice storms, tornadoes, hurricanes), landslides, earthquakes, mudslides, volcanic eruptions, floods, the activities of animals, and disease. Because of the great diversity of disturbances and their effects, it is almost always more useful to discuss specific kinds of disturbance rather than "disturbance" as a single phenomenon.



Courtesy of Steward T. A. Pickett

Steward T. A. Pickett

One way of classifying disturbances is by whether or not they completely remove a community, including any organic soil. These two aspects of disturbance lead to two categories of succession. **Primary succession** occurs when plants colonize ground that has not been previously vegetated. Examples include the establishment of plant communities on lava fields, earth laid bare by the retreat of a glacier, rock outcrops, sand dunes, and newly formed beaches and sandbars in rivers, or, at a much larger scale, islands newly emerged from the sea. Primary succession also occurs where human activity results in massive soil disturbance, as on mine wastes, roads blasted through rock outcrops, and other sites that are left with undifferentiated soil parent material (Walker 1999). A key aspect of primary succession is that the process begins with the development of soil, which forms as the plant community develops (see Chapter 4).

**Secondary succession** occurs when plants colonize ground previously occupied by a living community. In this case, soil already exists, and plant propagules, such as seeds and ramets, are readily available. Forest regrowth following a catastrophic fire is an example of secondary succession, as is the colonization of an abandoned agricultural field, called **old-field succession.** Once soils are well established, there is no fundamental difference between primary and secondary succession, as long as sources of colonizing plants are readily available. The effects of disturbance under these circumstances are particularly variable depending on the cause; although both remove vegetation, a fire will have very different effects from a hurricane, resulting in a different starting point and different successional trajectories.

Since disturbances are normal characteristics of all ecosystems, ecologists often discuss the **disturbance regime** of an ecosystem—the characteristics of the disturbances occurring in that ecosystem. We can describe disturbance regimes using three characteristics: intensity, size, and frequency. The **intensity** of a disturbance is the amount of change that it causes. A forest fire (Figure 13.5), for example, can be a light **surface fire**, which runs along the soil surface or ground vegetation and destroys only herbaceous plants or low-growing shrubs, or it can be a severe **crown fire** that spreads from the canopy of one tree to another, killing most of the canopy trees as well as most of the shrubs and other vegetation. The **size** (spatial extent) of a disturbance is the amount of area affected. A windstorm, for example, might blow down a single tree, while a hurricane might blow down hundreds or even thousands of trees over an entire section of a forest. And within that extent, the intensity of the disturbance can be very patchy; a fire might heavily burn one spot and completely miss another. The **frequency** of a disturbance—also called the **return interval**—quantifies how often, on average, it occurs in a particular place. These three characteristics are often correlated: small disturbances of low intensity are generally much more frequent than large, intense disturbances. These factors vary, however, depending on the type of community and disturbance. The timing of the disturbance can also be important. For example, a fire that occurs in early spring, before plants are actively growing, can have very different effects from one in late summer, when they are just beginning to set seed (Biondini et al. 1989; Howe 1995).

### Disturbance size affects which species can colonize

The size of a disturbance is a major factor in the types of species that can colonize a disturbed site. Disturbances create **gaps** in communities that can be filled

(A)

Courtesy of S. Scheiner

(B)

Courtesy of Tim Kozusko

(C)

Courtesy of S. Scheiner

**Figure 13.5** (A) A surface fire in northern Michigan, in which only ground-level vegetation was burned. (B) A crown fire in central New Jersey in 1986, in which entire trees were killed. (C) A crown fire occurred 3 years before this photograph was taken, killing most of the mature trees in this spruce-fir forest in the Rocky Mountains. At this time, only low-growing herbs and shrubs are found.

by colonizing species. In any given community, gaps of various sizes can exist at any given time, so succession may be occurring on a variety of scales. For example, mounds made by badgers (*Taxidea taxus*) in North American prairies are 0.2–0.3 $m^2$ in size. A study by William Platt and Michael Weis (1977) found that each year, about 0.01% of the prairie is turned over by the creation of new mounds. The mounds are colonized by nearby plants whose seeds have limited dispersal abilities. In contrast, a massive forest fire that burns thousands of hectares may leave no nearby seed sources. In this case, the initial colonists will be seeds that have traveled long distances by wind or animal dispersal or have remained dormant in the soil.

A forest can be thought of as a set of patches of widely varying sizes that have experienced disturbances of different types and intensities. The size of a patch has major effects on species composition, the successional trajectory, and ecosystem processes within that patch. In an experimental study on gaps of different sizes in the southern Appalachian Mountains in North Carolina, solar irradiance was two to four times higher in large than in small gaps, and soil and air temperatures were much higher (Phillips and Shure 1990). Standing crop biomass and aboveground net primary production (see Chapter 5) were three to four times greater in the largest than in the smallest gaps. Species richness was also greater in large gaps, and species composition differed among gaps of different sizes. We return to the question of how gap formation and gap size influence overall community species richness in Chapter 14.

Ecologists' thinking about disturbance has changed substantially over the past few decades. Early ecologists, for example, thought of disturbances as being unusual—or unnatural—occurrences that disrupted the ordinary and orderly processes of a community. In the past 30 years, however, disturbance has been recognized as a natural part of many communities. In the ecology of rangelands and other arid and semiarid systems, a similar debate has occurred, resulting in a more integrative view of community dynamics, with the dichotomy

of equilibrium and nonequilibrium views replaced by a single continuum in which system dynamics are dependent on temporal and spatial scales (Briske et al. 2003; Briske 2017). In these systems, succession is not about progression to forests, because trees are largely absent and rainfall is insufficient to maintain forests. These systems can move from being dominated by grasses to being dominated by shrubs, depending on the frequency and intensity of grazing, drought, and fire. All of these changes in viewpoints have affected the way ecologists understand mechanisms for species coexistence (see Chapter 10) and have also altered management practices, as we will see in the next section.

### Fire can cause disturbance

Fire is a major source of disturbance in many communities. Fires can vary tremendously in intensity, size, and frequency (see Figure 13.5). One measurement of fire intensity is the amount of heat transferred per unit area per unit time. The most intense forest fires can release as much as 500,000 kJ/m$^2$ in a few minutes, which is enough heat to melt an aluminum-block engine. Over its entire course, the Mack Lake Fire (discussed below) released approximately $3 \times 10^{12}$ kJ of energy, as much as 90 thunderstorms, or nine times the energy of the atomic bomb dropped on Hiroshima (Pyne et al. 1996). The speed with which a fire spreads is another aspect of its intensity. Surface fires can move quickly through an area. In a North American shortgrass prairie on a very windy day, a fire can move as fast as 22 km/h—as fast as an Olympic sprinter—for short periods. Crown fires have been clocked at 12 km/h.

Fire frequency is a key determinant of community structure and composition. In mesic prairies, fires can recur as often as every 2 to 3 years. As we will see in Chapter 16, these frequent fires destroy colonizing trees and so maintain the prairie as grassland. Chaparral communities (see Chapter 18) in southern California burn approximately once every 25 years. Average fire return intervals in different community types vary greatly—from every few years to once a century to once every millennium (Table 13.2). If weather conditions are favorable to fire, the probability of a fire occurring increases with the **fuel load**, the amount of combustible plant material (living and dead) in the community. In prairies, the grasses die back each winter, and highly combustible dead grass builds up over time. This buildup of material increases the probability of fire with time, as well as increasing the intensity of fire when it finally occurs. Likewise, when exotic grasses invade non-grass-dominated communities, they can contribute to increased fire frequency by creating fuel where it did not previously exist, which in turn appears to facilitate further grass invasion (D'Antonio and Vitousek 1992). In western North America, the European annual *Bromus tectorum* (cheatgrass) displaced native perennial grasses, increasing fire frquency, which further supressed native grasses, increasing both the abundance of *B. tectorum* and even more fires. In Hawaii, submontane woodlands that almost never experienced fire were invaded by *Schizachyrium condensatum* (Colombian bluestem) in the late 1960s and now burn on a regular basis. In many *Eucalyptus* forests in Australia, the pine barrens of the eastern United States, and chaparral in California and other parts of the world, similar processes of fuel buildup occur. Here the buildup is of living twigs, stems, and leaves of the dominant species.

In some other community types, the probability of fire does not increase with time. In these communities,

**TABLE 13.2** Examples of Fire Regimes in Various Community Types in North America

| Fire regime | Communities |
|---|---|
| Natural fires rare or absent | Pacific Northwest coastal forests; wetter regions of eastern deciduous forests; southwestern deserts |
| Infrequent, low-intensity surface fires with a ±25-year return interval | Most eastern deciduous forests; pinyon-juniper woodlands of the Southwest; some montane meadows in the Rockies and Sierras |
| Frequent, low-intensity surface fires with a 1- to 25-year return interval combined with high-intensity surface fires with a 200- to 1000-year interval | Sierra mixed coniferous forests; western montane-zone pine forests; southeastern pine forests; prairies of Nebraska and Oklahoma; sawgrass prairie fires in the Florida Everglades |
| Infrequent, intense surface fires with a +25-year return interval combined with crown fires with a 100- to 300-year interval | Pine forests and boreal forests in the Great Lakes region; lower-elevation forests of the Rockies; California redwood forests |
| Frequent, intense surface and/or crown fires with a 25- to 100-year return interval | Most boreal forests; higher-elevation western forests; chaparral from California to Texas |
| Infrequent crown fires with a return interval >100 years | Northwestern wet coastal montane forests; subalpine forests of the western mountains; rainforests of Hawaii |

*Source*: S. J. Pyne et al. 1996. *Introduction to Wildland Fire*, 2e. Wiley: New York, based on K. M. Davis and R. W. Mutch. 1994. In *M580: Fire in Ecosystem Management*, training course (Natl Adv Fire Res Inst: Tucson, AZ).

weather conditions, rather than fuel availability, limit fire frequency and intensity. In the boreal forests of North America, for example, the chance of a fire occurring in a given stand is roughly independent of the time since the last fire in that stand. There is usually adequate fuel for a fire, but weather is only sometimes conducive to forest fires (Johnson 1992). Interestingly, a roughly constant chance of fire also occurs in stands of California chaparral that have long been unburned: apparently, once the fuel load reaches a critical level, fuel availability no longer strongly affects the chance of fire (Keeley et al. 1999).

Some plant species are **pyrogenic**—that is, their accumulated leaf or twig litter tends to promote fire more than one would expect based on the mass of the litter alone. Examples include many *Eucalyptus*, some chaparral shrubs, and possibly some pines, all of which produce oils and other flammable chemicals. Mutch (1970) proposed that pyrogenicity might be an evolved adaptation. This proposal remains quite controversial, however, as the conditions required for such a trait to evolve as a direct adaptation to fire may be quite restrictive (see Chapter 7) (Kerr et al. 1999). Regardless of whether pyrogenicity is an evolutionary adaptation in itself or a by-product of selection on other traits, pyrogenic species reestablish themselves following a fire either because adults resprout or because the population has a seed bank (which may be located in the soil or in serotinous cones or fruits). Meanwhile, competing species are often killed by the fire. Thus, these pyrogenic plants create an environment that enhances their own persistence (or facilitates their invasion of a new community).

The problem of increasing fuel load with time led to a strenuous debate about how best to manage forests in the United States. For most of the twentieth century, the policy of the U.S. government was to suppress fire as much as possible. As a result, fuel loads and the density of young trees increased. Therefore, when fires did occur, rather than being light surface fires, they became raging crown fires, often killing mature trees and threatening human lives and property. In the 1970s, this management practice began to change for two reasons. First, the problems and dangers associated with accumulating fuel loads were recognized. Second, because of the increased understanding gained by ecologists, U.S. government agencies started to recognize that disturbance and fire are natural parts of ecosystems. In the absence of the natural disturbance regime, many properties of a community can change, including species composition. In the southeastern United States, for example, in the absence of fire, hardwood-dominated forests eventually replace longleaf pine forests.

The new U.S. government management practices include letting naturally occurring forest fires burn themselves out (if they are not threatening human lives or property) and setting **prescribed** or **controlled burns** to reduce accumulated fuel loads (to avert a much bigger, uncontrolled fire). Both of these practices are controversial, however. The extensive wildfires in Yellowstone National Park in the summer of 1988, which began naturally and were allowed to burn, were critical in turning public opinion against a policy of allowing fires to burn. Prescribed burns, too, have become a subject of controversy.

Prescribed burns are planned by foresters and set only after careful evaluation of weather and other conditions. Most have been safe and effective in reducing fuel load and in facilitating regeneration of fire-adapted species such as longleaf pine. They are used routinely to manage forests in many regions of the United States and elsewhere. There have been occasional but spectacular mistakes, however, in which controlled burns were set without the proper precautions and became uncontrolled wildfires. The Mack Lake Fire on May 5, 1980, in the Huron National Forest of Michigan began as a prescribed burn, set to help manage the jack pine community. Unfortunately, the fire got out of control. In 12 hours, the fire had burned 10,000 ha; tragically, it also killed a firefighter and destroyed 44 homes (Pyne et al. 1996). A prescribed burn in the spring of 2000 near Los Alamos, New Mexico, escaped control, burning thousands of hectares and destroying several dozen homes; fortunately, no lives were lost.

Managing forests and forest fires has become more complex and difficult as suburbanization and the rise in ownership of vacation and retirement homes in scenic areas—the wildland-urban interface—has led to development of increasing numbers of homes encroaching into forested wildlands. The Camp Fire in November 2018, which destroyed the town of Paradise, was the deadliest and most destructive in California history, burning more than 620,000 ha, causing over $8 billion in damage, including the destruction of nearly 19,000 buildings, and killing over 80 people; it is currently ranked as the sixth most deadly in U.S. history. The intensity and destructiveness of the fire was almost certainly made worse by the effects of global climate change (see Chapter 16). Furthermore, incidents such as the Mack Lake and Los Alamos fires have made decisions regarding fire management political and contentious. Many of these discussions are underlain by profound disagreements between environmentalists, logging corporations, and loggers over the extent to which the U.S. government should facilitate the logging of public forests. The logging industry proposes that it be permitted to thin forests to prevent severe fires from occurring. Critics of this proposal say that this is just an excuse to continue extensive logging of public lands, as the type of logging that is actually done removes most of the larger, commercially (and ecologically) valuable trees, leaving

the smaller trees that are most vulnerable to fire. An alternative that has been suggested by some environmentalists and foresters is to thin commercially valueless small trees and saplings, which are the most likely to contribute to catastrophic fires, but this proposal has not been enthusiastically embraced by the logging industry because it is not profitable. The results of these controversies have been protracted court battles without any clear resolution at present. Even without a political, economic, or environmental agenda to consider, there is also the problem of determining what a "natural" fire regime for a forest community might be and attempting to achieve that state without a catastrophic wildfire.

## Wind can cause disturbance

Wind can be another significant source of disturbance. **Windthrows** range from blowdowns of branches or large parts of trees to losses of single trees and neighboring groups of trees. Windthrows are important in many tropical forests (Figure 13.6) (Brokaw 1985), where trees can be very tall and are often connected by vines. As a result, when one tree falls, it often brings down others (Putz 1983). In the Mpassa forest of Gabon, Africa, for example, 51% of gaps were caused by single falling trees and were responsible for 38% of the total gap area (Florence 1981). Trees that fell in a domino-like fashion accounted for 14% of gaps and 16% of the total gap area. Once a gap forms, neighboring trees become more susceptible to being blown down. In Mpassa, such adjacent treefalls made up 13% of the gaps and 36% of the total gap area. In that forest, the average time between gap formation at any one spot is about 60 years, and the average size of a gap is 3 ha. Treefall in such forests is clearly a very important source of disturbance.

© Gregory G. Dimijian, M.D./Science Source

**Figure 13.6** A large *Dipteryx panamensis* (almendro tree, Fabaceae) has created a gap in the tropical forest at La Selva Biological Station in Costa Rica. This photograph was taken approximately 1 year after the tree fell.

Extremely powerful storms such as hurricanes and tornadoes, while rare, are an important source of wind damage. Hurricanes, typhoons, and cyclones are important in coastal regions. In the Caribbean, for example, hurricanes pass over any given patch of forest, on average, every 15–20 years. These storms can also be important in temperate regions. In the northeastern United States, hurricanes severe enough to blow down large tracts of forest recur once or twice a century. In 1938, a hurricane blew down 253,000 ha of forest in central New England (destroying most of the Pisgah forest mentioned in Box 13A). Tornadoes are another important source of wind damage, especially in the midwestern U.S. Tornado Alley (Figure 13.7). In the forests of northern Wisconsin, catastrophic windthrows create, on average, 52 patches each year, ranging in size from 1 to 3785 ha, with a return interval of 1210 years (Canham and Loucks 1984).

The gaps and patches created by all of these kinds of wind damage are important factors in forest community dynamics. Soil erosion by wind can also be a major disturbance factor in more open communities, from arid lands to plowed former grasslands (see Box 4B). Windstorms can also redistribute leaves and other organic material, creating large deposits in one spot and not another, thereby increasing the amount of variation in soil nutrients from place to place.

## Water can cause disturbance

Water can be an important source of disturbance in both its liquid state, in the form of floods and nonflood erosion (including soil erosion), and its solid state, as snow and ice. Floods are most important in **riparian** habitats (areas adjacent to streams and rivers) and in areas near swamps and bogs. In many such systems, annual floods are continually creating and destroying habitat. In Alaska, for example, the sandbar willow (*Salix exigua*, Salicaceae) grows on sandbars along rivers that are continually being destroyed and re-formed by snowmelt-caused floods. Heavy rainstorms can also lead to landslides. These types of disturbances can initiate primary succession because they tend to create new, previously unvegetated ground. Similarly, snow avalanches are responsible for disturbances in temperate

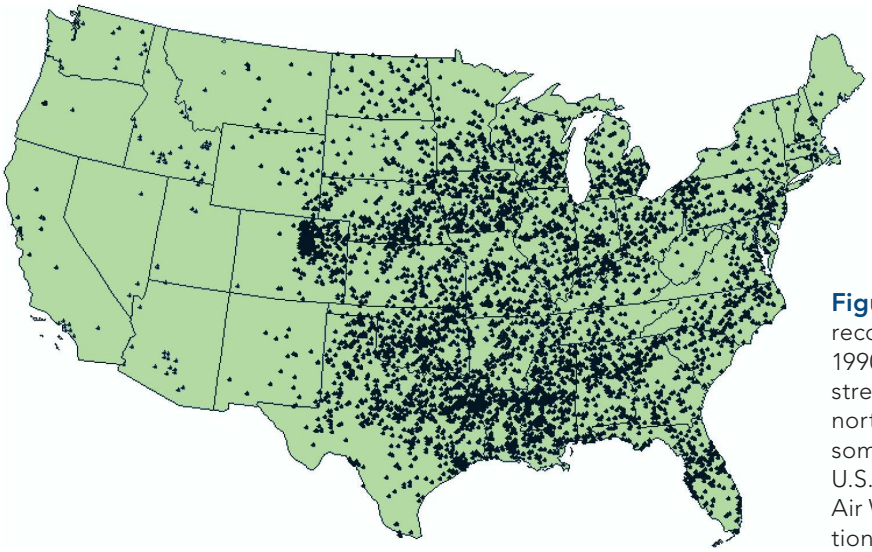


**Figure 13.7** Locations of all tornadoes recorded in the United States from 1981 to 1990. Tornadoes are most frequent in a region stretching from eastern Texas and Louisiana northward to Minnesota and Michigan, an area sometimes called Tornado Alley. (Courtesy of U.S. Air Force Combat Climatology Center, Air Weather Service and NOAA/Storm Prediction Center.)

mountain regions. In the Canadian Rockies, avalanches remove 1% of the forests each winter.

Ice storms are important sources of disturbance in many temperate regions. In temperate deciduous forests, such as those of the southeastern United States, they are often responsible for many small-scale disturbances, such as the removal of single branches from trees. Such disturbances create new patches in the forest. Ice storms can also cause large-scale disturbances. In January 1998, a massive ice storm covered large areas of New York and New England in the United States and Ontario, Quebec, and New Brunswick in Canada. Approximately 25% of the trees in the region were damaged moderately to severely in that single storm.

The combination of disturbance by wind and by water can be important in some communities, such as the cove forests of the southern Appalachian Mountains (Runkle 1985). These highly diverse forests occur in sheltered areas, near creeks, at middle elevations. Fire is very rare in these forests; windthrows are the major source of disturbance. Most gaps are caused by the deaths of single trees due to ice storms, lightning, or high winds. The average gap is small, about 31 $m^2$, although gaps can be as large as 0.15 ha when several trees fall together. On average, a single spot experiences a gap-causing disturbance once every 100 years.

Drought—the lack of water—can also be a disturbance that can cause devastating changes to both natural communities and agriculture. The combination of wind and drought resulted in enormous changes in the vegetation of the Great Plains during the Dust Bowl of the 1930s (Box 13B). Western China experienced a large-scale drought in the early twentieth century, and droughts have been happening in the Sahel region of northern Africa, just south of the Sahara, at least once a century for the past 400 years.

## Animals can cause disturbance

For the most part, disturbances caused by animals are small and often associated with herbivory (see Chapter 11). Sometimes the results of herbivory can be quite dramatic, however, as when a herd of elephants strips the foliage from a stand of trees. A single elephant can consume 225 kg of forage per day.

Even small animals such as insects can cause widespread disturbance during population outbreaks. The gypsy moth (*Lymantria dispar*), for example, was introduced from Europe into Massachusetts in about 1868, in an ill-advised attempt to establish a silk industry in North America by crossing it with the silk moth. The gypsy moth escaped from cultivation and spread rapidly throughout the eastern deciduous forests of North America; today, it can be found as far west as Oregon and as far south as Arkansas. During an outbreak, gypsy moth caterpillars can defoliate large tracts of forest, under some circumstances killing many of the trees (Figure 13.8). Bark beetles can kill hundreds of thousands of coniferous trees, causing as much change as a major forest fire (see Figure 11.3).

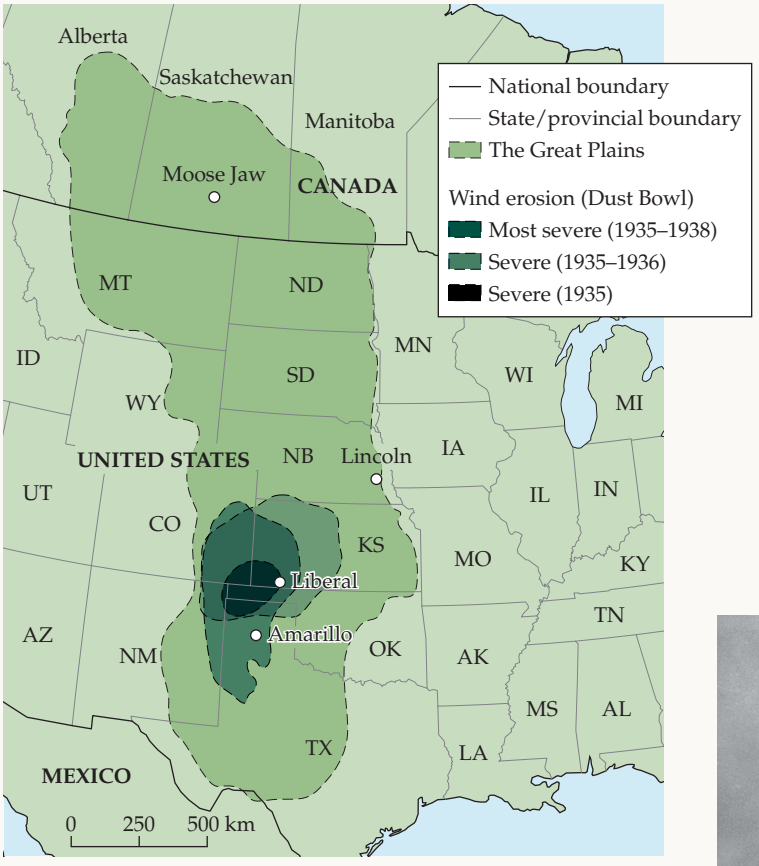
Many other kinds of animal activities can create disturbances (see Chapter 11). Large animals, such as the American bison in North America and the African buffalo in Africa, graze and also create wallows. In the Konza Prairie Research Natural Area (see Chapter 1), bison wallows are important habitats for annual species found nowhere else on the prairie.

## Earthquakes and volcanoes can cause disturbance

Seismic events such as earthquakes and volcanic eruptions are important in much of the world. Volcanic lava flows can create new islands in the ocean or bury existing lands. Seismic events can also have indirect effects by triggering landslides. In Chile, a series of devastating

## BOX 13B The Dust Bowl of the 1930s

For much of the decade of the 1930s, there was a severe drought in the Great Plains of North America. At the same time, the United States (and much of the rest of the world) was in the grip of a severe economic crisis, known as the Great Depression. Before the expansion of European settlement, the Great Plains was a short- and mixed-grass prairie, but starting in the mid-nineteenth century, nearly all was converted to cropland or rangeland for cattle grazing (McLeman et al. 2014). Only remnants of the great prairies are currently protected in preserves and parks. Years of limited rainfall combined with poor agricultural practices, including plowing marginally suitable grasslands and overgrazing, led to ecological disaster. A complex of economic, environmental, and social factors led to the collapse of commodity prices, widespread crop failure, high rates of unemployment, tremendous soil erosion, and dust storms. These were followed by the abandonment of family farms; bankruptcies by individuals, businesses, and local governments; and huge out-migration, particularly to the West Coast. The causes of the Dust Bowl are complex and include human actions and natural climatic fluctuation (severe La Niña conditions, see Chapter 16). Afterward land management practices were extensively reformed, but many of these reforms have gradually been dropped, and current practices are inviting a return of high levels of soil erosion (see Chapter 4). Combined with the depletion of irrigation water from rivers and aquifers, and warmer temperatures and precipitation changes due to climate change, these changes do not augur well for the future.




In the Great Plains, the greatest amount of wind erosion in the 1930s occurred in a region that included Colorado, Kansas, Oklahoma, and Texas. This region was dubbed the Dust Bowl. (After R. A. McLeman et al. 2014. *Popul Environ* 35: 417-440/CC BY. Great Plains boundaries based on those in D. J. Wishart. 2011. *Encyclopedia of the Great Plains*. University of Nebraska Press: Lincoln. Outline of the Dust Bowl region is based on USDA National Resource Conservation Service wind erosion maps for 1935, 1936, and 1938, viewable at http://www.nrcs.usda.gov/Internet/FSE_MEDIA/stelprdb1049472.png

Arthur Rothstein, 1915–1985

Farmer and sons in a dust storm, Oklahoma, 1936.

© J. Boval/Science Source

**Figure 13.8** Gypsy moth caterpillars (*Lymantria dispar*) defoliated this *Fagus grandifolia* (American beech, Fagaceae) tree in an outbreak that occurred in Connecticut in 1980. Deciduous trees do not always recover from such an attack.

Courtesy of U.S. Geological Survey

**Figure 13.9** The 1980 eruption of Mount St. Helens in Washington State created nearly 16 km$^2$ of bare ground, leading to primary succession.

earthquakes in 1960 set off thousands of debris avalanches, landslides, and mudflows, which covered 25,000 ha, or 2.8%, of the land surface of the province of Valdivia (Veblen and Ashton 1978). The Krakatoa volcanic complex in Indonesia has provided important data on primary succession because of its size and isolation. In 1883, Krakatoa erupted so forcefully that the explosion was heard 3500 km away. Only about a third of the original island remained above sea level after the eruption, but several new islands were formed from the volcanic material ejected. Continued volcanic activity has added to the land area. Early descriptions of Krakatoa following the main eruption say that it was essentially scoured of living things and that colonists arrived slowly from distant sources. It is impossible to judge whether this was truly the case, or whether the observers failed to look for the "biological legacies" that have proved to be important in studies of other regions. The eruption of Mount St. Helens in Washington State on May 18, 1980, was the most extreme in U.S. recorded history (although it was much smaller than Krakatoa). A large part of the top and side of the mountain was destroyed (Figure 13.9). The 500 km/h blast affected about 700 km$^2$ and created disturbances that included a 0.25 km$^2$ pyroclastic flow, 550 km$^2$ of trees blown down, 96 km$^2$ of trees that were burned or scored, an avalanche of 60 km$^2$ of volcanic debris, huge areas of mudflows, and tephra (ash and associated solid material), wiping out a total of 65,000 ha of forests, lakes, meadows, and streams (Turner et al. 1997).

## Disease can cause disturbance

The spread of disease can also be a source of disturbance. Chestnut blight was introduced to North America in the early twentieth century, leading to the death of chestnut trees (*Castanea dentata*, Fagaceae; see Figure 11.24) over most of the eastern deciduous forest of North America. Other examples of pathogen-caused disturbances are described in Chapter 11.

## Humans can cause disturbance

One of the most important causes of disturbance in natural communities today is human activity (see Chapter 19), including disturbances caused by domestic animals. Sometimes the disturbances we create are accidental or due to deliberate human actions. For example, approximately 35% of the forest fires in the United States, representing 19% of the area burned, are due to either human carelessness or arson.

Other human disturbances are more purposeful. The logging industry removes large numbers of trees from U.S. forests annually in a combination of **clear-cutting** (removal of all trees) and **selective logging** (removal of some commercially valuable trees, leaving others). After clear-cutting, timber companies are legally required to replant to replace the trees they

have removed. However, because the replanted trees are typically members of a single economically valuable species that grows quickly, the resulting forests are often dramatically different from the original forest communities. Following clear-cutting and replanting, most forests become plantations composed of a single tree species, with all of the trees usually of the same age and planted in a regular grid. Furthermore, logging roads can act as corridors for the movement of invasive species while creating barriers to the movement of other species (see Chapter 15).

In many tropical forests, clear-cut timber harvesting and clearing of land for farming has been extensive. Worldwide, in 2018 rainforest destruction was estimated at 1 ha (about equal to two U.S. football fields) gone every second, 87,000 ha (an area larger than New York City) lost per day, and 32 million ha (an area larger than Poland) destroyed over a year. But not only the forest that is destroyed suffers disturbance. The remaining areas of forest are also affected, as fragmentation of the landscape changes species migration patterns and population sizes (see Chapter 15 and Chapter 19).

Grazing by domestic animals—especially cattle, sheep, and goats—has drastically altered the composition of grasslands and shrublands over huge tracts of arid and semiarid countryside around the world (see Chapter 11), in many cases greatly accelerating soil erosion. Even the trampling associated with recreational hiking and other uses of wilderness areas can be substantial in areas with large numbers of visitors or in areas that are especially vulnerable, such as deserts and tundra.

Courtesy of Lawrence R. Walker

**Figure 13.10** A rock-strewn area of bare ground at the base of Franz Josef Glacier in New Zealand. This area has not yet been colonized by any plants.

## 13.3 Which Species Are Available for Colonization Affects Succession

### The dispersal capacity of species affects their colonization capability

All successional processes begin with propagules. Primary succession differs from secondary succession in the types of propagules involved and in their sources. In primary succession, there is generally little or no true soil—the substrate lacks structure and any substantial organic component (Figure 13.10). Most species are incapable of persisting under these circumstances and can colonize the site only after substantial soil development. Thus, a key issue in many primary successional settings is how the initially colonizing plants affect soil development. Generally, in primary succession, all propagules must be seeds or spores carried to the newly bare ground by wind or water. Animal dispersal plays a small role in primary succession because relatively few animals visit unvegetated sites. First to appear are species of mosses and lichens that specialize in colonizing bare ground. These species tend to be slow growing and to possess unique adaptations that enable them to live under the difficult conditions found at primary successional sites. In addition to the difficulties posed by the lack of organic matter, nutrients, and soil structure, other conditions, such as widely fluctuating temperatures and exposure to high winds, make survival at these sites difficult for most plants.

Colonization by dispersing seeds is also important in secondary succession, but the colonizers usually differ from those involved in primary succession. Because soil is already present, many fast-growing species can grow in the newly available patches. While newly disturbed areas such as gaps present many challenges to colonizing individuals, they also offer numerous advantages. Light levels are generally high because of the lack of a dominant plant canopy. Soil nutrients may also be high because of both reduced competition and the release of nutrients from decaying plants or plant

parts (e.g., leaves from the canopy trees killed in a large blowdown). The colonizing species in secondary succession are often (but not always) very fast growers with small, wind-dispersed seeds, such as *Taraxacum officinale* (dandelion, Asteraceae). Often these colonizing species are found at a site for only a year or two following the disturbance before being outcompeted by other species.

A lack of propagules can persist for a very long time after a disturbance. In the tropical rainforest of the Amazon basin of Ecuador, tree species richness is generally very high. However, one area of forest has noticeably fewer species. Nigel Pitman and associates (2005) concluded that the lower species richness in this area was due to a catastrophic flooding event that had occurred about 500 years previously. Even now, not all of the trees in the region have colonized the site.

Research on the Krakatoa archipelago by Susanne Schmitt and Robert Whittaker (1996) suggests that limitations on dispersal may be quite important to the successional process. Shoreline species recolonized the islands very rapidly, but these were largely water-dispersed species. Species that grow farther inland colonized the islands much more slowly. Wind-dispersed species, especially grasses, composites, and ferns, predominated among the first inland colonists, creating grasslands. Woody plants began colonizing the islands somewhat later; forests dominated the islands by the 1920s. The successional trajectories were similar among the islands, but there were also differences among islands and among sites within islands. These differences seemed to be due to chance differences in dispersal as well as environmental differences. Continued volcanic activity, resulting in the heavy deposition of ash on some islands, has contributed to increasing differences among the islands over time.

Although (by definition) primary succession occurs in unvegetated areas, dispersal distances may not be great enough to limit the ability of particular species to colonize a disturbed area. Around the Kilauea volcano in Hawaii, lava flows often leave small islands of vegetation (kipukas), which are important sources of propagules for the colonization of the lava (Figure 13.11). Most other primary succession, such as that on new beaches and sandbars in rivers, occurs on smaller disturbed sites, where dispersal distances may be short. In primary succession, at least initially, propagules that arrive at the site have few or no interactions with resident populations, because there are none. Competition is absent or of minor importance, and herbivory may be drastically reduced.

Probably the longest-studied site of primary succession is at Glacier Bay, Alaska. William Cooper (1923) set up permanent study plots there in 1916 in areas from which a glacier had retreated fewer than 40 years earlier. In this system, the first colonist of bare ground usually is a **cryptogamic crust**, a thin layer of mosses, fungi, various unicellular photosynthetic organisms, lichens, and photosynthetic bacteria (cyanobacteria). The cryptogamic crust limits erosion and helps to retain moisture, dead parts of the crust provide initial organic matter for soil formation, and dispersing seeds may lodge in cracks in the crust. The small shrub *Dryas drummondii* (Rosaceae) is typically the next colonist. A plant with nitrogen-fixing symbionts in its roots, *Dryas* forms dense mats of vegetation. In some regions of Glacier Bay, *Alnus sinuata* (alder, Betulaceae), another plant with nitrogen-fixing symbionts (see Chapter 4), also forms dense thickets.

Courtesy of S. Scheiner

**Figure 13.11** A lava field on the island of Hawai'i, showing a young tree fern, *Cibotium glaucum* (hapu'u, Dicksoniaceae), that has colonized the bare ground. This species is a Hawaiian endemic and is common on relatively new lava.

The rate of colonization is a key determinant of how species richness changes during succession. An analysis of 62 successional sequences by Kristina Anderson (2007) found that new species entered communities mostly in the early years; although colonization never completely halted, it was much slower in later years. In contrast, the rate of species loss from a site was about the same early and late in succession. The net result was that

species richness would rise quickly early on, and then continue to rise at a slower pace as time went on. These patterns depended on the type of succession, with the rate of colonization falling rapidly for secondary succession and more slowly for primary succession or in places with harsh abiotic conditions.

### Species can emerge from the propagule pool

Soil (or aerial) seed banks (see Chapter 8) can be an important source of propagules for secondary succession. Some species that specialize on early successional habitats have seeds that are able to remain dormant for many years and germinate under specific conditions that follow a disturbance. In many annual and short-lived perennial herbs, seed dormancy is broken only by exposure to sufficient light, which occurs only after soil disturbance. You may have observed this in a garden: after you turn the soil over, many weedy species seem to appear suddenly. Other cues can also be important; for example, some seeds germinate in response to the chemicals in ash or smoke from a fire.

The composition of the soil seed bank can be quite different from that of the plant community growing above it (Oosting and Humphreys 1940; Livingston and Allessio 1968). In an upland oak-pine forest in Michigan, for example, *Arenaria serpyllifolia* (sandwort, Caryophyllaceae) was common in forest seed banks yet was absent above ground. In contrast, *Conyza canadensis* (horseweed, Asteraceae), which has wind-dispersed seeds, was one of the most common species following a fire but was absent from the soil seed bank (Scheiner 1988). A gap in this forest might be quickly colonized by both species: *Arenaria* from the seed bank and *Conyza* by dispersal of wind-borne seeds from other areas.

Vegetative propagules can also be an important source of regrowth following disturbance. Many plant species have adaptations to fire. One such adaptation is the location of the meristems (see Chapter 8) in a place protected from fire (Pausas et al. 2018). Aspens (*Populus tremuloides* and *P. grandidentata*) in northern temperate forests specialize in growing in early successional habitats. These trees can spread by "root suckering," that is, through the growth of new stems as offshoots from rhizomes. Fire stimulates their tendency to produce new shoots. Following an experimental burn in northern lower Michigan, there were 24,750 aspen stems per hectare, and these shoots grew to an average height of over 1 m in a single summer (Figure 13.12).

Prairies, where fires recur frequently, are dominated by grasses and forbs, which have meristems at or below the ground surface. A light, quickly moving surface fire is unlikely to harm these meristems, although the tops of the plants may be destroyed. (Having meristems at ground level also protects these plants from damage by grazers, which may have been the primary selective force acting on this adaptation; see Chapter 11.) Several tree and shrub species have the ability to resprout from roots, rhizomes, or buds underneath the bark if the aboveground portions of the plant are killed or severely damaged by fire. Examples from the frequently burned pine barrens of the eastern United States are *Pinus rigida* (pitch pine, Pinaceae) and *Quercus ilicifolia* (scrub oak, Fagaceae). Other species with such adaptations include *Eucalyptus* spp. (Myrtaceae) in Australia, *Populus tremuloides* (quaking aspen, Salicaceae) at high latitudes or high altitudes in North America, and *Adenostoma fasciculatum* (chamise, Rosaceae) in southern California.

Though volcanic eruptions can be massive disturbances, it is a mistake to think that they always create large areas devoid of life. In the eruption of Mount St. Helens, recovery was very heterogeneous, depending on the form of the disturbance, the availability of seed sources, topography, and microclimate. The initial revegetation was dominated by "biological legacies"—buried roots, small stands of surviving plants, and seeds from nearby untouched areas (Frenzen et al. 1986; Franklin 1990; del Moral and Titus 2018). In other areas there were no surviving propagules, and almost all individuals came from windblown seeds (Dale and Adams 2003). Early successional colonists such as *Chamaenerion angustifolium* (fireweed, Onagraceae, Figure 13.13) and *Lupinus latifolius* (broadleaf lupine, Fabaceae) were found in the first years. Woody plants including *Alnus rubra* (red alder, Betulaceae, a nitrogen-fixing tree) and fast-growing willows including *Salix lasiandra* and *S. scouleriana* (Salicaceae) colonized subsequently. By 15 years after the eruption, conifers had become patchily established, including *Pseudotsuga menziesii* (Douglas fir, Pinaceae), *Abies amabilis* (Pacific silver fir, Pinaceae), *A. procera* (noble fir, Pinaceae), and *Tsuga heterophylla* (western hemlock, Pinaceae). The most severely disturbed areas recovered slowly, and some areas remain poorly vegetated (Dale and Adams 2003). In many of the affected areas, forests have still not returned.

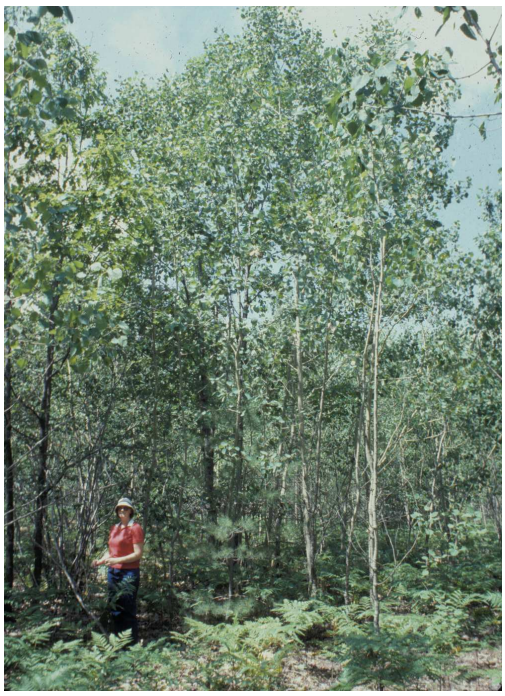
For some species, the propagule pool sits above ground in fruits or cones that release their seeds following a fire. Some pines have a trait called **serotiny**, in which they retain their seeds in tightly sealed cones for many years, releasing them only after exposure to fire. The cones are sealed by resin, which melts during a fire, releasing the seeds. Following a fire, these trees can release many years of accumulated seed production. The sealed cones, therefore, serve as an aboveground

(A)

(B)

(C)

(D)



(E)

(F)

Photos courtesy of S. Scheiner

◀ **Figure 13.12** Disturbance and early succession in a forest in northern lower Michigan. (A) This site was clear-cut and burned in August 1980. (B) One year following the fire, the site is covered by *Populus grandidentata* (big-toothed aspen, Salicaceae) about 1 m tall that has resprouted from roots in response to the fire. (C) The same view 3 years following the fire. (D) An adjacent site that was clear-cut and burned 29 years earlier. The site is still dominated by *P. grandidentata*, now grown to full size. (E) An adjacent site that was clear-cur and burned 47 years earlier. Other tree species have begun to enter the canopy, notably *Pinus strobus* (eastern white pine, Pinaceae). (F) The forest 60 years after fire. The aspen are thinning out and being replaced by *Acer rubrum* (red maple, Sapindaceae), *Quercus rubra* (red oak, Fagaceae), and *P. strobus*. Saplings of the last species are particularly noticeable.

seed bank. Cleared mineral soil may be required for germination in many serotinous species. Among the pine species with serotiny are *Pinus contorta* (lodgepole pine) in western North America, *P. banksiana* (jack pine) in north-central North America, and some populations of *P. rigida* in the eastern United States (Figure 13.14). Serotiny also occurs in fire-adapted taxa unrelated to pines, such as many Australian species in the Proteaceae.

Courtesy of J. Gurevitch

**Figure 13.13** *Chamaenerion angustifolium* (fireweed, Onagraceae) is often an early colonist after major disturbances in much of North America.

(A)

(B)

(C)

Photos courtesy of J. Gurevitch

**Figure 13.14** Serotiny occurs in some *Pinus rigida* (pitch pine, Pinaceae) populations in the pine barrens of Long Island, New York. (A) Ordinary open cones on a typical, nonserotinous tree. (B) Closed, serotinous cones on a tree belonging to a frequently burned population with a high frequency of serotiny, which is a genetically determined trait. (C) A serotinous cone that has opened as a result of a forest fire, releasing its seeds.

## 13.4 Species Performance Determines the Pattern of Successional Change

### Species vary in their life histories

The unfolding of species' life histories (see Chapter 7) is a contributor to successional change. This mechanism is at the heart of one theory of succession. In 1954, Frank Egler described two alternative scenarios for succession (Figure 13.15). The first, known as **relay floristics**, was largely a Clementsian view, in which integrated communities replace one another like runners in a relay race. Egler proposed a second scenario that he thought was more likely: the **initial floristic composition** hypothesis, which emphasized the process of colonization and differences in the life-spans of species. Focusing on succession in abandoned fields in the eastern United States, Egler claimed that all species reach a site early in the successional process. Early successional species, such as annuals and perennial herbs, dominate the site because they grow quickly. Tree species also arrive early but do not dominate until much later because they grow more slowly. Thus, in Egler's view, a particular successional sequence is a direct consequence of the initial composition of the plant community. Put differently, Egler's theory was that succession is simply a process of the unfolding of various life histories that proceed at different rates; species neither facilitate nor inhibit other species. In this sense, Egler's theory of initial floristic composition resembles Connell and Slatyer's tolerance model of succession (see the next section).

There are numerous examples of life history differences causing shifts in dominance in some communities. For example, *Populus grandidentata* (big-toothed aspen) is an early successional species found in upland forests with poor soils around the North American Great Lakes (see Figure 13.12). It is replaced by a variety of species, including *Quercus rubra* (red oak). This pattern occurs because aspen grows very quickly. Its quick growth means that it can colonize rapidly, but its wood is not very strong and is susceptible to damage by fungal attack and by high winds. In this region aspen trees (individual ramets) rarely live beyond 60–70 years. In contrast,

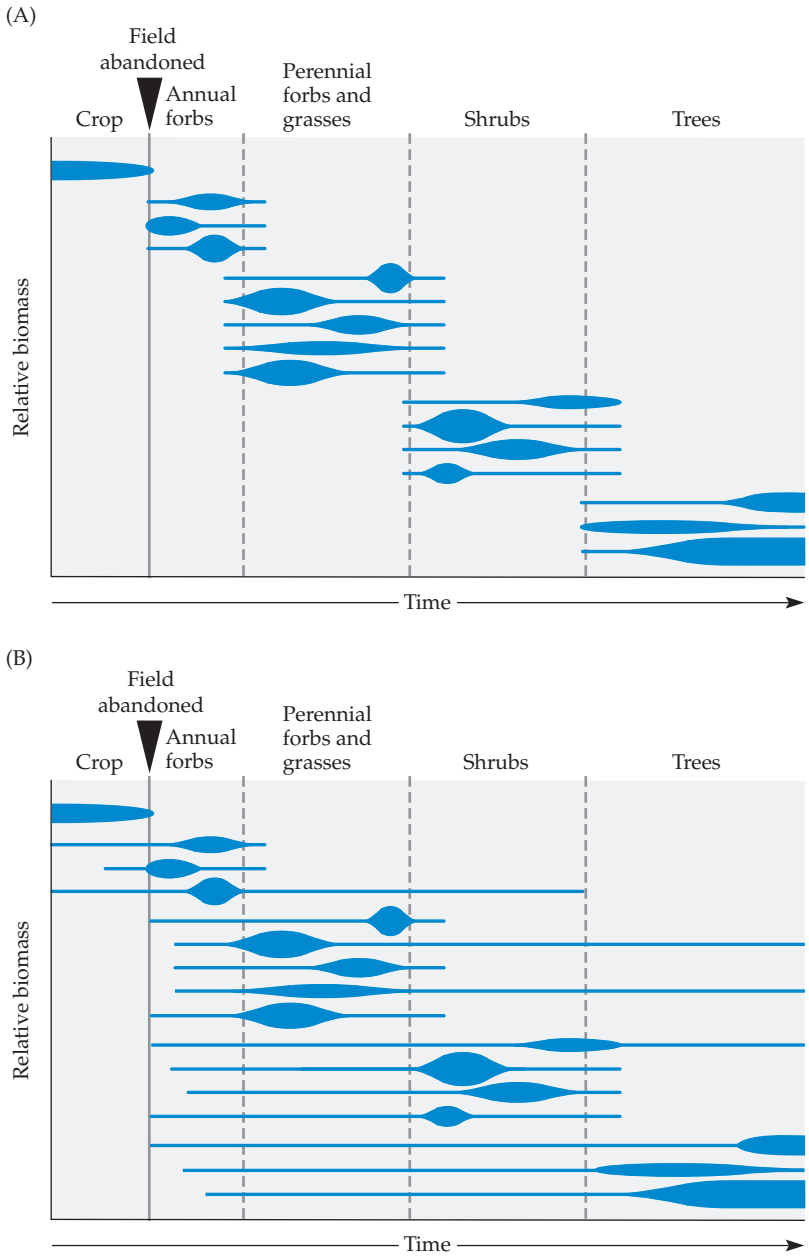



**Figure 13.15** F. E. Egler's theories of succession, stylistically diagrammed for a hypothetical abandoned field in North Carolina. Each line represents a single species of the vegetation type indicated. The thicker the line, the more important the species at a given time. (A) According to relay floristics, groups of species replace one another like runners in a relay race. (B) Egler's modification of relay floristics was the initial floristic composition theory. Here, all of the species are present at the beginning of succession, which is simply a process of the unfolding of their various life histories. (After F. E. Egler. 1954. *Vegetatio* 4: 412–417.)

oaks grow very slowly, producing the strong, dense wood prized by furniture makers. Oaks may colonize an aspen stand very quickly, but they will not dominate for decades. In other cases, however, later dominance is clearly the result of a species' inability to colonize an early successional site. For example, *Tilia americana* cannot colonize new sand dunes. It requires other species to stabilize the dunes and begin the process of soil formation.

Other types of life history patterns are important adaptations to particular types of disturbance. *Pinus palustris* (longleaf pine) is a dominant tree in forests of the southeastern United States, where surface fires recur approximately every 5 to 7 years. Young seedlings are especially vulnerable to these fires. *P. palustris* has an unusual life history that appears to be an adaptation to these frequent fires: After the first year or so, the young plant exists as a small tuft of needles growing at ground level, looking very much like a clump of grass (Figure 13.16). The apical meristem sits just below the ground, where it is often undamaged if a low-intensity fire occurs. During this time, the plant grows a large root system and creates a large nutrient store. Finally, when its nutrient store is large enough, it grows very quickly. Within a few years, the tree is big enough, and has bark thick enough, to be quite fire-resistant. Thus, its growth pattern minimizes the number of years during which it is vulnerable to low-intensity fire.

© Gilbert S. Grant/Science Source

**Figure 13.16** A seedling of *Pinus palustris* (longleaf pine, Pinaceae) looks very much like a clump of grass—hence the term *grass stage*. The apical meristem is located out of sight, just below the ground, where it is protected from fire.

Life history patterns can also include changes in growth form. *Quercus velutina* (black oak) in the eastern United States and *Pinus ponderosa* (ponderosa pine) in the western United States have very thick bark that protects the trees' cambium (the living, growing part of the trunk just under the bark) during a surface fire. As they grow, these species also tend to shed their lower branches, which could serve as a "fire ladder," with the result that fire cannot as easily spread to the canopy.

### Species interactions are central to species replacement during succession

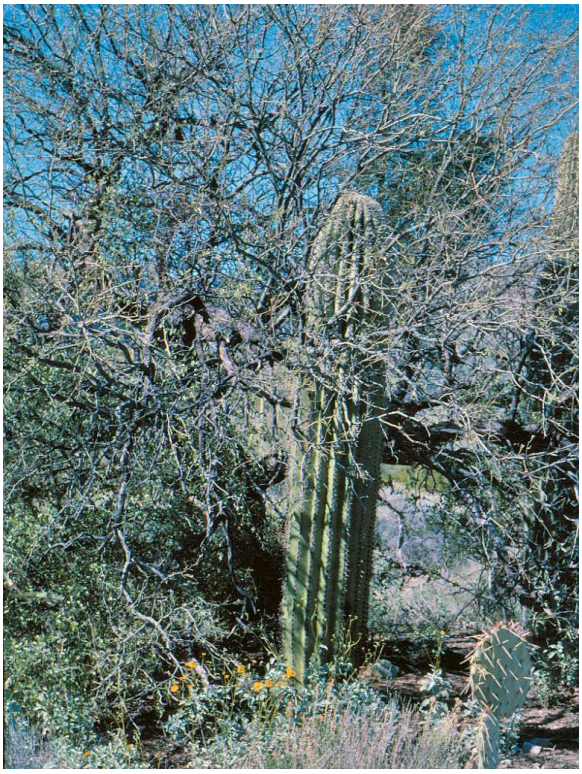
The first comprehensive successional theory was that of Clements (1916); it was part of the superorganism concept, his overall theory of the nature of plant communities (see Chapter 12). In Clements's view, some species played the role of preparing a site for later occupation by other species. Interactions among species were seen as benefiting the functioning of the whole community. Some species existed in the community for a limited time before bowing out and allowing others to take over. Clements believed that these species were somehow designed for this purpose. Today, we recognize that Clements's ideas were profoundly non-Darwinian. He developed them at a time when the Darwinian theory of evolution by natural selection was still a hotly debated issue. Today, no scientist would accept the notion of a species existing only for the good of the community. Although mutualistic interactions can be favored by natural selection (Wilson 1980), their evolution is complex, and their role in succession is much more limited than envisaged by Clements.

How, then, do populations of successional species interact with one another? To develop a set of testable hypotheses, Joseph Connell and Ralph Slatyer (1977) proposed three possible mechanisms by which succession might occur. While we present these mechanisms as alternatives, keep in mind that all three may operate in any given community and that they probably interact with one another (Walker and Chapin 1987). First, early successional species might **facilitate** the colonization and growth of later successional species. This hypothesis looks at succession from a viewpoint akin to Clements's, albeit within an evolutionary context and in a more mechanistic fashion. Second, a process of **inhibition** might drive succession. An early successional species might inhibit colonization by later successional species by monopolizing resources such as light, water, or nutrients. Succession would eventually occur when later species outcompeted earlier species. Finally, a process of **tolerance**, in which species neither help nor hinder colonization by other species, might determine succession.

There is evidence that all three of these mechanisms (see Table 13.1, proposition 6) operate in some settings.

Facilitation is best demonstrated during primary succession. Some species are adapted to eking out a living under difficult conditions. Some lichens, for example, can grow on bare rock. They do so by slowly dissolving away the rock to gain nutrients, thus beginning the process of soil formation. Another example of facilitation during primary succession occurs along lakeshore dunes, such as those of the North American Great Lakes. Species such as *Elymus canadensis* can colonize newly deposited sand, with its very low levels of nutrients and poor water-holding capacity. In areas of active sand dune formation, the fibrous root systems of this grass act to stabilize the sand (see Figure 13.3). Once the sand is stabilized, other species with smaller root systems are able to colonize the site; the establishment of these species is thus facilitated by *Elymus*.

Secondary succession in harsh environments also provides examples of facilitation. In the Sonoran Desert, seedlings of *Parkinsonia microphylla* trees (palo verde, Fabaceae) survive poorly in the open. Instead, they germinate and grow in the shade of shrubs such as *Ambrosia deltoidea* (Asteraceae). Because of this role in the community, the latter are called nurse plants (see Chapter 10). Similarly, *Carnegiea gigantea* (saguaro, Cactaceae) germinate and grow under both *Parkinsonia* and *Ambrosia* (Figure 13.17), so *Parkinsonia* shifts roles to become a nurse plant. The eventual replacement of early successional species by later ones in these instances is not a result of one species voluntarily giving way to another, as Clements suggested. The later species eventually outcompete the early species—in fact, there is good evidence that saguaros eventually kill off their nurse trees. The early successional species persist in the landscape because they are able to colonize and live in ephemeral habitats.



Courtesy of Timothy P. Craig

**Figure 13.17** In Arizona's Sonoran Desert, a saguaro cactus (*Carnegiea gigantea*, Cactaceae) grows in the shade of its nurse plant, a palo verde tree (*Parkinsonia microphylla*, Fabaceae). Once the cactus grows large enough, it no longer requires protection and will displace the tree by outcompeting it for water.

Inhibition and tolerance are the easiest successional processes to demonstrate and observe. Inhibition comes about because of competition for light, water, and nutrients (see Chapter 10). One or more of these resources is always in limited supply in a given community. In a desert, the limiting resource is often water. In very sandy soil, the limiting resource is often nitrogen (see Chapter 4 and Chapter 5). In a forest understory, the limiting resource is usually light. Competition for these limiting resources results in reduced growth and reproduction. Thus, competition reduces a species' ability to colonize and grow in a site.

## Resource availability can change during succession

Successional trajectories are determined both by the initial resources available in a site and by how those resources change through time due to the changing vegetation (see Table 13.1, proposition 9). One example is the role of plants with nitrogen-fixing symbionts in primary succession following deglaciation. As described earlier, both *Dryas drummondii* and *Alnus sinuata* harbor such symbiotes. Not surprisingly, it appears that these early nitrogen fixers facilitate the later colonization of *Populus trichocarpa* (cottonwood) and *Picea sitchensis* (Sitka spruce, Pinaceae), which do not fix nitrogen. Spruce is much more successful as a colonizer of alder thickets than of earlier successional stages, no doubt partly because of the soil nitrogen, soil organic matter, and mycorrhizal fungi that the alder provides. Some spruce trees are inhibited by, while others are facilitated by, the presence of cryptogamic crusts, *Dryas*, and alder, as well as by other spruce trees. Thus, both facilitation and inhibition occur in this system. Different sites at Glacier Bay follow different successional trajectories; in particular, alder does not occur in some areas, mainly due to its limited dispersal ability (Chapin et al. 1994).

Competition is less intense if the competing species are adapted to somewhat different conditions. For example, some species grow best under high-light conditions, while others do best under low-light conditions. During

forest succession, light levels decrease as trees grow up and the canopy fills in. In this case, species adapted to high light intensities tend to occur early in succession (or in gaps in a mature forest), while species adapted to low light intensities usually occur in older communities. In the temperate forests of the northeastern United States, *Prunus pensylvanica* (pin cherry, Rosaceae), a species that grows best in high-light conditions, occurs in early successional forests. Because of its high light requirement, its seedlings cannot grow up underneath its own canopy. Instead, stands of pin cherry are often colonized by *Acer saccharum*, which can grow under much lower light intensities. Growth under these conditions is actually very slow, and a sugar maple sapling would never become a full-sized tree if it were always growing in the shade of another tree. Rather, the strategy of the sugar maple is one of tolerance. It can survive, growing slowly, under low-light conditions. Then, if a canopy tree happens to fall and create an opening, the sugar maple sapling can grow quickly to fill that gap. Because the sugar maple is already at the sapling stage, it can reach the canopy before any of the other trees that germinate in that gap. So pin cherry inhibits—slows down—the colonization of sugar maple, while sugar maple is able to colonize because of its tolerance of low light.

## 13.5 The Pathway of Succession Can Vary

### Succession may or may not be predictable

One of the most important contributions of the initial floristic composition theory was that it called attention to the process of colonization. While earlier theories of succession saw colonization as an important process, they viewed it as a **deterministic** one—a process with a fixed outcome. Colonization, however, is a process fraught with random elements. These elements include the timing of propagule arrival, the number of propagules that arrive at a site, whether those propagules land in a **safe site** (a favorable spot for their germination and growth), and whether climatic conditions are favorable for their establishment when they arrive. Two identical sites might experience different successional sequences due to differences in these chance events. Thus, how often succession is predictable is an open question. (Recall that the use of chronosequences to study succession depends on the assumption that succession is predictable and repeatable.)

One way of thinking about the predictability of succession is by dividing it into two components: the starting point of succession, and the path (or trajectory) that succession takes to its end point (Figure 13.18; see Table 13.1, proposition 8). The starting point encompasses the physical conditions following a disturbance and the propagules that either survive the disturbance or arrive at the site. The early ideas about succession saw all sites in an area as starting from approximately the same starting point and then following a single trajectory to the same end point. Once we recognize the possibility of unpredictability, however, three other outcomes are possible. The following descriptions assume that basic conditions (e.g., climate, soil type, slope, aspect) are similar among sites.

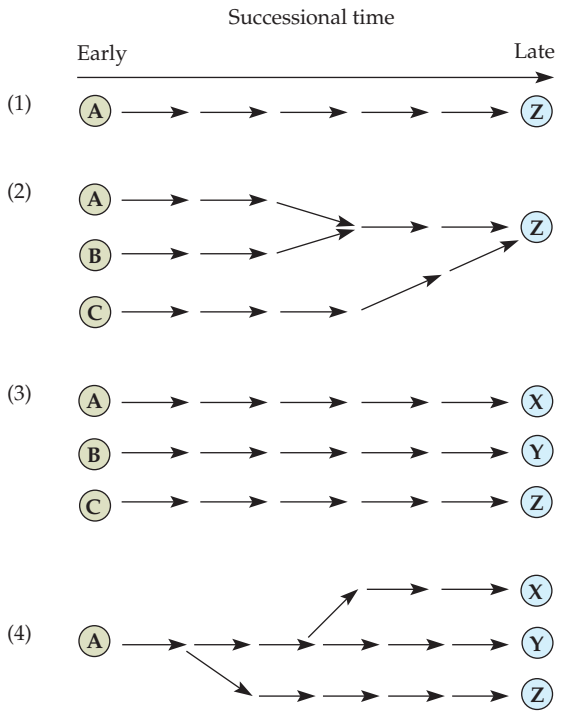



**Figure 13.18** Predictable and unpredictable successional trajectories. (1) Early ideas about succession saw all sites in an area as starting from approximately the same starting point and following a single predictable trajectory to the same end point. (2) Different starting points might still lead to the same end point. In this scenario, deterministic processes dominate, and the outcome is predictable. (3) Different starting points might lead to various unpredictable end points. (4) Even similar starting points might lead to different unpredictable end points.

First, different starting points might still lead to the same end point. While early successional communities might differ considerably in composition, and while the exact successional sequence might differ, all sites in an area might eventually converge on the same type of community. In this scenario, deterministic processes dominate over random ones in the long run. Second, different starting points might lead to different end points. Given different starting points, different successional sequences might unfold. The initial floristic

composition hypothesis would fit under this scenario. Here, random processes dominate in setting the stage for a given successional sequence, although deterministic processes take over thereafter. Third, even similar starting points might lead to different end points. In this scenario, random processes dominate during the entire course of succession.

Currently we have limited data to suggest which of these three scenarios occurs most often. The best evidence that we have favors the first two scenarios, suggesting that deterministic processes dominate in the long run. A study of tropical rainforest communities in the Amazon basin of eastern Peru, for example, found that species composition tended to converge among sites over time (Terborgh et al. 1996). Similarly, in Oosting's study of old-field succession in North Carolina, fields of similar age had very similar sets of species across many different sites. Other factors can complicate the pathway to convergence, however. In a study of succession on mudflows resulting from the eruption of Mount St. Helens, carried out 16 years after the eruption, an area that was relatively close to an intact forest—a source of many propagules—showed evidence of convergence to the composition of that forest, while an area farther away from the forest did not (del Moral 1998). Thus, dispersal rates and opportunities do affect the successional trajectory. In this case, however, since the entire successional process takes several centuries, even the more distant site may converge, given enough time.

Additional evidence of the long-term effects of the starting point is provided by examining succession on lands formerly dominated by human activities. Different land use histories can have long-term consequences for successional trajectories and lead to very different communities. For example, forest succession will be different depending on whether the land in question was burned or cut for timber, or used as pasture, or plowed for crops (Motzkin et al. 1999; Hall et al. 2002). Kerry Brown and Jessica Gurevitch (2004) found that forest clearing led to long-term losses in tropical forest diversity even after forest regrowth. Repeated cutting of the temperate hardwood forests of the northeastern United States has led to a decline in the number of species due to factors such as preemption of space by invasive plants, the dispersal limitations of herbs, and permanent changes to the soil (Bellemare et al. 2002).

All of the examples cited above are observational in nature. To experimentally test these ideas, Tadashi Fukami and colleagues (2005) took former agricultural fields in the Netherlands and sowed them with different mixtures of species, including replicates of some of the mixtures. After 8 years, those treatments that received the same initial seed mixtures remained similar, while those that had different starting conditions diverged in species composition. Interestingly, however, when they examined the functional roles of those species, the treatments had converged. Functionally the communities were similar even though those functions were being carried out by different species.

### Understanding successional processes is critical for community restoration

The predictability of succession is of particular importance to ecologists who are attempting to restore communities—for example, returning agricultural fields to prairie, or surface-mining sites to woodland. Given that a goal for restoration has been selected (itself a complex topic, but one outside the scope of this book), how do ecologists proceed? This sounds like a simple question, and a simple answer would be to plant the species that would dominate the community if the goal were achieved. This approach is not often used, however. In most restoration projects, the species planted are those that dominate at an earlier successional stage than the goal of the restoration project, because the early successional species create the environmental conditions necessary for the persistence and growth of the late successional species. As a simple example, consider restoring an area used for surface mining to mature woodland. Many of the late successional tree species will not become established in bright sunlight, so it may be necessary to first plant some early successional species. Clearly, using successional processes in restoration efforts can work well only if succession itself is fairly predictable. The predictability of succession in restoration efforts largely depends on two things: establishing conditions favorable for the success of the goal (such as shade for woodland species) and ensuring a source of propagules of the desired species.

How important are successional changes to soils in establishing conditions favorable for the goal? Some cases are well understood. In the North American tallgrass prairie, for example, the native soils are relatively nitrogen poor, and many prairie plants are adapted to those soils (Morgan 1997). Sites that have been farmed have often been fertilized, so exotic invasive plants that are adapted to high-nitrogen conditions tend to dominate abandoned farms (Maron and Jefferies 2001). One restoration method for mitigating the effects of the added soil nutrients is "reverse fertilization"—adding large quantities of organic material to the soil, which promotes microbial activity and immobilizes much of the nitrogen (Averett et al. 2004). Native prairie plants that are grown under these conditions have lower concentrations of leaf nitrogen, which thereby reduces the rate at which nitrogen is cycled through the soil (Wedin and Tilman 1996) (see Chapter 5).

These studies suggest that the regional history of land use can have strong effects on the restoration process. Jennifer Fraterrigo and colleagues (2005) studied stands of Appalachian forest in North Carolina that had been used for pastures or timber extraction, but had been abandoned for more than 60 years, and compared them with stands of old-growth forest. They found that historical land use had strong effects on the spatial heterogeneity of the soils: Former pastures had similar levels of carbon, potassium, and phosphorus while logged stands had increased and highly variable levels of calcium. These changes in the spatial patchiness of nutrients are likely to have strong effects on the resulting communities due to their effects on competitive interactions (see Chapter 10 and Chapter 15).

As we saw above, some successional changes involve the soil communities—the fungi, bacteria, and microarthropods—and these can be manipulated in restorations. A restoration project by Roger Smith and colleagues (2003) of meadows in the United Kingdom was speeded significantly by planting nitrogen-fixing legumes that promoted the growth of soil fungi. Julie Korb and associates (2003) found that mechanical thinning of stands of *Pinus ponderosa* (Ponderosa pine, Pinaceae) in northern Arizona, followed by burning, led to substantial increases in the density of arbuscular mycorrhizae and hastened the reestablishment of grasses and herbs in the forest understory. Much remains to be learned about the restoration of soil communities and its interaction with aboveground plant communities.

No restoration program can re-create a community if there is no source of seeds for some of the species. Unless seeds are hand sown, they must come from nearby remnants of the targeted community. Some principles from the theory of island biogeography (see Chapter 15) are useful here: Colonization is most likely from nearby remnants and from large remnants. If establishment and persistence in a site depend strongly on which individual arrives first (as in the lottery model; see Chapter 10), and sources of seeds are relatively distant or small, it may be difficult to predict the outcome of succession.

The pattern of change may depend on the history of the system. In some cases, simply reversing or eliminating the cause of a change may not cause the system to quickly return to its previous state, a process called **hysteresis** (Lockwood and Lockwood 1993). For example, Thomas Valone and colleagues (2002) studied the effects of cattle grazing in an arid grassland in northern New Mexico by surveying two areas that had been fenced to exclude grazers, one for 20 years and one for 39 years. Prior to the building of the fencing, intense overgrazing had rapidly converted the grassland to a shrub-dominated system. But halting that grazing did not immediately return the system to a grassland. After 20 years, the system was still dominated by shrubs. It was only in the site that had been fenced for 39 years that the abundance of perennial grasses increased substantially and the abundance of shrubs decreased. In some cases change is irreversible if, for example, the removal of plant cover causes severe erosion and the loss of topsoil (see Chapter 4) greatly alters what species are able to grow there.

## 13.6 Ecologists Have Reconsidered the Concept of Climax

The concept of the climax—the hypothesized static, deterministic end point of succession—once dominated ecologists' thinking about succession. Early ecologists thought of succession as an orderly progression of **seral** (preclimax) communities leading to a predictable end point. According to this idea, once a community has reached the climax state, it stops changing, unless a disturbance occurs that resets the community back to an earlier seral stage. There are serious difficulties with this concept, which have become increasingly apparent over the years.

The notion of a successional climax was devised by Hult (1885) and developed over the next several decades. One of its proponents was Frederick Clements (1916). His ideas may have been influenced by classical Greek philosophy, a major part of the intellectual tradition of his day, which emphasized idealized types in nature. Ecologists of the time viewed the natural state of the community as being unchanging and viewed disturbance as an unnatural, external process. As we have seen in this chapter, however, various kinds of disturbances are intrinsic to most communities. As plant ecologists have recognized this fact in recent decades, many have called into question the entire concept of the climax as the end point of succession.

The result has been a conceptual shift—to considering patterns of stasis and change at different spatial scales. Consider a landscape dotted with many different communities, each containing many different patches. A single small patch may always be in a state of flux as populations change and one species replaces another (as in van der Maarel's carousel model; see Chapter 10). At this scale, change may seem directionless and at the constant mercy of unpredictable factors, such as climate and herbivory. In contrast, the community as a whole may be undergoing a slow successional process of gradual replacement of one set of species by another. At this larger scale, disturbances may still occur every century or so, resetting the successional cycle. Finally, the entire landscape may consist of a mosaic of communities at different points along the successional cycle. Although each community is changing, the proportion of the entire

landscape at each successional stage remains about the same. Thus, a dynamic equilibrium may exist at the scale of the landscape, even though no single community is at equilibrium. We will return to this issue of hierarchies of scales when we discuss landscapes in more detail in Chapter 15.

Today, plant ecologists recognize that communities and landscapes never attain a constant, unchanging state. Rather, communities may attain a dynamic equilibrium, if an equilibrium is reached at all. Long-term climate changes (see Chapter 17), introduction or evolution of new species, and landform changes such as erosion, mountain building, and continental movements mean that the world is always in flux. At best, these changes occur slowly enough that communities or landscapes are at quasi equilibrium. Thus, one's view of change and equilibrium depends on the scale considered.

In this chapter and the previous ones, we addressed a number of questions about the nature of communities. We began by asking whether communities are entities in their own right, or merely collections of populations that happen to co-occur. We described a number of ways in which species interact. These interactions can be direct or indirect, positive or negative. For example, competition for common nutrients is a direct, negative interaction. In contrast, the stabilization of a sand dune by *Elymus canadensis*, allowing colonization by other species, is an indirect, positive interaction. These interactions help to shape communities. Succession occurs, in part, because of these interactions. Thus, a community is more than simply the sum of its constituent species. How those species interact is also important in shaping community composition and structure.

## Summary

- The process of succession has many implications for our understanding of the nature of communities.
- Succession starts with disturbance, and all communities experience disturbances.
- Those disturbances can be small (a limb falling from a tree, a badger creating a mound) or large (a forest fire, a volcano erupting).
- Disturbances can occur every year, or once in a millennium, and can be an essential factor determining community composition and dynamics.
- Disturbances can result from a multiplicity of factors—fire, wind, rain, snow, animals, disease.
- After a disturbance, new species may colonize the disturbed site. These species then interact with one another, so the community changes over time, resulting in the process of succession.
- Various processes play roles in determining successional trajectories. These include facilitation, inhibition, tolerance, and initial floristic composition.
- Within a given community type, the successional trajectory may be predictable and lead to a single end point. If disturbances are relatively infrequent, the community may reach a state of dynamic equilibrium.
- Earlier ecologists thought that a static end point to succession, called the climax, was the usual state of affairs in plant communities. There are now good reasons to think that even a state of long-term dynamic equilibrium may be exceptional.

Александр Лещёнок/CC BY-SA 4.0

14

# Local Abundance, Diversity, and Rarity

All species are rare when they first evolve; some go on to increase in numbers and spread, while others remain rare. In this chapter, we address a series of questions that raise fundamental issues in ecology and also have major implications for applied ecology and conservation. Ecologists have asked what determines the relative abundances of the individuals in a community ever since ecology was a scientific discipline. What determines how many species can coexist in a community? We also want to understand how species interact, and whether that has anything to do with their abundances. Are abundant species competitively superior to less common species, or are they more abundant for other reasons? We want to know why some species are rare, others common, and still others invasive, in order to understand the fundamental properties of ecological communities, but also for practical reasons. Are rare species more vulnerable than common species to becoming threatened or extinct due to anthropogenic environmental change? Our focus in this chapter is mainly on the abundances of the individuals in a species as well as the number of species at the scale of communities; we return to some of the same issues at larger scales in Chapter 19.

G. Evelyn Hutchinson asked a deceptively simple question in the title of his 1959 paper, "Homage to Santa Rosalia, or Why Are There So Many Kinds of Animals?" Ecologists are still asking this question, although we have learned some of the answers. We are concerned not only with the numbers of individuals, but also the number of species that can coexist in a community. We know that species richness (see Chapter 12) in a community might be regulated largely by deterministic processes, such as competition (see Chapter 10) and herbivory (see Chapter 11), or by chance processes such as weather events (see Chapter 8), or which species are able to reach

Above: *Larix siberica* (Siberian larch, Pinaceae) trees are abundant and geographically widespread deciduous conifers, here with foliage turning yellow in early Autumn in the Altai Mountains, Siberia, Russia in central Asia.

the community (see Chapter 13). How do interactions among the species in a community affect the ability of a new species—particularly an exotic invasive species—to become integrated into the community? Both biotic and abiotic factors can make a plant community more susceptible or more resistant to invasion by exotic species.

The processes that determine local abundances within a single community have been extensively studied by ecologists. Landscape and regional processes can also be important in determining the relative numbers of individuals per species, as well as the number of species in communities. In this chapter, we examine the interaction between local and regional processes from a local perspective. In Chapter 15 we look at landscape-level processes that determine population dynamics and species distributions, and in Chapter 19 we consider abundance and diversity at regional and global scales.

## 14.1 Are Dominant Species Competitively Superior?

An observer strolling through almost any plant community will notice that not all species are equally abundant. In most temperate communities, a great number of the individuals (or biomass) in a community will belong to one or a few species, while many other species will be rare. The most conspicuous and abundant species (numerically or in terms of biomass) in a plant community are called the **dominants**. Dominant species are often used to characterize plant communities: a cattail marsh, a pine barren, an oak-hickory forest, and a saguaro "forest" are all characterized by their most conspicuous or abundant plants. But what do we know about these dominants?

The term *dominant* is also used to refer to competitive superiority. Are numerically dominant species really competitively dominant as well? While plant ecologists have generally assumed that this is the case, few studies have directly tested that assumption. Differing views on the consistency of competitive hierarchies have been examined in Chapter 10. Here we consider whether numerical abundance is determined by competitive ability. Random processes, such as which species happen to occupy a site first, can also contribute to the pattern of species abundances at a site. In addition, any current pattern of abundances in a community may be temporary. For example, if a site is undergoing succession following a disturbance (see Chapter 13), one species may be more abundant than another because it tends to dominate during early succession, but it may be displaced as time passes.

Timothy Howard and Deborah Goldberg (2001) used field and garden experiments on herbaceous perennials in a Michigan old field to test whether rankings of competitive ability were related to observed species abundances in the field. They found that responses to competition at the stage of seedling germination and early growth were strongly related to species abundance. The effects of adult competitors on germination and seedling growth were small for the most abundant species, *Poa compressa* (bluegrass, Poaceae), and large for the least abundant species, *Achillea millefolium* (yarrow, Asteraceae). In contrast, the competitive ability or survivorship of adult plants had only a limited relationship to species abundance. Other researchers who have examined the effects of competitive ability at single life history stages (mostly early seedling growth) have generally reported positive relationships between abundance and competitive ranking.

Another approach to testing whether numerical dominants are really competitive dominants is to remove the dominant species from a community and observe the effects on community structure. For example, Pua Kutiel and associates (2000) found that removal of dominant shrubs on coastal dunes in Israel led to greater abundance of annual species and therefore higher diversity. David Wardle and colleagues (1999) removed the non-native $C_3$ grass *Lolium perenne* (English ryegrass, Poaceae) from a New Zealand grassland, resulting in increased species diversity. Removal of the $C_4$ grasses did not have that effect, presumably because no other species in the community were growing at the same time of year as those species (see Chapter 2). While this experimental work is an invaluable contribution to assessing what the relationships are between competitive dominance and numerical abundance in plant communities, we do not yet have enough information to generalize about those relationships or know when they hold.

### There are many ways to be rare but few ways to be common

Species differ greatly, not only in how abundant they are in any one place, but in the number of places where they are found. Some species have a wide geographic distribution, while others are **endemic** (found in only one area, usually where they first evolved). An understanding of rarity and the processes responsible for it is central to the preservation of endangered and threatened species, and also for the conservation of biodiversity from local to global scales. Not all rare species are endangered; as we will see shortly, many species have always had small population sizes, regardless of human impacts. However, endangered and threatened species are rare by definition, although they may have been common at one time. They may be rare and threatened only regionally, or everywhere. Threatened species may be protected locally, nationally, or not at all (see Chapter 19). In the United States and many other countries, the official listing of a species as endangered is a long and complex process that eventually initiates laws and regulations concerning how the species and its habitat

**TABLE 14.1** A scheme for describing eight categories of commonness and rarity based on three traits

| Geographic range | | Wide | | Narrow | |
|---|---|---|---|---|---|
| Habitat specificity | | Broad | Restricted | Broad | Restricted |
| Local abundance | Somewhere large | Common | Predictable (habitat specialists) | Unlikely | Endemic |
| | Everywhere small | Sparse | | Unlikely | Rare on all counts |

*Source*: D. Rabinowitz. 1981. In *The Biological Aspects of Rare Plant Conservation*, H. Synge (Ed.), pp. 205–217. John Wiley: Chicester, UK.

should be protected. Thus, the topic of rarity has legal, political, and economic consequences. The IUCN Red List is a global inventory of the conservation standing of plants and animals, based on expert evaluation of quantitative criteria to rank the extinction risk of species. Although the IUCN Red List does not have legal standing, the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES), is a multilateral international treaty that aims to protect endangered plants and animals.

## Being rare can vary over space and time

While a plant species may be rare or common, native or invasive, these categories are not necessarily fixed characteristics of the species. A remarkable example is *Pinus radiata* (Monterey or radiata pine, Pinaceae) (Lavery and Mead 1998). Once a rare endemic limited to a narrow strip of the California and northern Mexican coasts, Monterey pine is extensively planted for timber in South Africa, New Zealand, Chile, and other countries in the Southern Hemisphere. Pines are not native to the Southern Hemisphere. *Pinus radiata* has now escaped from timber plantations in many areas in the Southern Hemisphere to become a major invasive pest. In their native habitat, the trees do not have desirable properties for timber and are not grown commercially. Much of their native habitat has been eliminated by development, and the remaining populations, while protected, are experiencing a decline due to an introduced fungal pathogen (pitch canker, *Fusarium subglutinans* f. sp. *Pini*). Naturalized populations of this species now vastly exceed those in its native habitat. We will return to the factors hypothesized to make a species invasive after taking a look at some ecological aspects of rarity and commonness.

Provided by Cornell University

Deborah Rabinowitz

Deborah Rabinowitz (1981) famously identified three different aspects of species distributions: geographic range (wide or narrow), habitat specificity (broad or restricted), and local abundance (somewhere large or everywhere small; Table 14.1). When range, specificity and local abundance are characterized as dichotomies—either one thing or the other—these three aspects result in eight possible combinations of characteristics, and a given species could display any one of those combinations. Only one of the eight possible combinations—wide range, broad habitat specificity, and large local abundance somewhere—is classified as commonness. The other seven each constitute some form of rarity.

For example, the vast taiga of western Siberia, encompassing over 1.6 million km$^2$ of uplands and marshes, has several very abundant species and low overall species richness. The forests are dominated by just a few conifers: *Picea obovata* (Siberian spruce), *Abies sibirica* (Siberian fir), *Pinus sibirica* (Siberian pine), *P. sylvestris* (Scots pine) and *Larix sibirica* (Siberian larch, chapter introductory photo). These are quintessentially common species: they have enormous ranges and are locally abundant, dominating the areas where they are found. They are somewhat habitat specific, being excluded from the wetlands, but their upland forest habitat is very extensive. In contrast, the small herbaceous perennial *Myosotis alpestris* (alpine forget-me-not, Boraginaceae; Figure 14.1) has a remarkably broad circumboreal global distribution, with

Courtesy of Julia Postol

**Figure 14.1** *Myosotis alpestris* (alpine forget-me-not, Boraginaceae) at high elevation above Lake Louise in Banff National Park, Alberta, Canada.

populations ranging from the Rocky Mountains of the U.S. and Canada, to Alaska, eastern Asia, Siberia, the Himalayas, and west to the mountains in parts of Europe and in northern Great Britain. However, despite its wide distribution, it is found in somewhat limited habitats within its range and is generally patchily distributed rather than abundant where it is found. So, despite its very wide geographic distribution, it is "rare" or at least uncommon on two of the three measures.

While all possible combinations of local and regional patterns of commonness and rarity are conceivable, what combinations do we actually find in nature? There is often a positive relationship between abundance and geographic range (Figure 14.2): species with large population sizes tend to be widespread, while species with small geographic ranges tend to have small population sizes. Similarly, species that can live in a wide variety of habitat types tend to have large geographic ranges because in any given area there is likely to be some suitable habitat. While these relationships are the typical, exceptions exist. Some species are habitat specialists with small geographic ranges, yet they dominate the communities they are found in. An example is the mint *Mentha cervina* (Lamiaceae), a specialist of montane wetlands of the eastern Mediterranean. While it is found in a very limited range of sites in central Spain, it exists at high densities where it is found.

Tropical forests also generally do not have species that are common on all three measures. While many tropical tree species are widely distributed, usually no one species is locally dominant or abundant. Only occasionally do tropical forests have areas in which a single tree species is dominant (contributing over 50% of the individual trees in the area). Hans ter Steege and collaborators (2019), using a database of 1946 forest-inventory plots and over 5000 tree species, determined the frequency of these "monodominant" areas in Amazonia. They found that they were exceedingly rare. Only 2.6% of all plots were dominated by a single species, and even plots where over 20% of the individuals belonged to a single species were uncommon (18% of plots). Most species (97%) made up fewer than 20% of the individuals in a plot (Figure 14.3). The vast majority of Amazonian forests are highly diverse, with individuals from many species present in a site and no dominant species, except in a very small number of areas.

Commonness and rarity are the results of processes that operate at local, regional, and even global scales.

(A)

Courtesy of G. Sánchez-Colomer

(B)

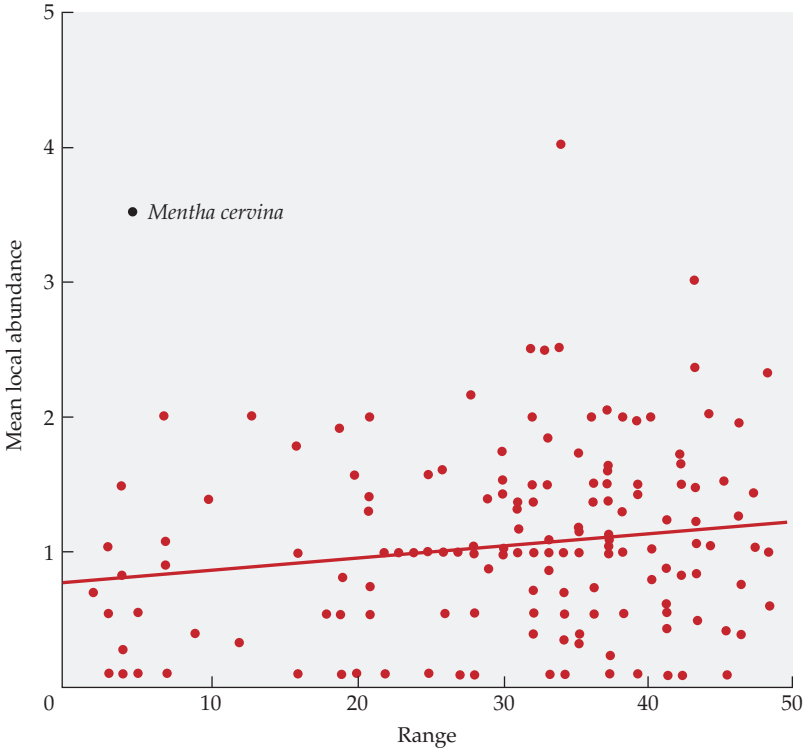



**Figure 14.2** (A) A wet meadow in the mountains of Spain dominated by sedges. Groundwater makes this meadow substantially wetter than the surrounding woodland, which is dominated by evergreen oaks and annual grasses. (B) Correlation of mean local abundance and geographic range for vascular plants in montane wet meadows in Spain ($r_s = 0.17$, $P < 0.01$). *Mentha cervina* is an exception to the general trend that species with the largest geographic ranges also have the highest abundances where they are present. (B after J. M. Rey Benayas et al. 1999. *Cons Ecol* 3: 5. )

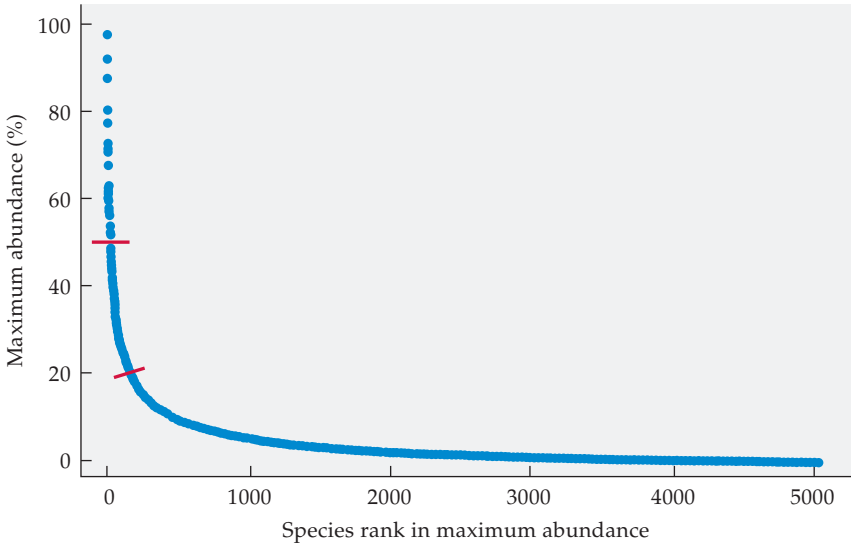



**Figure 14.3** The maximum relative abundance that each of 5029 species reached in at least one of 1946 inventory plots in Amazonia. Species are ranked from high to low maximum abundance. The top red line indicates the species that comprised at least 50% of the trees in a plot; the lower red line indicates those that comprised 20%. Dominant species were just 3% of the total. (After H. ter Steege et al. 2019. *Sci Rep* 9: 13822/CC BY 4.0.)

Local differences in species abundances may or may not be reflected at larger scales, and these patterns may change through time. Plants that were once common may decline in numbers, while rare species may, over time, become common or even dominant, particularly in new or changed environments. Local and regional patterns may be uncoupled or may be linked in a variety of ways. We will return to these issues in discussing the process of biological invasions, below.

In an analysis of the native British flora, Rabinowitz and colleagues (1986) found that 36% of species were common, with wide geographic ranges and large local abundances. Almost 60% were restricted to particular habitats, while 15% had narrow geographic ranges and 7% had small maximum population sizes (Table 14.2A). This study was the first to analyze these patterns and is often used as a basis for generalizations about rarity. However, other researchers have found that these patterns vary depending on how one looks at the flora. A study by José Rey Benayas and colleagues (1999) considered the entire set of vascular plant species inhabiting a specialized habitat type, wet meadows in the mountains of central Spain (Table 14.2B). In this study, the category with the largest number of species was the one with narrow geographic ranges, restricted habitat specificities, and small local abundances. In another study, Nigel Pitman and colleagues (1999) examined tree species in a highly diverse tropical rainforest in Manú National Park in the western Amazon basin and in the surrounding region in eastern Peru (Table 14.2C). Most of the tree species in this study were found across a wide range of habitats (fewer than 25% were habitat specialists), and all had broad geographic distributions. This study contrasts strongly with the study of the British flora by Rabinowitz and colleagues, in which many more species were habitat specialists (Ricklefs 2000).

**TABLE 14.2** Percentages of species falling into each of eight commonness and rarity categories in three studies

| (A) Vascular plants of the British Isles ($n$ = 160) | | | | |
|---|---|---|---|---|
| Geographic range | Wide | | Narrow | |
| Habitat specificity | Broad | Restricted | Broad | Restricted |
| Local abundance large | 36 | 44 | 4 | 9 |
| Local abundance small | 1 | 4 | 0 | 2 |
| **(B) Vascular plants of montane wet meadows in Spain ($n$ = 220)** | | | | |
| Geographic range | Wide | | Narrow | |
| Habitat specificity | Broad | Restricted | Broad | Restricted |
| Local abundance large | 14 | 7 | 12 | 13 |
| Local abundance small | 2 | 10 | 2 | 39 |
| **(C) Trees of a tropical rainforest in Manú National Park, Peru ($n$ = 381)** | | | | |
| Geographic range | Wide | | Narrow | |
| Habitat specificity | Broad | Restricted | Broad | Restricted |
| Local abundance large | 68 | 19 | 0 | 0 |
| Local abundance small | 7 | 6 | 0 | 0 |

*Sources*: (A) D. Rabinowitz et al. 1986. In *Conservation Biology: The Science of Scarcity and Diversity*, M. E. Soulé (Ed.), pp. 182–204. Sinauer: Sunderland, MA; (B) J. M. Rey Benayas et al. 1999. *Cons Ecol* 3: 5; (C) N. C. A. Pitman et al. 1999. *Ecology* 80: 2651–2661.

We can also consider rarity from a taxonomic perspective. A study by Mark Schwartz and Daniel Simberloff (2001) of rarity in the vascular plants of the United States and Canada found that taxonomic groups (genus, family, order, and class) with many species also tended to have higher than average numbers of rare species, while smaller groups had fewer rare species than expected. This pattern would be expected if large taxonomic groups were more likely to produce new species through allopatric speciation (see Chapter 9), in which case, on average, the new species would be isolated, have small range sizes, and hence be more likely to be rare. Conversely, groups with high rates of extinction might lose their rare species first, so small groups would be left with common species only.

Are rare plant species more vulnerable to extinction? While this has not been tested directly for plants, in a study of rare and common vertebrates, animal species considered rare with respect to all three forms of rarity showed particularly strong declines in habitats that were highly anthropogenically disturbed. Species that were common across all three measures (geographic range, habitat requirements, and abundance) increased in disturbed environments (Sykes et al. 2020), and perhaps a similar pattern might hold for plants.

### What makes a species common or rare?

The patterns described above suggest that ecological specialization may be an important cause of rarity. Species that are ecological specialists tend to have low abundances and small geographic ranges. This observation raises a broader question: Why are there both ecological specialists and ecological generalists? The traditional explanation is that there are trade-offs that result in commonness or rarity. By specializing, according to this hypothesis, a species can better exploit a particular niche. The hypothesis predicts that local abundance should be negatively correlated with the number of habitats occupied, because a specialist should do very well where it is found, but be found in only a few places. However, this prediction is contrary to what is typically observed (see Figure 14.2).

The "environmental control" or "superior organism" theory was proposed and later elaborated by James Brown (1984, 1995) to account for differences among species in abundance. Brown began with the premise that species vary in their ability to make use of all of the resources in their environment. This premise is built on G. E. Hutchinson's (1957) concept of the multidimensional niche. A niche is the set of all the conditions and resources that determine the ability of the individuals of a species to survive and reproduce. A fundamental niche is the set of conditions and resources (see Chapter 9) in the absence of competitors (or, by extension, predators, pathogens, and parasites).

Brown assumed that some species have broad fundamental niches, while others have narrow ones. The center of a species' geographic range, he proposed, is the place where it can utilize the widest range of resource combinations and, therefore, the widest number of habitats. In this region, it should be at high abundance. As the distance from this center increases, the conditions favoring the species become less common because, from its perspective, environmental conditions are deteriorating (since environmental variables tend to be correlated with one another). Therefore, the species' local habitat specificity narrows, and its abundance decreases. Eventually, the edge of its geographic range is reached. Those species that begin with a broad niche will have large geographic ranges, while those that begin with narrow niches—that is, with restricted environmental tolerances—will be rare and have small geographic ranges. This theory has similarities to the centrifugal theory of community organization (see Chapter 10).

A major limitation of Brown's theory is that it does not account for the persistence of inferior performers. According to the theory, the superior performers should simply overwhelm the inferior ones, eventually replacing them. In all likelihood, both mechanisms (the existence of trade-offs and the superior adaptations of some species) play a role in determining patterns of commonness and rarity.

Evidence regarding Brown's theory is mixed. An analysis of the abundance and occurrence of plant species in 74 landscapes across the globe found support for the theory (Scheiner and Rey-Benayas 1994). Two predictions of the theory were confirmed: first, there was a positive correlation between the number of patches a species occupied and its local abundance, and second, the most frequent class of species in any given landscape was the species found at only a single site, especially in middle- and low-latitude areas. In contrast, Mark Burgman (1989), in a test using the flora of the mallee vegetation of Western Australia, did not confirm the major predictions of Brown's theory. This region has a mediterranean climate and highly variable soil conditions. The vegetation of the region is characterized by high species richness, with many endemic species, and is dominated by low shrubs and herbaceous perennials. Burgman's analysis strongly refuted Brown's "environmental control" theory. It found no support for the hypothesis that rare species have smaller fundamental niches than do widespread, abundant species. The low abundances of the rare plants in Burgman's study were not caused by limited environmental tolerances, nor were the high abundances of the common

plants attributable to wide environmental tolerances. Burgman explained these results by saying that even if a species can tolerate a broad range of environmental conditions, if those combinations of conditions and resources are found in only limited geographic areas or patches, those species will be limited as well.

The flaw in Brown's theory, therefore, could be the assumption that the range of variation in environmental conditions is positively related to the geographic extent of the range. Imagine, for example, that a plant species can tolerate a wide range of soil textures from sandy to loamy soils but not soils high in clay, and in the environment in which it is found, there are a few patches with loamy soils, large areas with clay-loam and clay soils, but no sandy soils. Then, even though the species could do well on many different soil types, in that environment it would be restricted to a small number of patches and would be rare. The rare species in Burgman's study appeared to be rare because, although they could tolerate a wide range of environmental conditions, those conditions were found at few sites. In contrast, the conditions tolerated by the common and widespread species were found in many places in Western Australia.

Kevin Gaston (1994) and William Kunin (1997) listed a number of hypotheses that had been proposed to explain what makes a species rare. They suggested that the explanations for the immediate causes of rarity fall into three categories: ecological specialization, lack of dispersal, and historical accident. While various explanations have been proposed for why species become ecologically specialized (see also other chapters in Kunin and Gaston 1997 and Rosenzweig and Lomolino 1997), the causes of specialization are not well understood, and the evidence connecting specialization with rarity is limited.

A few authors have considered species characteristics associated with commoness and rarity. Colleen Kelly (Kelly 1996; Kelly and Woodward 1996) searched for correlates of commonness and rarity in the floras of Great Britain and Crete. She asked whether there are particular attributes of species that give them wide geographic ranges. She looked at a number of attributes, including growth form, pollination syndrome, and dispersal mechanism. Only growth form (trees were more common than shrubs) and pollination syndrome (wind-pollinated plants were more common than animal-pollinated plants) were found to explain geographic range in the two floras.

Courtesy of Colleen Kelly

Colleen Kelly

## 14.2 Biological Invasions Are a Worldwide Concern

Ecologists and environmentalists have become alarmed at the tremendous spread of invasive species in the past few decades because they are often associated with the displacement of native species and the alteration of community and ecosystem processes. Many definitions and alternative terms for **invasive species** have been proposed, but the term generally refers to species that have become established in large numbers or biomass outside of their native ranges and are self-reproducing. Most invasive species are non-natives or **exotics** that are far from their native ranges, often from other continents. Invasive species may also be native to a region but aggressively spreading there; for example, disturbance or climate change may favor rapid increase. But generally the discussion about biological invasions concerns non-natives. In almost all cases, non-native species have not arrived in their new environment by natural dispersal, but were brought either purposely or inadvertently to a new habitat by people. A large proportion of the plants that have become problematically invasive were deliberately introduced, planted, and cultivated by people (Figure 14.4). Between 2000 and 3000 non-native plant species have become established in the United States, many within the last 100–200 years. In Europe, by contrast, many non-native species were introduced and became established in medieval or even prehistoric times; these ancient introductions are called **archaeophytes** in contrast to recent introductions, sometimes called **neophytes**.

Biological invasions affect plant communities in two ways: through the effects of the spread of non-native plants on organisms, communities, and ecosystems, and through the effects of invasive pathogens and herbivores (especially insects) on native plant species (see Chapter 11). Ecologists are concerned with a number of questions regarding invasiveness. In this section we ask what makes species successful invaders and what factors make a community susceptible or resistant to invasion by exotic species. We will then consider what the impacts of invasive species are on native species and communities.

Invasiveness is related to, but not exactly the same as, being **weedy** or **ruderal**. Weeds are generally defined as noncultivated species that proliferate in agricultural settings, interfering with crop production, although the term has been used in a broader context on occasion. Weeds are often exotic species but may sometimes be natives that thrive in the conditions of high nutrient and light availability and frequent soil disturbance typical of agricultural fields. Ruderal species are those that possess

(A)

©Manfred Ruckszio/Shutterstock.com

(B)

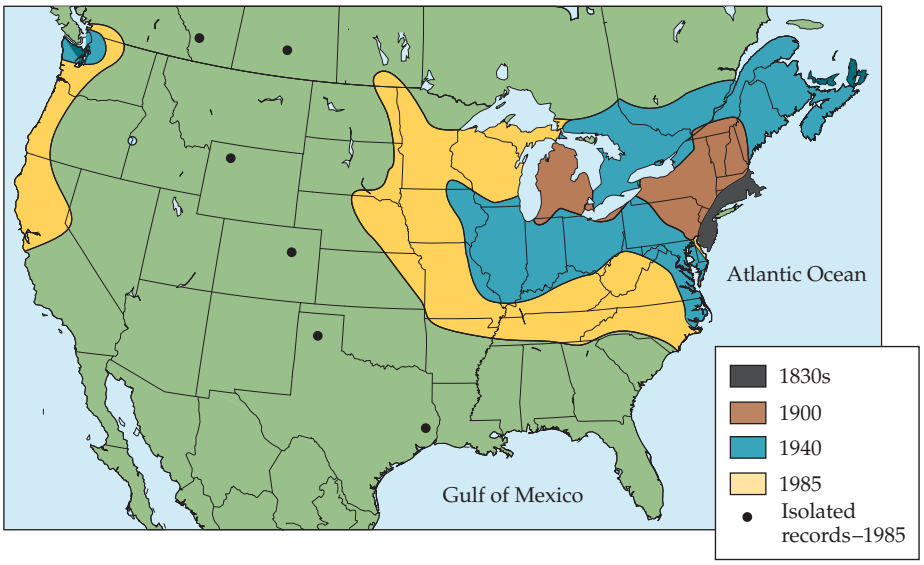



**Figure 14.4** (A) *Lythrum salicaria* (purple loosestrife, Lythraceae) is an invasive wetland species that now dominates many marshes and lakes in the northeastern United States. (B) Purple loosestrife spread across North America over a 150-year period. It may have been introduced into the U.S. in the 1700s, but it did not begin spreading for a half century or more. (After D. Q. Thompson et al. 1987. *Spread, imact and control of purple loosestrife [*L. salicaria*] in North American wetlands.* U. S. Fish and Wildlife Service: Washington, DC; R. L. Stuckey. 1980. *Bartonia* 47: 3–20.)

traits that allow them to do well in disturbed habitats, including agricultural fields and roadsides, as well as areas subject to frequent natural or anthropogenic disturbances, such as foot or automobile traffic, rockslides, or flash floods. The word *ruderal* derives from *rude*, in the sense of "wild" or "rough." While *weedy* and *ruderal* are often used synonymously, the former is typically used in agriculture and the latter is used in ecology.

## Why do some species become invasive?

For a long time ecologists have struggled to predict which plants might become invasive. Charles Elton (1958) and Herbert Baker and G. Ledyard Stebbins (1965) attempted to answer this question in the mid-twentieth century, tabulating the characteristics and attributes of species known to be successful invaders. One problem with this approach is that failed invasions tend to be unknown and unstudied, leaving us with only half the picture.

Many hypotheses have been proposed to explain biological invasions. The typical hypothesis offers a single explanation for the cause of invasions (Catford et al. 2009), although it is very likely that most invasions are a result of multiple factors. Explanations for why non-native species become invasive have included an abundance of introduced propagules (the propagule pressure hypothesis), the inherent superiority of non-native invasive species, their superior plasticity, preadaptation to the novel environment, evolutionary adaptation in the novel environment, their hybrid origins, filling an empty niche in the novel environment, and others (Lowry et al. 2013). There are many cases of each of these factors contributing to invasiveness, although generally only one factor is considered at a time.

Successes in predicting invasiveness may be particular to specific regions or systems. While many plant traits have been suggested to be correlated with invasiveness, results have often been inconclusive. Traits associated with rapid growth often appear to be correlated with the success of invasive species in disturbed, high-resource environments. For example, a study by Michelle Leishman and associates (2007) comparing plants in low-nutrient native vegetation with plants in disturbed, nutrient-enriched sites in New South Wales, Australia, found large differences in plant traits. The invasive plants that thrived in the disturbed, nutrient-enriched sites had higher specific leaf area (SLA), high foliar nitrogen (N) and phosphorus (P), and high N:P ratios. The native species that dominated low-nutrient, undisturbed sites had the lowest values for these traits.

Marcel Rejmánek and David Richardson (1996) quantified the traits of 24 species within a single genus—*Pinus*, pines—12 species that are highly invasive and 12 that have been widely planted worldwide but have never become invasive. Because the pines are closely related to one another and are ecologically and physiologically similar

in many respects, this was a powerful approach for isolating those characteristics that might be closely linked to invasiveness. The researchers analyzed ten life history traits and identified three that were most closely related to invasiveness: reproduction at younger ages, smaller seeds, and large seed crops produced at shorter intervals. Rejmánek and Richardson then applied this information to 34 other pine species that were not in the original data set but whose invasive status was known, and they accurately predicted which of those species would be invasive. They also offered a tentative set of rules for predicting woody plant invasiveness (Table 14.3).

Plant invaders may have the ability to increase rapidly for several reasons that allow them to become established and multiply in their new environment. They may escape herbivores and pathogens that control them in their native environments (see Chapter 5 and Chapter 11). Human disturbances may reduce competition from other plants, or alter resource availability (e.g., increase light and soil nutrients). "Empty niches" may exist in the new communities in which invaders may be able to proliferate. For example, *Lonicera maackii* (Amur honeysuckle, Caprifoliaceae) has become a major invasive shrub in forests in the southern U.S. that have no native shrubs.

The creation of invasive plant taxa by hybridization is an intriguing avenue for invasiveness. Hybrids between species or regionally adapted ecotypes that do not overlap spatially in their native environments but meet in the novel environment may have superior traits due to heterosis, including dominant alleles masking deleterious recessive traits, an increase in genetic variation, and other advantages (see Chapter 9). *Spartina anglica* (cordgrass, Poaceae, also called *Sporobolus anglicus*) is a major invasive species in Britain that has spread along the coastline, dominating large stretches of tidal mudflats. It is a nineteenth-century hybrid of the European native cordgrass *Spartina maritima* (small cordgrass) and the introduced North American species *Spartina alterniflora* (smooth cordgrass). In the late 1800s it was widely planted to prevent erosion, but it later spread and became highly invasive. Another case of hybridization leading to invasiveness is *Rorippa* (yellowcress, Brassicaceae) in Germany. Walter Bleeker (Bleeker 2003; Bleeker and Matthies 2005) studied the process of hybridization between the native *R. sylvestris* and the invasive *R. austriaca*. In some locations, hybrids had increased fitness relative to either parental species, while in other locations they had reduced fitness. One hybrid, *R.* × *armoracioidies*, is spreading in northern Germany. Another instance of hybridization and invasiveness is the hybrid *Bryophyllum daigremontianum* × *B. delagoense* (mother of millions, Crassulaceae), one of the most problematic invasive plants in Australia. Studies of *Fallopia japonica* (Japanese knotweed, Polygonaceae), a fast-spreading invasive plant in North America and Europe, suggest that hybridization with two congeneric species may have been important in the evolution of its invasiveness (Bailey et al. 1995).

The process by which a species makes the transition from introduction to invasive status can be thought of as consisting of several phases (Mack et al. 2000). Many invasive species have a long **lag phase** (Figure 14.5), during which their numbers are low and they are not particularly noticeable. This lag phase may be brief or may last well over a century. Because most introduced species never become invasive, disappearing during this lag phase, it is difficult to tell which introduced species are headed for extinction and which will become invasive. Populations of successful invaders increase during the lag phase even though their total numbers are still low. Many invasive plant species are introduced not once, but multiple times, and at multiple sites, during the lag phase. At some point, however, the population begins to grow rapidly and at an accelerating pace in both numbers and area occupied; at this point, the species has become invasive. Ultimately the invader reaches new geographic and ecological limits at which its population stabilizes.

**TABLE 14.3** Tentative rules for predicting which woody seed plants are likely to become invasive

| | | Opportunities for vertebrate dispersal | |
|---|---|---|---|
| *Z* | Type | Absent | Present |
| >0 | Dry fruits and seed mass >2 mg | Likely invasive | Very likely invasive |
| | Dry fruits and seed mass <2 mg | Likely invasive in wet habitats | |
| | Fleshy fruits | Unlikely invasive | Very likely invasive |
| <0 | All | Noninvasive unless dispersed by water | Possibly invasive |

*Source*: M. Rejmánek and D. M. Richardson 1996. *Ecology* 77: 1655–1661.
*Note*: $Z = 19.77 - 0.15 - 3.14 - 1.21S$, where $M$ = mean seed mass (mg), $J$ = minimum juvenile period (years), and $S$ = mean interval between large seed crops (years). $Z$ is thus a function that integrates several life history traits.

## What makes a community susceptible to invasion?

Various predictions have been made regarding the factors that make communities vulnerable to invasion. Some kinds of communities appear to be highly susceptible to

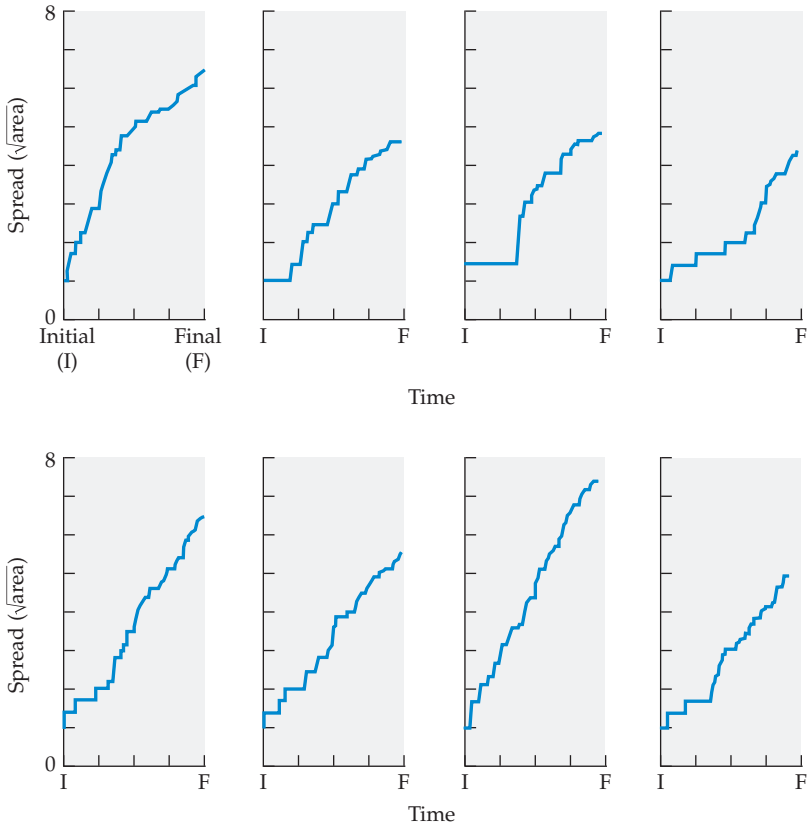



**Figure 14.5** Some outcomes for a spatially explicit simulation model of the spread of an invasive species in a new environment. In this model, about half the time, the rate of spread is slow at first and then becomes more rapid as the invading species begins to more fully occupy the new habitat. This pattern corresponds to the lag phase seen in many actual plant invasions that have been documented, in which the rate of spread is unnoticeable at first, but at some point becomes very rapid. Lags can last years, decades, or longer. (After A. Hastings. 1996. *Ecology* 77: 1675–1679.)

invasion by exotic species, while others appear better able to resist invasion. In his classic study reviewing invasions by plants and animals, Charles Elton (1958) hypothesized that invaders were most likely to become established in disturbed and species-poor communities (the biotic resistance hypothesis). Other factors that have been hypothesized to affect the propensity of a community to become invaded include the absence of natural enemies in the novel environment (the enemy escape hypothesis), complex interactions with other species in the novel environment (mutualisms, facilitation, and interactions with soil microorganisms that favor the invader and harm native species—the novel weapons hypothesis), fluctuating resources, and climate change, among others (reviewed by Catford et al. 2009 and Lowry et al. 2013).

While most ecologists continue to believe that disturbance is a critical factor in promoting invasion, there is conflicting evidence on its role, and its effects may differ among communities (Lodge 1993). In a native grassland community in Kansas, for example, Melinda Smith and Alan Knapp (1999) showed that increased disturbance (annual burning) strongly decreased invasion by exotic species. Furthermore, disturbance clearly encompasses more than a single factor; it may reflect changes in factors as diverse as soil surface microclimates, predator-prey relationships, resource availability, and competitive interactions (Orians 1986). Rigorous data appropriate for testing the importance of disturbance, or of its many different effects, are limited (Rejmánek 1989; Lodge 1993). In examining the role of disturbance in promoting or inhibiting invasion, elucidating the effects of specific factors (such as soil disturbance, logging, or fire) may be a more satisfactory approach than lumping those factors together into a general category.

Likewise, evidence about whether species-poor communities are indeed more vulnerable to invasion has been contradictory. Jonathan Levine and Carla D'Antonio (1999) evaluated data and theory on native diversity as a factor in biological invasions. They found that while both classic and recent theory predicted that more diverse communities would be more resistant to invasion (e.g., Elton 1958; Tilman 1999), data from experimental and descriptive studies were ambiguous. Some studies supported this predicted relationship and others refuted it. Two studies in grasslands showed negative relationships between vulnerability to invasion and both species and functional group diversity (Tilman et al. 1997; Symstad 2000). Other recent studies, however, have shown the opposite pattern. Thomas Stohlgren and his colleagues (1999) studied the relationship between plant species richness and invasions in the Colorado Rockies and in the central Great Plains of North America across a wide range of plant community types (including forests, meadows, and grasslands) and over a range of spatial scales. They concluded that exotic species primarily invaded areas of high species richness (Figure 14.6). In a survey of 184 sites distributed globally, Mark Lonsdale (1999) also found a positive relationship (on a landscape scale) between the degree of invasion and native species richness. The same positive relationship between native species richness and invasion by exotics holds in systems as diverse as the South African fynbos (Higgins

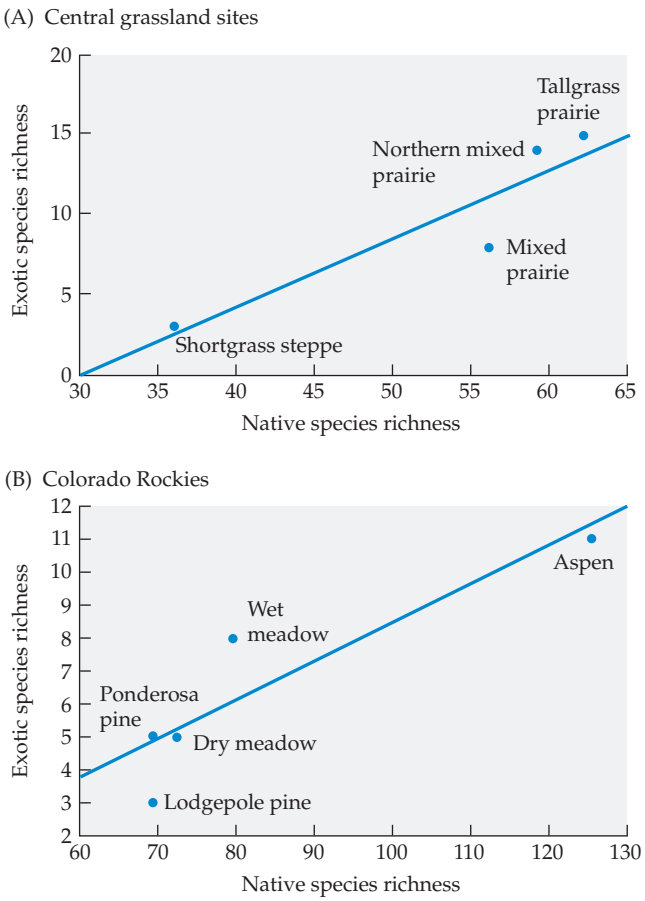



**Figure 14.6** The relationship between native species richness and exotic species richness in the Central Great Plains grasslands (A) and in the Colorado Rockies (B) of North America. The results are based on lists of species in 1000 $m^2$ plots of each vegetation type. (Results of linear regressions for grasslands were $r^2 = 0.85$, $P = 0.078$; for Rockies, $r^2 = 0.81$, $P = 0.038$.) (After T. J. Stohlgren et al. 1999. *Ecol Monogr* 69: 25–46.)

et al. 1999), New Zealand mountain beech forests (Wiser et al. 1998), Great Plains grasslands (Smith and Knapp 1999), and forests in southern New York State (Howard et al. 2004). In a review of plant invasions globally, Jonathan Levine (2000) suggested that, in general, the most diverse plant communities are the most likely to be invaded. In contrast, in a different review, David Hooper and associates (2005) concluded that greater species richness tended to decrease invasions.

It has been hypothesized that the relationship between native and invasive species richness varies with spatial scale, with a negative relationship at local scales (hypothesized to be due to biotic resistance) and a positive relationship at large spatial scales (hypothesized to be due to environmental characteristics that favor the support of both more native and more non-native species). In an effort to resolve all of these seemingly contradictory results and test the scale hypothesis, Shijia Peng and colleagues (2019) carried out a systematic review and meta-analysis of the correlation between non-native invasive species and native species in 101 observational studies over a wide range of plot sizes. They showed that there was a positive relationship between native and invasive species richness that became greater as plot size increased, but that was not negative at any size. Even for very small plots, the average correlation over all studies was not negative. However, despite the strong overall pattern, the results were very heterogeneous, and at small sizes, some individual studies showed strong negative correlations.

Variation in resource availability has been proposed as an underlying cause for the positive association between native and invasive species diversity. A study of Long Island, New York, forest communities found that soil characteristics are critical in the facilitation or inhibition of invasion by exotic species (Howard et al. 2004). There was a strong positive relationship between both nitrogen and calcium levels and plant invasions, and a strong positive relationship between native and non-native species diversity in these forests. These results echo the pioneering experimental work of Richard Hobbs (1989) in Australia, which showed that disturbance alone did not make a plant community susceptible to invasion but that disturbance coupled with the addition of soil nutrients greatly facilitated invasion. Michelle Leishman and Vivien Thomson (2005) carried out laboratory and field experiments on noninvasive and invasive species with both non-natives and natives in very-low-fertility soils near Sydney, Australia. These sites have a very high diversity of native species and few invasives, except where anthropogenic nutrients from wastewater increase soil fertility. In their experiments, non-native invasive plant species had higher survivorship and growth compared with other plant groups when soil nutrients were added, providing evidence that the success of alien invasive plants is limited by the naturally low-nutrient soils but strongly facilitated by nutrient addition.

Invaders may alter ecosystem characteristics in ways that favor their own subsequent increase as well as fostering the invasion of other species. Peter Vitousek and Lawrence Walker (1989) investigated the invasion by *Morella faya* (formerly *Myrica faya;* fire tree, Myricaceae) of the new volcanic soils in the U.S. Hawaii Volcanoes National Park on the Island of Hawaii. The soils there are extremely poor in nitrogen, and *M. faya* is a symbiotic, actinorhizal nitrogen-fixing tree (see Chapter 11). Nitrogen fixed by this symbiosis not only contributed to the rapid spread of *M. faya*, but also altered the characteristics of the entire ecosystem by dramatically changing the nitrogen cycle and increasing the availability of soil nitrogen for other species.

Another case of invasive species changing ecosystem characteristics occurs when invasive grasses increase the probability of fires, and speed up their frequency. There are many instances of non-native grass invasions in many parts of the world, including the western U.S., Australia, and Central and South America. In the northern Intermountain West and Great Basin of the western U.S. and Canada, the invasion of the Eurasian annual grass *Bromus tectorum* (cheatgrass, Poaceae) has caused numerous changes in community composition and ecosystem characteristics (Mack 1981; Knapp 1996). Perennial grasses have disappeared over enormous areas, fire frequency and intensity have increased substantially, diversity of both plant and animal species has decreased, and severe soil erosion has been exacerbated. The value of the land for both cattle grazing and wildlife has been diminished.

Another instance of an invasive grass changing ecosystems and communities is the invasion of the $C_4$ perennial grass *Schizachyrium condensatum* (bush beardgrass, Poaceae) in Hawaii (the species is widely available for horticultural plantings in the U.S. although banned from shipping to Hawaii). Since the 1960s this species and other invasive grasses have replaced native trees and shrubs in woodlands, greatly facilitating an increase in fire intensity and frequency. The fires kill the native plants, but the grasses recover and are able to spread and dominate the communities (D'Antonio and Vitousek 1992).

Invasiveness may also be related to other aspects of community composition, such as the absence of herbivores or pathogens. Both herbivores and pathogens can have profound effects on community structure (see Chapter 11). Invasive species often grow more vigorously following introduction to a new range (Baker and Stebbins 1965). A commonly invoked explanation for this phenomenon is that natural enemies (e.g., competitors, predators, and pathogens) that are present in the native range are absent in the new range (Keane and Crawley 2002; Mitchell and Power 2003; Colautti et al. 2004). Lorne Wolfe (2002) tested this idea by measuring herbivory and fungal attack on *Silene latifolia* (bladder campion, Caryophyllaceae), both in its native Europe and in North America. He found a dramatic decrease in both among North American populations (Figure 14.7). Furthermore, the North American populations had evolved so that they invested less in defensive traits such as trichomes. Instead, they grew faster and produced more flowers (Blair and Wolfe 2004).

Release from the effects of enemies may interact with local resource availability to determine invasiveness (Blumenthal 2005). The reason that these two factors may interact is that species living in high-resource habitats might be more susceptible to enemies. These fast-growing species have tissues with a high nitrogen content (from rubisco and other proteins; see Chapter 2)

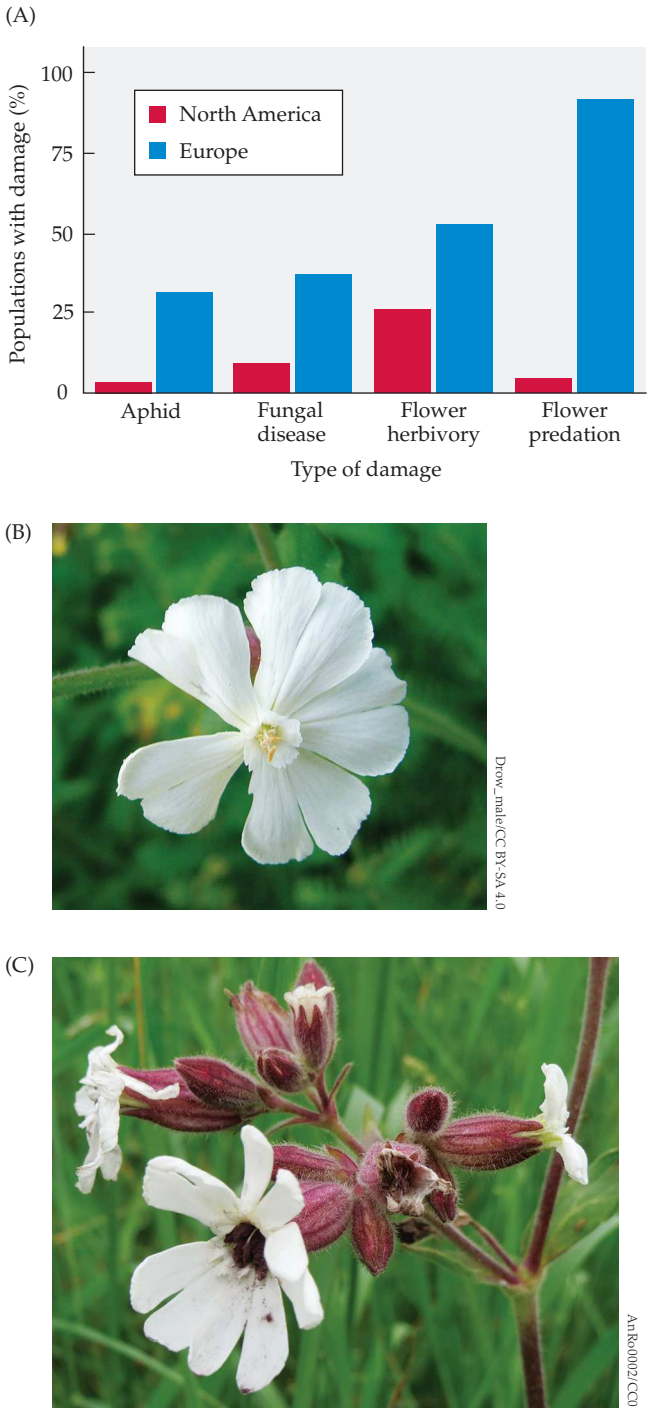



**Figure 14.7** (A) The percentage of populations found with phloem-feeding aphids, infections by the anther smut *Microbotryum violaceum*, herbivory on flower parts, or predation on fruits or seeds in 50 European and 36 North American populations of *Silene latifolia* (white campion, Caryophyllaceae). In all cases, enemy attack was substantially lower on North American populations ($P < 0.01$). (B) A flower of *S. latifolia*. (C) *S. latifolia* flowers infected by anther smut fungi; notice the center of the flower blackened by fungal spores. (A after L. M. Wolfe. 2002. *Am Nat* 160: 705–711.)

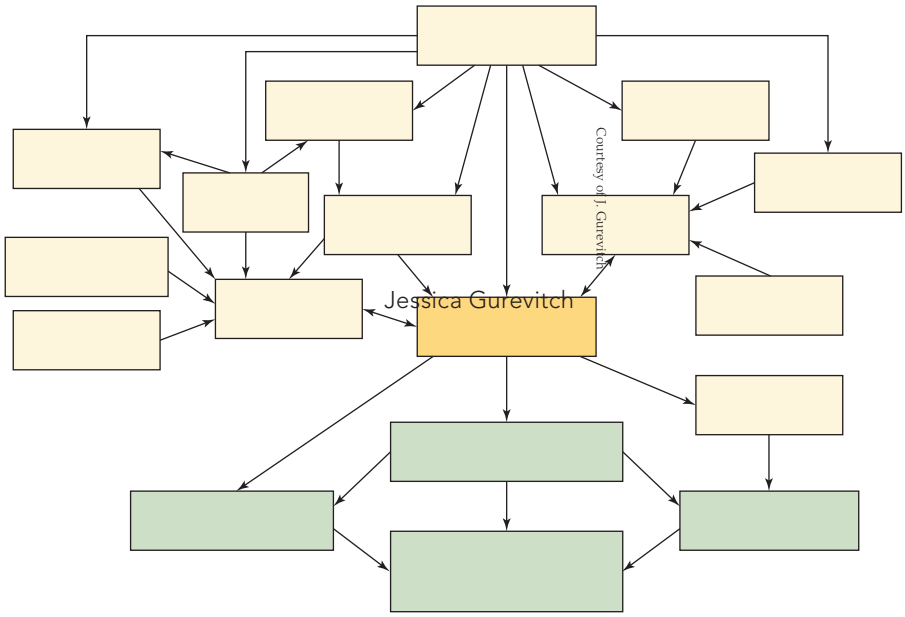
that are very attractive to herbivores. At the same time, they tend to grow fast and invest less in defenses. In their native range, these species are controlled by herbivores. In their introduced range, however, their rapid growth makes them highly invasive.

### Efforts have been made to integrate explanations for invasiveness and susceptibility to invasion

Biological invasions are most obvious and most problematic when they spread from local to larger spatial extents. By chance alone, a species may be introduced to a new site to which it is well adapted and become highly abundant, as the Monterey pine has done in the Southern Hemisphere. Many other species are introduced, and even widely planted, and may not become invasive, or they may become invasive after a lag period. Examples of the introduction of potential invaders are abundant. The species *Larix siberica* (Siberian larch) discussed above as a dominant and geographically broadly distributed species in its native habitat in Eurasia is being widely planted in the western U.S. and Canada for lumber. Will it become invasive? We do not know, but that certainly is conceivable. A species that is not invasive now, if introduced and planted in large numbers, might eventually grow in population numbers, spread, and become abundant across a broad region. Other widely planted species appear to never become invasive. *Syringa vulgaris* (common lilac, Oleaceae) is native to eastern Europe and elsewhere in temperate Eurasia. It was first introduced to North America in the seventeenth or eighteenth century; George Washington and Thomas Jefferson recorded planting lilacs in their gardens. Although it has been long and very abundantly planted, and produces viable seed, it has never become invasive in North America. It is unclear why that is the case.

When an introduced species transitions to becoming invasive, events at a local level spread to a regional scale. Conversely, processes at a regional scale (such as gene flow among sites; see Chapter 9) may result in a species becoming adapted to varying conditions across a wide geographic range. A high average level of adaptation, coupled with metapopulation and landscape demographic processes (see Chapter 15), may result in the species reaching high abundances everywhere it occurs. In this manner, regional-scale processes can determine local-scale patterns. However, in many other cases, this does not happen at all and a species does not increase or spread beyond where it was planted; ecologists do not always know and cannot always predict which species will spread and which will not. Horticulturalists and foresters generally do not consider the potential for spread in choosing to plant non-native species, focusing instead on fast growth, high survival rates, lumber or decorative qualities, and other nonecological considerations.

How can the various hypotheses and explanations for invasiveness and for community invasibility be integrated? Jessica Gurevitch and collaborators (2011) proposed a conceptual framework that puts the demography of the increase of the invasive species in a central role (Figure 14.8), affected by organism traits, acted on by evolutionary change and by population interactions (such as with herbivores or competitors) occurring in the context of the regional species pool and other factors, and resulting in rapid population increase, range expansion, and the establishment of dominance. By situating the study of biological invasions simultaneously within population, community, and ecosystem ecology and evolutionary biology, this conceptual framework integrates invasion biology into a more general understanding of fundamental biological processes.




Landscape processes

Abiotic and climate

stem
sses

## Invasive species may alter many community properties and threaten biodiversity

The spread of exotic and invasive species is increasingly threatening many ecosystems, although some communities and regions have few invasives. Biological invasions have serious consequences, both for humans and for the preservation of native species and natural communities (Lodge 1993). Human activities such as deliberate and accidental species introductions, forest fragmentation, and logging have been factors in the invasion and spread of non-native species. Jeffrey Dukes and Hal Mooney (1999) raised the provocative suggestion that global climate change, including nitrogen deposition and rising atmospheric $CO_2$ levels, may preferentially enhance the success of invaders, but little has been concluded in the decades since that paper was published (Hellmann et al. 2008). While land use change remains a dominant factor in facilitating biological invasions, many land and natural resource managers are concerned about including information on the effects of climate change for invasive species management, although they report that their organizations are not well prepared to do so (Beaury et al. 2020). Better communication between researchers and managers about predictions for the spread of invasive species and their response to climate change is important but not often realized.

Courtesy of S. Scheiner

**Figure 14.9** A hillside in New Zealand covered with the invasive species *Ulex europaeus* (gorse, Fabaceae). It was introduced from Scotland as a type of hedge but has since became a major blight. This species has also invaded the montane grasslands of Horton Plains National Park, Sri Lanka.

Exotic plant species may have substantial negative effects on the retention and restoration of native species and communities. New Zealand, for example, now has at least as many naturalized exotic plants as native species (Allan 1961; Webb et al. 1988) (Figure 14.9). These introduced species have modified the vegetation of large areas in New Zealand. While native species have not been driven to extinction there by the exotic species, the native plants appear to have been locally displaced and diminished in abundance by the invaders in many cases.

Evidence for the impacts of invasive plants on plant communities is extensive and well documented. Invasive plants have invaded forests worldwide and altered grasslands, deserts, and other biomes. Invasive plant species can decrease the diversity of native plants and disrupt ecosystem processes, altering nitrogen cycling, fire regimes, and other fundamental characteristics of ecosystems (Vitousek et al. 1996; Higgins et al. 1999; Mack and D'Antonio 2001). They can exact enormous economic costs, including decreases in timber growth rates and forest value, and are certainly altering many ecosystems and contributing to global change. A review by Petr Pyšek and colleagues (2012) suggested that invasive plant species frequently have negative effects on native plant survival and species richness, change community productivity and soil resources (often increasing soil nutrients and nutrient cycling to the detriment of native species), and that invasive grasses increase fire frequency. In a more rigorous quantitative global meta-analysis, Montserrat Vilà and collaborators (2011) concluded that invasive plant species increased productivity; negatively affected the abundance, diversity, and fitness of native plants; and altered soils and nutrient cycling. Nitrogen-fixing plants increased nitrogen pools and nitrification (see Chapter 5) but did not affect litter decomposition or other ecosystem processes. Contrary to the analysis by Pyšek and collaborators, Vilà and colleagues found no difference in effects on islands and continents.

Courtesy of Montserrat Vilà

Montserrat Vilà

Many ecologists and conservationists have hypothesized that plant invasions are a major cause of plant species extinctions (Wilcove et al. 1998). The replacement of native plants with non-native horticultural species in suburban landscapes has been shown to result in reductions in bird and lepidopteran abundance and

species richness (Burghardt et al. 2009). However, other ecologists have found that the role of invasive plants in the decline and ultimate extinction of native plants is in many cases poorly documented (Sax et al. 2002; Gurevitch and Padilla 2004). Invasive plant species have often been associated with serious declines in native species populations, but other factors that harm the native plants also commonly occur at the same time. For example, damage by feral pigs and habitat destruction are often associated with plant species' threatened status in addition to the presence of invasive plants. Globally, many more plant species are threatened by grazing and trampling by domesticated livestock as by competition from non-native invasive species (Gurevitch and Padilla 2004). Without question, habitat alteration, fragmentation, and destruction have reduced biological diversity far more and caused much greater threats to the viability of native plant species than invasive plants have. It is possible that the negative effects of invasive plants have not yet fully materialized but will in the future. More well-documented, rigorous work is critically needed.

## 14.3 Species Richness and Abundances Differ Greatly among Communities

As we have just seen, species can be rare or common within a community; they may be long-term or short-term dominants, or they may be invading the community. If we shift our focus from individual species to the entire plant community, we recognize that communities differ greatly in the number of species present and in the degree to which individual species are dominant. Total biomass also differs greatly among communities, as does productivity—the rate of energy flow through a trophic level. What are some of the causes of these contrasts among communities? Are variations in these different aspects of communities causally related?

### Abundance curves illustrate community structure graphically

One way of describing hierarchies of relative abundance is the use of **abundance curves**, also known as dominance-diversity curves (Figure 14.10, and see Figure 14.3). Abundance curves offer an alternative to diversity indices (see Chapter 12) for expressing the number and relative abundances of the species in a community (or sample, or stand). Because they are graphical presentations, rather than attempts at condensing everything into a single number, they avoid some of the limitations of diversity indices. Abundance curves can provide an excellent graphical contrast among different kinds of community structures—a curve representing a community with a high degree of dominance by a single species and low species richness will have a very different shape from one representing a highly diverse community without strong dominance, with many common species, and with a high total species number. However, these curves are more useful for descriptive purposes than for quantitative comparisons.

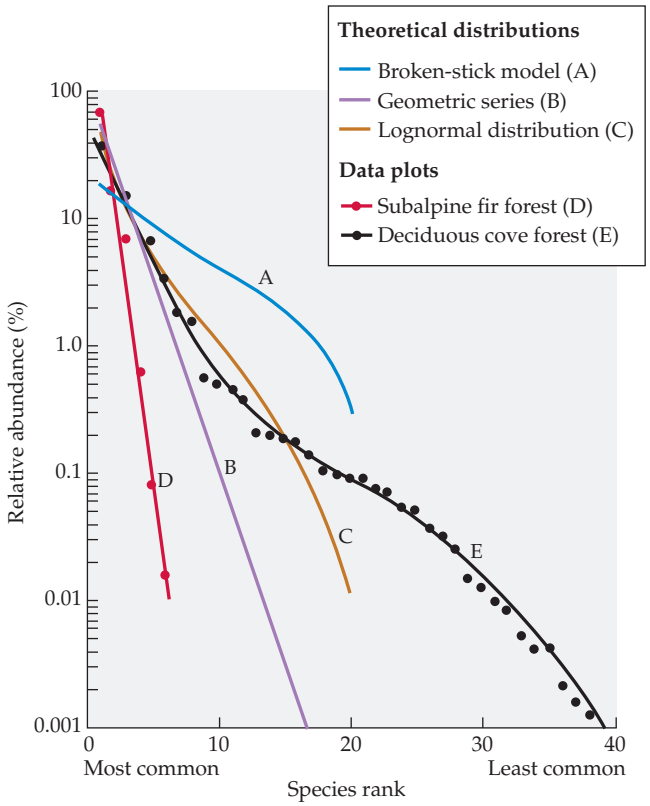



**Figure 14.10** Examples of both theoretical and empirical abundance curves. The vertical axis is the relative abundance of a species in a community—that is, the percentage of the community that it represents. Besides numbers of individuals, the relative abundance could be measured in terms of biomass, cover, frequency, or a combination of measures. As is often the case, the vertical axis is shown at a log scale. The horizontal axis arrays the species of the community in rank order, from most common to least common. The most common species has rank 1 (left side), the next most common species has rank 2, and so forth. Curves A–C represent three theoretical distributions. Curves D and E represent two sets of empirical data for vascular plant communities in the Great Smoky Mountains of Tennessee: subalpine fir forests and deciduous cove forests. (After R. H. Whittaker. 1975. *Communities and Ecosystems*. 2nd ed. Macmillan: New York.)

During the 1960s and 1970s it was suggested that abundance curves could serve as indicators of the processes responsible for community structure, but this effort proved disappointing. Robert MacArthur (1960)

proposed that the shape of a curve was created by the way in which species divided up the available energy in a system. Each species was hypothesized to monopolize a random fraction of that energy. This theory is referred to as the broken-stick model because the resulting pattern is what would be obtained if one took a stick (representing the available energy) and broke it into pieces at random (see Figure 14.12, curve A).

Others demonstrated, however, that one could obtain an identical pattern through a strictly deterministic process. Depending on the assumptions used, curves of any shape can be generated, but such assumptions are extremely difficult to verify. Assume, for example, that the most dominant species can preempt 50% of the available energy in the system, that the next most dominant species can preempt 50% of the energy that remains, and so forth. This deterministic process would produce a geometric series (see Figure 14.10, curve B). On the other hand, if dominance were determined by many independent factors that affected different species in different ways, the curve would be lognormal (see Figure 14.10, curve C).

The problem with deciding among these alternatives (random division, preemption, or many independent factors) is that all three models (broken-stick, geometric, and lognormal) could be fit to any given set of data. Therefore, it is not possible to decide which process is responsible for a given dominance pattern by examining the shape of the abundance curve for a given community. Instead, ecologists have shifted their attention to working out the processes causing dominance patterns within different communities.

Various hypotheses have been proposed to explain the observed relationships between relative species abundances and the productivity and diversity of communities. We briefly discuss some of those hypotheses here; others are reviewed throughout the rest of this section of the textbook (Part III). Hypotheses explaining differences in species diversity at local and regional scales have been reviewed by Michael Rosenzweig (1992), Michael Huston (1994), Robert Ricklefs and Dolph Schluter (1993), and David Tilman (1982). We examine patterns and possible explanations for differences in species diversity at larger spatial scales in Part IV.

## Productivity and diversity are related in complex ways within communities

Productivity can affect community structure in a variety of ways. If rates of herbivory and decomposition differ little between communities, a community with higher productivity will have a greater total biomass, or **standing crop** (the total living biomass in an area): individual plants will be larger, or there will be more of them, or both. In most terrestrial ecosystems, this strong positive correlation between standing crop and productivity sets the stage for many important processes that result in structural differences among communities.

Differences in productivity can have dramatic effects on physiognomy and community structure. If higher productivity, for example, gives rise to larger individuals, the community may have trees instead of shrubs, or shrubs instead of herbaceous perennials. We will examine these differences in more detail when we look at global vegetation patterns in Chapter 18. More subtly, higher levels of productivity can lead to larger, more vigorous individuals within a species, or to a change in the particular species with a given growth form occupying a site (e.g., a large tree such as eastern white pine, *Pinus strobus*, instead of a small tree such as jack pine, *P. banksiana*, or even the replacement of pine species with hardwoods).

Differences in productivity can also result in differences in the number and diversity of species in a community. The nature and causes of the effects of productivity on species diversity have been disputed by ecologists for many years (we visit the converse question of the effect of increased diversity on productivity later in this chapter). The nature and shape of the relationship between productivity and diversity in plant communities depend on spatial scale. Here we look at the nature of this relationship at local to regional scales. In Chapter 10 we saw how competition affects this relationship, and in Chapter 19 we examine the patterns observed at larger spatial scales.

Frank Preston (Preston 1962a,b) argued that when the total area (or, by extension, the resources) available to support species in a given region is small, then the numbers of individuals of the rarest species will be too low to maintain the existence of those species in that region. Therefore, as the available area (or resources) increases, the number of species will increase.

David Wright and his colleagues (1993) extended Preston's theory to explicitly predict that an increase in productivity should lead to an increase in species diversity. They began by assuming that greater productivity reflects a greater amount of energy available to the system, resulting in a greater number of individuals at a site. If a site can support more individuals, then more species can each have the minimum number of individuals necessary to prevent their extinction. Wright and colleagues allowed, however, that the shape of the relationship between productivity and diversity could be unimodal (a hump-shaped curve), or even negative, under certain circumstances or at smaller scales. If increased productivity resulted in larger individuals, for example, rather than a greater number of individuals, there might be no change in the number of species, or it might even decrease if the larger individuals crowded out smaller ones. Such

declines in individual numbers and species richness are typically seen after fertilization in old fields and other herbaceous communities (Figure 14.11).

Many authors have asserted that at local and regional scales, there is generally a unimodal, or hump-shaped, relationship between productivity and diversity (e.g., Grime 1973; Whittaker and Niering 1975; Huston 1979; Tilman 1982; Rosenzweig 1992; Huston and DeAngelis 1994). That is, diversity increases as productivity increases up to a maximum at intermediate productivity, but then decreases with further increases in productivity. James Grace (1999) reviewed evidence for this relationship at small spatial extents in herbaceous communities and concluded that the data were consistent with such a unimodal pattern. However, all possible relationships have been found in at least some communities (Figure 14.12). The issue of scale in productivity-diversity relationships is explored in Chapter 19.

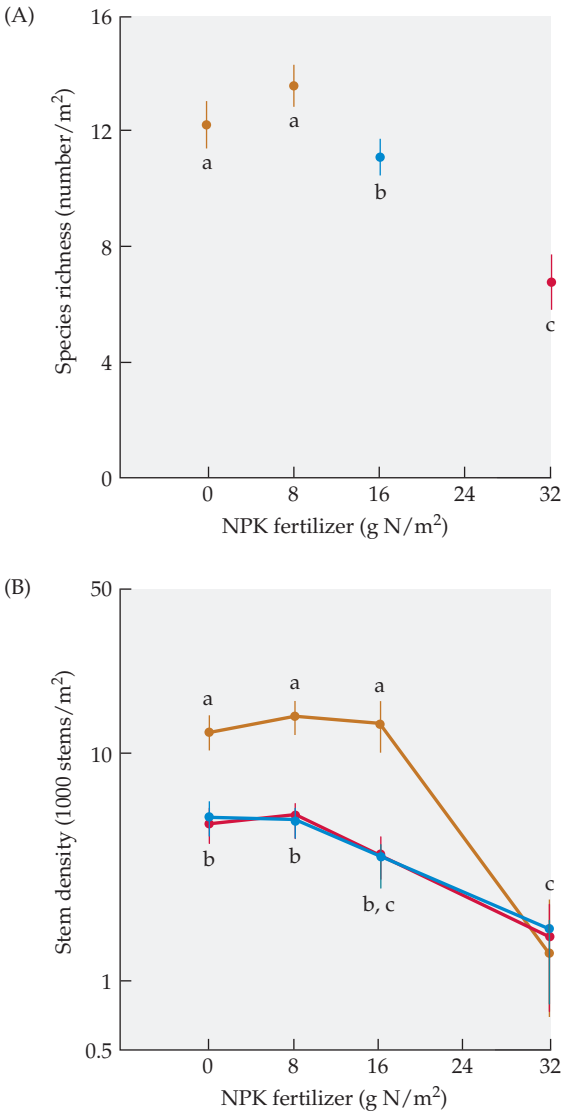



**Figure 14.11** The results of experimental fertilization of 72 plots in an old field in northwestern Pennsylvania. A slow-release NPK fertilizer was applied in May and June, and all aboveground biomass was harvested and measured in September. (A) Species richness declined by about 50% at the highest level of fertilizer treatment. (B) Stem density declined nearly 80%, with the greatest decline in the experimental block that had the highest stem density at the lowest level of fertilizer treatment. Means (bars are ± standard error) with different letters are significantly different ($P < 0.05$). (After M. H. H. Stevens and W. P. Carson. 1999. *Ecology* 80: 455–465)

Michael Rosenzweig (1971) called the pattern of decline in diversity at high productivity the "paradox of enrichment." While first proposed for predator-prey interactions, this pattern has been found for many different taxa, including terrestrial plants. Plant ecologists have tested the generality of this paradox experimentally by fertilizing plots. All of these experiments (reviewed by Grime 1973, Huston 1979, and Tilman 1982) have been done in herbaceous plant communities because these communities are likely to respond to treatment within short time frames. In most cases, a decline in diversity is found after fertilization. However, this decline occurs even when very-low-productivity sites are raised to intermediate values (Gough et al. 2000). It may be that such short-term declines in diversity are due to changes in competitive dominance among species (see Chapter 10), while an increase in diversity requires the migration of new species into experimental plots, a longer-term process.

A number of reasons have been suggested for the increases in diversity observed as productivity increases from very low to moderate levels. As mentioned above, one explanation is that more energy or resources can support more individuals, leading to a greater number of species (Wright et al. 1993). Another is that very few species are adapted to survive the difficult conditions that underlie low community productivity (such as low temperatures, low water availability, infertile soils, or short growing seasons) (Grime 1973). According to this hypothesis, as conditions become more favorable and productivity increases, more species are able to survive because the specialized adaptations necessary to survive in stressful conditions are not required. Yet another explanation is that in nutrient-poor environments, the best competitor for the limiting soil nutrient should dominate, while at intermediate levels of fertility (and productivity), spatial heterogeneity in soil nutrients should lead to the coexistence of multiple species due to niche differences among those species (Newman 1973; Tilman 1982; Tilman and Pacala 1993).

There are many contradictory explanations for the often-reported decline in diversity at high productivity (Scheiner and Willig 2005). J. Philip Grime (1973) proposed that highly competitive species dominate the most

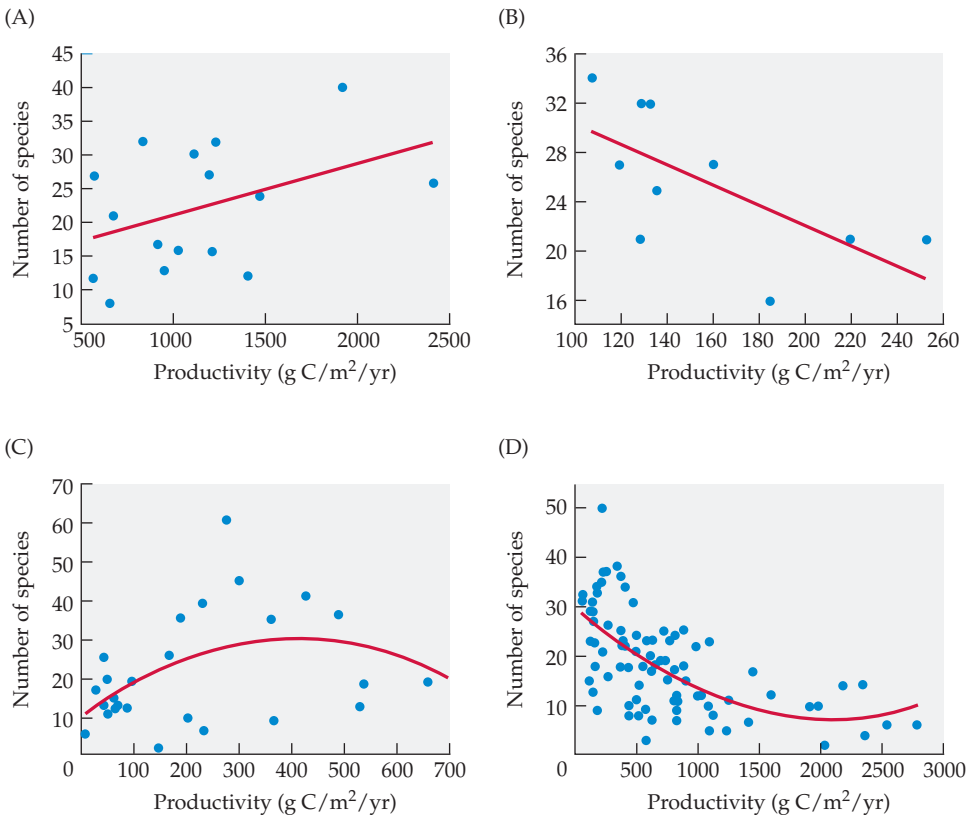



**Figure 14.12** The relationship between productivity and species richness can have different shapes in different communities. All relationships are statistically significant. (A) Aboveground yearly productivity versus total number of tree, shrub, and herb species in a deciduous forest in the Great Smoky Mountains. (B) Aboveground yearly productivity versus total number of herb species in a grassland in western North Dakota. (C) Standing crop versus total number of ground-layer species in grasslands and deciduous forests in Estonia. (D) Standing crop versus total number of herb species in a fen in the United Kingdom. (A data from R. H. Whittaker. 1966. *Ecology* 47: 103–121; B data from R. E. Redmann. 1975. *Ecol Monogr* 45: 83–106; C data from K. Zobel and J. Liira. 1997. *Oikos* 80: 325–332; D data from B. D. Wheeler and S. C. Shaw. 1991. *J Ecol* 79: 285–301.)

favorable environments and suppress other, less competitive species (see Chapter 10). David Tilman (Tilman 1982; Tilman and Pacala 1993) theorized that in the most fertile and productive environments, individual plants should be much larger than in other environments. Julie Denslow (1980) suggested that because habitats of intermediate productivity are far more common than those with very low or very high productivity, more species should have evolved to survive in the conditions presented by those intermediate habitats, and far fewer should be adapted to conditions found in environments of very high or very low productivity.

Courtesy of Nancy Larson

David Tilman

As you might gather, the question of the relationship between productivity and diversity is an active area of research (e.g., Forrester and Bauhus 2016; Jactel et al. 2018; Storch et al. 2018). Because more than one process may contribute to the observed patterns, with different processes dominating in different circumstances, resolving the causes of the patterns is difficult. While it is satisfying to resolve such controversies and reach a conclusion, it is the process of working out complex matters such as these new insights and discoveries are made, and that makes ecology an exciting and challenging science.

## Trade-offs and specialization contribute to diversity in heterogeneous environments

Why does diversity exist? Another way of asking this question is, Given that a positive relationship exists between the number of individuals and the number of species in a community, why is it that adding individuals leads to more species, rather than just more individuals of a single species?

One reason why a single species does not replace all the others in a community by outcompeting them is that different species are best adapted to different conditions. It is not possible to be best at everything simultaneously, and environments vary in both space and time. Adaptations to different conditions combined with environmental variation allow species to divide up the environment in various ways. Understory species, for example, are able to tolerate low light levels, while canopy species are superior performers under conditions of full sunlight. Moreover, different understory species are adapted to the conditions in their environments in different ways (see Chapter 2). In areas of outcrops and patchy, rocky soils, some species are able to dominate pockets of deeper soil, while others are found in the shallower soils on ridges and among boulders. Some plants are able to tolerate saline soils, or soils with heavy metals, while

others are restricted to soils free of these materials. And because plants change their environments, they may create conditions that other species can exploit.

Temporal heterogeneity may also offer opportunities for specialization. In the Sonoran Desert, which has both a winter and a summer rainy season (see Chapter 18), some annuals are specialized to germinate in response to winter rains in December, January, and February and grow in cool spring temperatures. They flower in March and are dead by May as soils become dry. Others germinate in response to summer rains in July and August, grow in the hot temperatures of summer, flower in September, and die in October as temperatures cool and soils again dry. Because these two sets of species have different phenologies, they do not compete with each other. Some species specialize on conditions that occur frequently in time or space, while others specialize on rare environments. Specialization leads to species coexistence and, thus, increased diversity in the community. Explaining the maintenance of species diversity on the basis of environmental heterogeneity and niche differentiation also may depend on differences in species' competitive abilities under different conditions (see Chapter 10).

## Disturbances might maintain community diversity

Disturbances can create gaps in mature communities, particularly in forests (see Chapter 13). Researchers have proposed two different reasons why these gaps might increase community diversity. The first is that some species might be specialized to take advantage of the different conditions afforded by gaps of different sizes (the niche partitioning hypothesis). If this is the case, certain species will be maintained in the larger community by being superior at colonizing and dominating small gaps, for example, while others will dominate intermediate or large gaps, and still others will be found only in intact forest. The second reason why gaps might increase total forest diversity is based on chance events. According to this second hypothesis, the plants that dominate any particular gap are those that happen to be there when the gap forms, or those with propagules nearby. Species coexistence in the case of niche partitioning would depend on competitive interactions, with different species being competitively superior in different gap types. Competitive exclusion would be prevented because many competitively inferior species would be able to persist by chance, essentially by being in the right place at the right time. The two hypotheses are not mutually exclusive, and both processes may play a role in a given community.

Nicholas Brokaw and Richard Busing (2000), reviewing the evidence, concluded that gaps are critical for maintaining species diversity in both temperate and tropical forests. The evidence pointed to a major role for chance in allowing species coexistence. There are clearly some species that are better than others at dispersing to and growing quickly in gaps (pioneer species). However, while there is some evidence for niche differentiation in different kinds of gaps, this evidence is much weaker than the evidence for the role of chance, and niche partitioning among gaps of different sizes probably does not play a large role in maintaining diversity in forests. Other studies have agreed with these conclusions for both tropical and temperate forests (e.g., Busing and White 1997; Brown and Jennings 1998), although Stefan Schnitzer and Walter Carson (2001) found evidence for niche partitioning in tropical forest gaps.

Gaps are not the only result of disturbance that may affect diversity. If disturbances reduce the abundance, distribution, or height and spread of competitive dominants, competitively inferior species may be able to coexist with them. But if disturbances are too frequent or severe, few species will be able to survive. This observation led to the formulation of the **intermediate disturbance hypothesis** (Connell 1978), which posits that species diversity should be highest at intermediate levels of disturbance. According to this hypothesis, competitive exclusion reduces species diversity at low levels of disturbance. When disturbances are very frequent, most existing species are destroyed and have insufficient time available for full recolonization before the next disturbance. The intermediate disturbance hypothesis has been tested in many different systems (not just plant communities), and a conceptual model was developed and formalized (reviewed by Huston 1994). In a survey of 197 studies, Robin Mackey and David Currie (2001) found that the most common pattern was no relationship between disturbance and diversity, although unimodal, positive, and negative relationships were all frequent (Figure 14.13). A subsequent review of experimental studies by Randall Hughes and her collaborators (2007) found a similar lack of generality. There are complex technical problems with the hypothesis, which

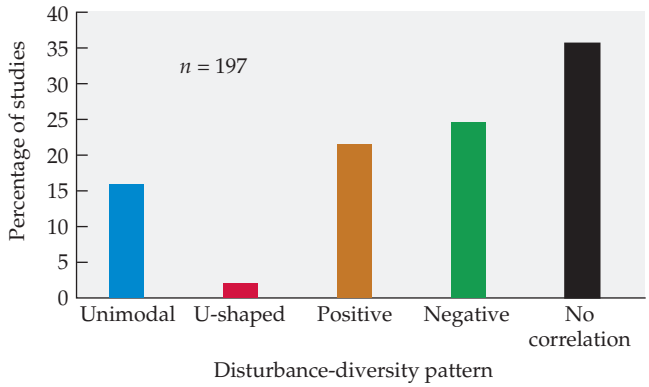



**Figure 14.13** Percentage of studies showing various relationships between disturbance and species richness. (Data from R. L. Mackey and D. J. Currie. 2001. *Ecology* 82: 3479–3492.)

limit its generality (Fox 2013), and while it is clear that disturbance promotes diversity in some cases, there is not a convincing case that diversity varies with disturbance in any simple way for most communities.

## 14.4 Does Increased Diversity Enhance Community Productivity or Stability?

In the previous section, we looked at some of the evidence for the effects of productivity on diversity. The direction of causation can also be reversed: Some ecologists argue that productivity, as well as other ecosystem properties, should be enhanced by increased diversity. (*Diversity* in this context can refer to either species diversity or diversity of functional groups; see Chapter 12.) It is easy to conceptualize how this might occur. Imagine, for example, a forest with no spring ephemerals; no species are growing actively in the early spring. If a spring ephemeral species were introduced into the forest, it would not displace any of the current species; rather, it would occupy an empty niche. It might use up nutrients that another herbaceous species growing later in the year would use. But if light were generally limiting for those species, the effect of the spring ephemeral on the growth of the later species might be small relative to the amount of new biomass it would add to the community.

This idea has been the subject of heated debate among ecologists. Shahid Naeem and his colleagues (1996), David Tilman (1996), and others have been strong supporters of the idea that an increase in biodiversity increases productivity as well as affecting other ecosystem properties. The idea was criticized by David Wardle and other ecologists (Wardle et al. 1997, 2000), who argued that increased productivity and the other ecosystem properties attributed to increased biodiversity are mostly a result of traits of the particular dominant species and functional groups present, rather than of increased diversity itself. Eventually a consensus was reached (Hooper et al. 2005) concluding that, in general, increases in species diversity do lead to increases in productivity, but as with many ecological patterns, this relationship is variable and contingent on many other factors.

The effect of diversity on productivity has been tested experimentally in herbaceous communities (old fields and prairies). A large experiment by Andy Hector and colleagues (1999) involved sowing seeds of the same species at two sites each in Germany, Greece, Ireland, Portugal, Switzerland, Sweden, and the United Kingdom. Among plots within each site, they varied the number of species and the number of functional groups. At the end of the growing season, they counted the number of species present in each plot and harvested all of the aboveground biomass to measure productivity. They found that greater diversity was related to greater productivity at all of the sites (Figure 14.14A). Most of this effect could be attributed to differences in the number of functional groups present in a plot (Figure 14.14B). Other experimental tests of the effects of species or functional group diversity on ecosystems have also been carried out (e.g., Tilman 1996; Hooper and Vitousek 1998), but the results have been subject to very different interpretations (Naeem 2000; Wardle et al. 2000). As with the effects of productivity on diversity, it is likely that the effects of diversity on productivity will depend in some as yet unknown ways on spatial and temporal scale.

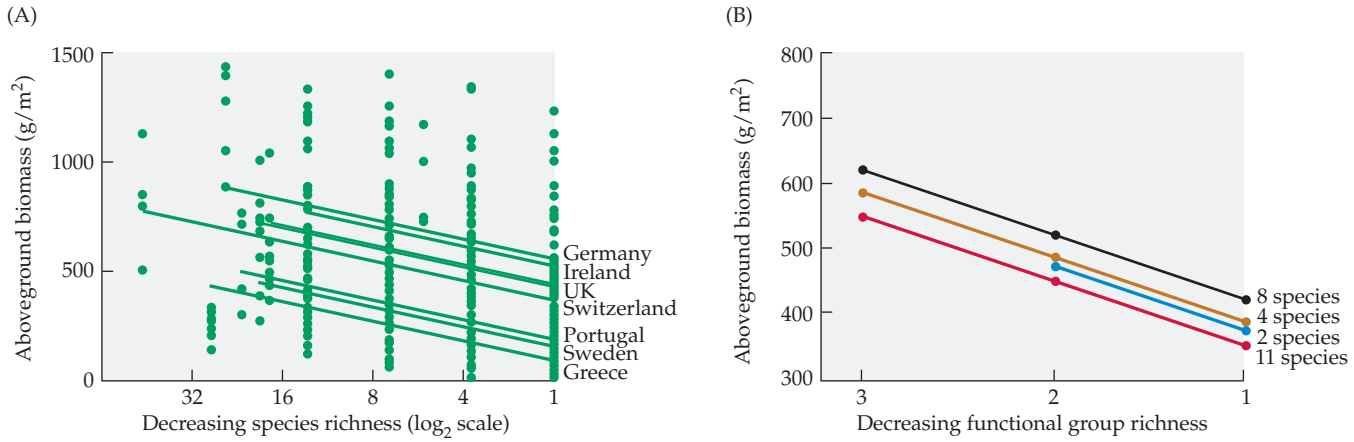



**Figure 14.14** The relationship between plant diversity and aboveground productivity in experimental plots at seven sites across Europe. (A) Species richness. The points represent individual plots; the lines are slopes from a series of regressions. (B) Biomass decreases as functional group richness declines. Assemblages with 11 species occurred only at one site in Great Britain, but those with 2, 4, and 8 species occurred at all sites. Points are averages over all sites containing a given number of species. Because the Great Britain site was less productive, the aboveground biomass for the 11-species assemblages was lower than the average of the other assemblages. (After A. Hector et al. 1999. *Science* 286: 1123–1127.)

These effects, if substantiated in other systems, could have important implications for conservation efforts. A failure to conserve large tracts of wildlands could lead to a decrease in plant species diversity (see Chapter 15). If such a diversity loss led to a decrease in primary productivity, the effects could potentially reverberate through the food chain, leading to extinctions of herbivores and carnivores.

### Community dominance and diversity can affect ecosystem processes

Is there any relationship between the characteristics of a plant community, such as diversity, and ecosystem processes such as nitrogen and phosphorus cycling? Considerable efforts have been made to determine and understand the potential connections between these important ecosystem and community properties (Tilman et al. 1997; Loreau et al. 2001; Hooper et al. 2005).

David Hooper and Peter Vitousek (1998) set out to answer this question experimentally in a California grassland. The grassland was located on serpentine soil (see Box 4A). The plants were categorized into four functional groups that differed in phenology (seasonal growth pattern), root-to-shoot ratio, and other traits relevant to their use and cycling of nitrogen and phosphorus. The four functional groups were early-season annual forbs (herbaceous plants), late-season annual forbs, nitrogen fixers, and perennial bunchgrasses. Experimental plots contained plants of a single functional group, combinations of functional groups, or no plants. The researchers measured pools and fluxes of nitrogen and phosphorus, as well as other ecosystem variables, in the experimental plots.

Hooper and Vitousek found that plant functional group diversity increased the total use of nutrients, as they had hypothesized, because of seasonal differences in plant growth and resource use among the functional groups. They also found, however, that functional group richness did not decrease nitrogen leaching, although much more nitrogen was lost when plants were totally absent than when plants of any of the functional groups were present. Ecosystem nutrient retention was determined at least as much by the indirect effects of plants on microbial activity and subsequent nutrient retention by microorganisms as by the direct effects of plant nutrient uptake. The researchers concluded that the identity of the particular plants present explained more about ecosystem nutrient cycling than could be explained by considering the number of functional groups present.

### Diversity has been hypothesized to increase stability

Yet another contentious issue of long standing in ecology is whether community diversity is related to the stability of communities through time (Hooper et al. 2005). The answer depends on which of several senses of *stability* one has in mind. In mathematics, **dynamic stability** means that a system tends to return to its original state after a small perturbation. Robert May (1973), using a mathematical model, predicted that increased diversity would have the opposite result: after perturbation, a more diverse community should be less likely to return to its original state than a less diverse community. Subsequent theoretical studies have slightly modified this conclusion; the robust theoretical result is that increasing diversity in ecological models is never likely to increase the dynamic stability of a community, but it may not decrease it.

Several ecologists have studied a quite different type of stability: they have attempted to determine whether more diverse communities have less variation in productivity from one year to the next. The hypothesis is simple: with more species present, it is more likely that in any given year, regardless of environmental fluctuations, some species will have a "good" year. In other words, the ability of different species to grow well under different conditions should mean that more diverse communities vary less in productivity (McCann 2000; Cottingham et al. 2001; Loreau et al. 2002).

In a study of prairie communities in Minnesota, David Tilman and John Downing (1994) found that more diverse communities (those with a greater number of species) maintained higher productivity, even during a period of drought (Figure 14.15). Some species were better adapted to low water conditions than others; in a drought year, those species constituted a greater percentage of the total biomass than in other years. Their biomass increased when drought caused other species to grow poorly. Ecologists have debated whether this finding represents an important emergent property of communities (Tilman 1997) or a statistical inevitability of sampling larger groups of species (Doak et al. 1998).

### Diversity, rarity, and commonness vary with spatial extent

In this chapter, we have focused on the role of local processes in determining community structure. However, local diversity, dominance, and abundances might be determined not only by the interactions among species within a site, but also by regional processes. If a species is common throughout a region, for example, then a continual influx of migrants into a site might keep its densities high. We explore this process in more detail when we look at metapopulation models in Chapter 15.

Similarly, the number of species found at a site might depend, in part, on the regional species pool. If we compare the number of tree species in deciduous forests in the eastern United States, Europe, and China, for example, we see a curious result. Although we can find sites with nearly identical climates and soil conditions, we will not find the same numbers of species. Species

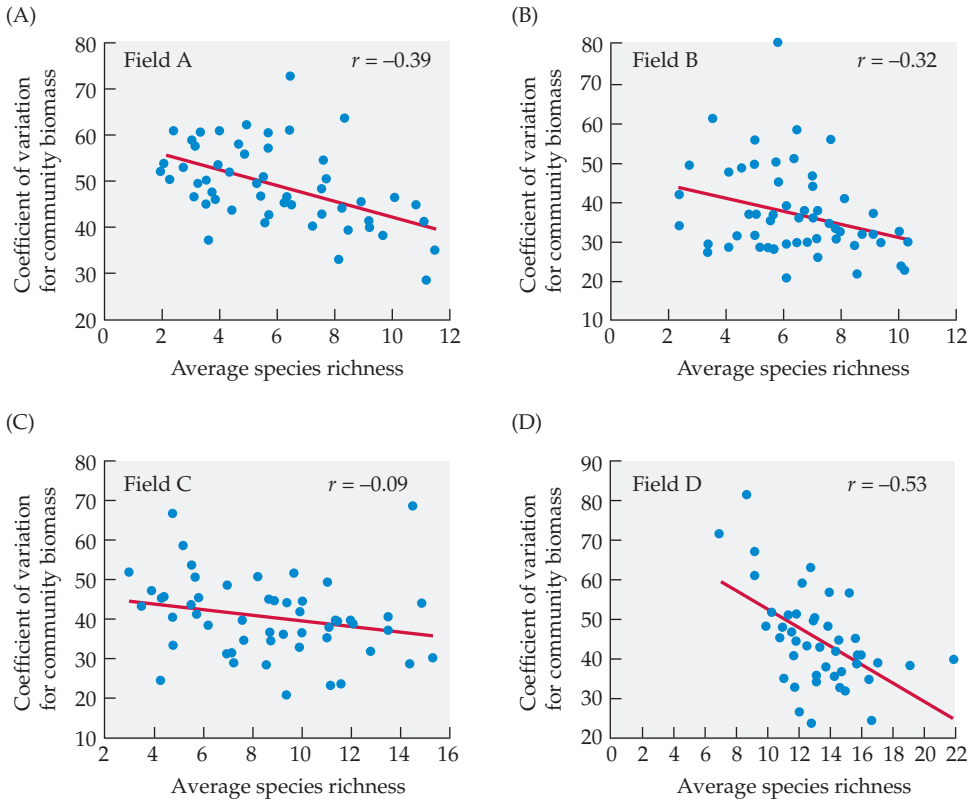



**Figure 14.15** Relationship between plant species richness and stability of productivity (variability in total biomass among years) for 4 × 4 m² plots in four fields in a prairie in Minnesota from 1984 to 1994. A severe drought occurred in 1987–1988. (After D. Tilman. 1996. *Ecology* 77: 350–363.)

richness will be highest in China, somewhat lower in the United States, and lowest in Europe. These local differences reflect differences in the regional species pool. Europe has the fewest tree species because of extinctions during the Pleistocene glaciations. China, on the other hand, has the most species because its temperate zone is adjacent to tropical land areas that have been a source of new species over geological time (Qian and Ricklefs 1999). These regional differences in species richness are a result of nonequilibrium processes that operate over very long time scales. Determining how local and regional patterns and processes are linked is currently one of the grand challenges in ecology (Lawton 1999). We take up these issues in more depth in Chapter 19.

## Summary

- In most communities, most species are rare, and only a few are common.
- The most conspicuous or numerically abundant species in a plant community are called the dominants.
- Species may be rare in different ways and for different reasons, and a species' status as rare or common can sometimes change dramatically over time.
- Rare plant species can be limited in terms of geographic distribution, local population size, or range of habitats occupied.
- A common explanation for rarity—specialization and narrow environmental tolerances—does not appear to be a good general explanation for why species are rare.
- Various hypotheses have been proposed for the ability of some plant species to become invasive and for communities to be invasible, but multiple factors rather than single explanations are likely to contribute to both invasiveness and invasibility.
- Invasive species have a range of effects on native species, communities, and ecosystems.
- Changes in productivity can alter the richness and diversity of species in a community.
- Species diversity can be related to the frequency and severity of disturbance, and gaps created by disturbance are important in maintaining species diversity in temperate and tropical forests.
- The question of the effects of species diversity on ecosystem properties is still under active investigation and remains controversial.

# PART IV

# From Landscapes to Planet Earth

Courtesy of S. Scheiner


**15** Landscapes: Pattern and Scale 421

**16** Climate, Plants, and Climate Change 447

**17** Paleoecology 495

**18** Biomes and Physiognomy 513

**19** Global Biodiversity Patterns, Loss, and Conservation 543

Courtesy of S. Scheiner

15

# Landscapes: Pattern and Scale

Communities exist as components of a larger landscape. The populations that make up a community usually are linked to populations of the same species in other communities by migration (see Chapter 8). Communities are also linked by the movement of air, water, and nutrients (see Chapter 5). **Landscape ecology** is the study of the spatial distributions of individuals, populations, and communities and the causes and consequences of those spatial patterns. In this chapter, we examine ecological patterns and processes related to spatial distributions and how they vary with spatial scale. Science involves the search for patterns and for the processes that cause those patterns (see Chapter 1). What might seem to be a third activity—prediction—depends on having identified the processes that explain the observed patterns. These two aspects of science are especially apparent in landscape ecology.

How can we objectively describe and quantify differences in community composition and structure? How are species distributed among communities? Can we relate these differences to variation in the environment? These types of questions and analyses are some of the oldest in plant ecology. Comparing communities across a landscape and looking for environmental causes of differences and similarities goes back to the origins of plant ecology at the end of the nineteenth century. While technological advances have enhanced ecologists' ability to collect and analyze complex data to use in making these comparisons, many of the basic underlying questions and approaches have changed little in more than 100 years.

While the roots of landscape ecology go back to the origins of ecology, it gained its current identity only in the 1980s. The term *landscape ecology* was coined by Carl Troll, a German geographer, in 1939. Until 1980, the field was mostly confined to Europe, where it was most closely allied with the floristic-sociological approach of

Above: A lava field on the island of Hawai'i.

Braun-Blanquet and emphasized fixed patterns (Naveh and Lieberman 1994). In the early 1980s, Richard T. T. Forman (1995) and others brought landscape ecology to North America, modifying the European tradition by adding the study of dynamics and of spatial scale (Turner 2005). The theoretical roots of some of these efforts go back to the 1960s, to island biogeography theory and metapopulation theory (discussed in this chapter). The coalescence of landscape ecology into its modern form has also relied on contemporary developments in the field of geography. Other threads that have contributed to the emergence of landscape ecology in the last two decades include the development of spatially explicit statistics and modeling (Liebhold and Gurevitch 2002; Wagner and Fortin 2005; Dale and Fortin 2014), as well as geographic information systems (GIS), and advances in satellite imagery and other means of remote sensing and analysis. As with much of ecology, new perspectives, methodology, and conceptual advances continue to develop (Fortin and Agrawal 2005; Lausch et al. 2015).

Many ecological phenomena are best studied at the landscape level. For example, fires—even very intense fires—can be quite patchy on a local scale: severely damaged trees may be found near lightly scorched or unburned trees. As burn intensities increase, new spatial patterns emerge: intense fires tend to "spot"when flying embers ignite material at a distance from the initial fire, and complex fire behavior can lead to fire vortices and other strange phenomena with spatial consequences. The spatial arrangements of individual trees, groups of trees, lakes and rivers, and topography, along with the condition of the vegetation and atmosphere, all combine to determine fire spread and intensity. Thus, we can gain unique insights into fire ecology and the resulting responses of plant populations by studying how the elements in a landscape are arranged.

Landscape ecology is concerned with questions of pattern and process. Typical pattern questions are, How are individuals, species, and communities distributed within a landscape? How many types of patches are there in a landscape? What are the sizes, shapes, and distributions of those patches? Beyond describing these patterns, we want to know what are the processes responsible, and what outcomes are caused by these patterns. We might ask how spatial patterns affect the movement of individuals, materials, and energy among patches and across the landscape. How do different ecological processes and patterns change as the size of the area changes? Earlier generations of ecologists assumed that one could get reasonable answers to ecological questions by taking averages over the landscape; today, one of the ultimate questions for landscape ecology is, When does spatial pattern matter, and when can we ignore it? The answers to these questions have important implications for conservation and landscape management. The design of nature reserves, for example, is sometimes explicitly based on landscape ecology theory. Realistically, political and economic realities often intrude on scientific considerations in the establishment and management of nature reserves; nevertheless, it is important to know what one is aiming for in making management decisions. We return to issues of conservation and management at the end of this chapter.

The name *landscape ecology* can be misleading—landscape ecology does not necessarily concern large (to humans) areas of land, but rather, it involves questions of spatial pattern and scale at any extent. For example, Christopher Foster and colleagues (2019) asked how local abundance patterns of flower-visiting beetles respond to landscape variation in southern England. They determined which elements in this urban and agricultural mosaic landscape affected habitat suitability for the beetles and how those elements shape the beetle community at extents of single meters to hundreds of meters. Landscape connectivity (below) did not shape the community, but gardens were found to be important for maintaining beetle biodiversity. Typically, however, studies in landscape ecology are concerned with spatial extents of hundreds to many thousands of square kilometers.

## 15.1 Understanding Scale Is Critical to Understanding Ecological Processes

### Patterns and processes can vary with scale

One of the most important ways in which landscape ecology has influenced people's thinking across the entire field of ecology is the emphasis on scale, which now permeates ecology. How do we specify (and think about) scale? Scale is defined by the area, volume, or duration of the data (Figure 15.1). It is composed of the elements of grain, extent, and focus. **Grain** is the minimum **resolution** of the data, for example, a 10 x 10 m square, a 1 $km^2$ area, or a 1 week period in which or during which species are counted. The sampling unit is the area, volume, or duration over which the data are actually collected. While the grain is often equal to the sampling unit, comparisons among data sets may require standardizing the data to a common grain, either extrapolating to a smaller grain or aggregating to a larger grain. **Extent** is the largest spatial or temporal scope of the observations. The extent of a study might be across a local region of 100 $km^2$ or across the entire globe. **Focus** is the way in which grains are aggregated or summed (more on this below). Often ecologists use

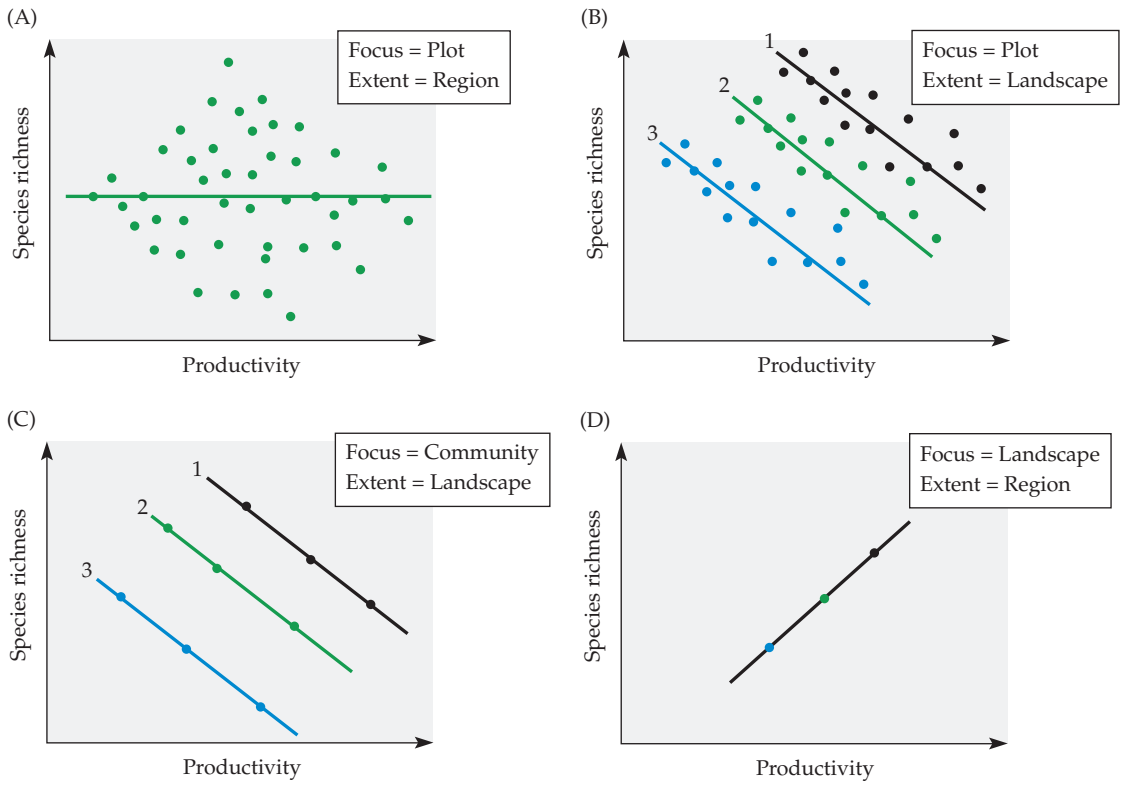



**Figure 15.1** An illustration of how changing the focus and extent of a study can alter the pattern observed—in this case, the relationship between species richness and productivity. Each diagram represents the same region. In this region there are three landscapes, with three communities in each landscape and five sampling plots in each community. In all cases, the grain of the study remains the plot. Species richness is measured as the total number of species in each plot. (A) Each point represents one plot, so the focus is the plot, and the extent is the entire region. An analysis of the relationship between species richness and productivity results in a slope of zero, or no relationship. (B) Each point represents one plot, so the focus is the plot. Each landscape is analyzed separately, so the extent is the landscape. Analyses of the relationship between species richness and productivity for each landscape result in a negative slope. (C) Each point represents the average of the plots in a community, so the focus is the community. Each landscape is analyzed separately, so the extent is the landscape. Analyses of the relationship between species richness and productivity for each landscape result in a negative slope. (D) Each point represents the average of the plots in a landscape, so the focus is the landscape. The extent is the entire region, and the relationship between species richness and productivity is positive. (From S. M. Scheiner et al. 2000. *Evol Ecol Res* 2: 791–802.)

the shorthand of **scale** when they might specifically mean "extent" or "extent and grain," and we will use this shorthand to describe general ideas. Ecologists conventionally speak of local, landscape, regional, continental, and global scales, and they use area to distinguish these scales—for instance, local scales are often thought of as being up to $10^2$ km$^2$, landscape or regional scales as being about $10^2$ to $10^8$ km$^2$, and continental or global scales as being $10^8$ km$^2$ and larger. It is usually important, however, to define scale in terms of the particular ecological processes involved. For example, $10^2$ km$^2$ may be about the scale at which most pollen dispersal occurs for pines or some other plants with wind-dispersed pollen, but it is much too large a scale for pollen dispersal by animals such as ants, beetles, or Australian possums.

When we study an ecological pattern, we must establish the scale at which we are conducting our analysis. However, we must point out a potentially confusing difference in how people in different fields discuss scale. Ecologists generally equate scale with the physical dimensions of the phenomenon being studied: a process that occurs over 10 km is at a larger scale than one that occurs over 10 cm. We follow this practice in this textbook. Some landscape ecologists refer, instead, to "coarse scale" and "fine scale" for larger and smaller physical areas, respectively (Forman and Godron 1986). However, cartographers and remote sensing researchers use exactly the opposite terminology, in which scale is the relative size at which an object of a standard size appears in a map or photo. Thus, a world map has a much smaller scale than a regional map (that is, a meter is much smaller on a

world map). It is therefore important to be explicit about how one is using the term *scale* (e.g., Turner and Gardner 1991), and it is often useful to be clear about whether one is discussing extent, grain, or focus.

## Scale interacts with environmental heterogeneity

A great deal of recent interest in ecology has concerned how ecological patterns and processes vary as a function of the scale at which they operate and are studied (Figure 15.2). The same ecological questions can have very different answers when studied within a small local area and across a landscape or region—that is, at different spatial extents and grains. Likewise, one's perspective can change dramatically when studying an ecological process over a single growing season of a few months or over a period of decades or centuries. Different things may be going on at different scales, and expanding one's focus to more than one scale can be richly rewarding. In a study of a local community, for example, we might see that competitive interactions keep individual plants of a particular species at a distance from one another (see Chapter 10). At a larger extent, we might notice that the plants are grouped together across the landscape, because individuals that are too far apart from any others never become pollinated and fail to leave descendants or because the seeds have limited ability to disperse (see Chapter 6). At a regional or continental extent, the plants may exist in several large but separated enclaves, determined by patterns of glaciation and species migration thousands of years in the past. The adaptations of a plant species can be investigated over a range of different environments that are found within an area of a few square kilometers or over a range of the environments across an entire region. One's understanding of the global range of a species may depend on knowing both the history of that species—its distribution during the last glacial period (see Chapter 17)—and competitive interactions within communities.

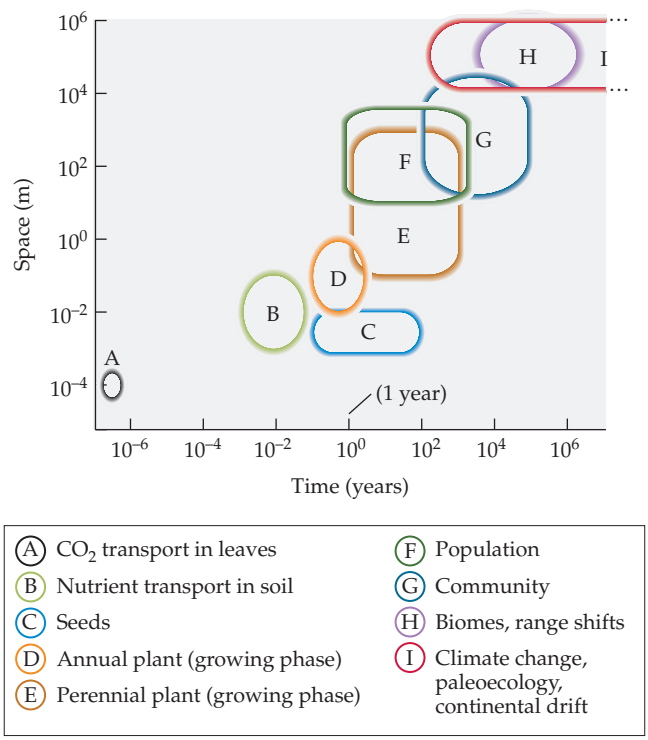



**Figure 15.2** Ecologists study patterns and processes across a wide range of scales in space and time. The processes of plant physiology, such as the diffusion of $CO_2$ molecules in a leaf, occur over the shortest distances ($10^{-4}$ m) and times ($10^{-7}$ years). Moving up the hierarchy, the domain of the whole plant and its birth, growth, reproduction, and death encompasses slightly longer distances (100 m) and times (100 years). At yet larger scales ($10^1$–$10^2$ m and $10^1$–$10^2$ years), we enter the realm of populations and communities and their changes over the course of years and decades. Finally, ecologists study patterns that stretch across the entire globe ($10^5$ m) and thousands or even millions of years ($10^3$–$10^6$ or more years).

One of the reasons that awareness of scale is increasingly central to all areas of ecology is the recognition that the world is a very heterogeneous place and that this variation cannot be ignored. Even over very small distances, conditions can change in ways that may be important to living organisms. Environmental conditions are a particular concern in plant ecology because plants cannot move—or, at least, mature terrestrial plants generally are firmly rooted in place, although their offspring may be dispersed some distance away. So, the environment immediately surrounding an individual plant is overwhelmingly important to its survival, growth, and reproduction.

The **habitat** of a population or species is the kind of environment it generally inhabits, and it includes the set of **biotic** (living) and **abiotic** (nonliving) factors that influence it in the places one usually finds it (see Part I). But the conditions in the immediate surroundings of an individual plant—its **microhabitat**—may differ considerably from the average conditions in the general habitat (Figure 15.3). Factors operating to distinguish a microhabitat from others around it include the composition of the soil; the microclimate of the immediate area; the presence, size, and identity of neighboring plants; and other organisms in the immediate surroundings (grazers, pollinators, seed eaters or dispersers, and mutualistic or pathogenic fungi or bacteria).

Similarly, the environment varies from moment to moment. Temporal heterogeneity exists at many scales and has major effects on plants. Variations in conditions, from day to night or summer to winter; across periods of wet years, cold years, or snowy years; and at longer

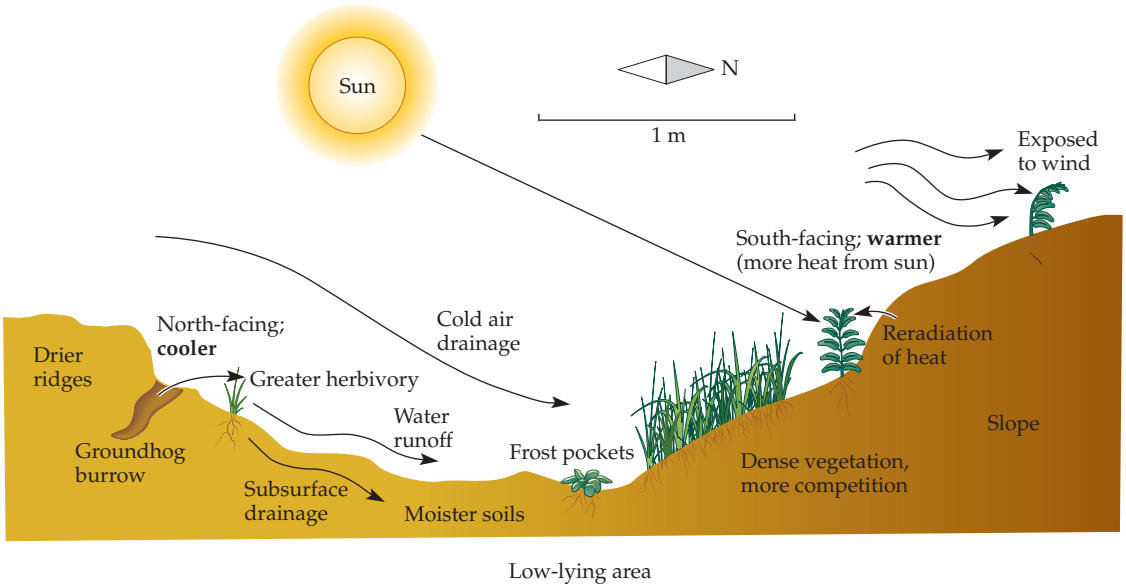



**Figure 15.3** The environment in a particular microhabitat can differ in a number of ways from conditions in the surrounding area. Individual plants experience the conditions in their immediate microhabitat, not the average conditions in the general area. The grass on the north-facing slope at the left of the diagram experiences cooler temperatures, and perhaps greater herbivory due to its proximity to the groundhog burrow. The plants in the low-lying area may experience frost sooner in autumn and later in spring than those in the surrounding areas due to cold air drainage; here, the soils are moister, and competition for light and soil moisture may be more intense, than on the ridges. Other potential effects of a microclimate are also illustrated.

scales as climate or other environmental factors change over thousands of years, all have important influences on plants. Depending on the ecological process being studied and the organisms involved, it may be the small-scale, moment-to-moment variation that matters most (such as fluctuations in light levels in a small forest gap on a partially cloudy day), or it may be large-scale, long-term average conditions (such as $CO_2$ concentration in the atmosphere), or it may be the interplay between processes occurring at different scales (such as $CO_2$ flux in a forest canopy over the course of a day or a season).

## Processes and patterns may vary as grain and extent change

At a different grain, extent, or focus, different relationships may predominate (see Figure 15.2 and also Figure 19.16 and Figure 19.17). Individual herbivores consume plants at a very local scale, but an elephant herd has effects over hundreds to thousands of meters, and a locust swarm can devour plants across an entire region. Population dynamics occur locally, but range shifts and biological invasions occur across much larger areas. The process of species extinction can occur locally, across an entire region, or globally; its implications at these different extents are clearly very different. A species that disappears locally can reappear due to dispersal from elsewhere, but global extinction is forever.

Large-scale patterns may be based on processes that operate at smaller scales, and local processes can scale up to larger areas. Large-scale patterns may simply be the sum of local-scale patterns, or there may be new phenomena that emerge across large areas. Scaling up and down from one set of patterns to another is a major current focus of research in many different areas of plant ecology, from the modeling of carbon fluxes (leaf-canopy-ecosystem-global) to determining the causes of diversity patterns.

Jonathan Levine (2000, 2001) found that at different scales, different processes explained the success of plant invasions. In northern California, *Carex nudata* (Cyperaceae), a sedge, forms tussocks (a tightly clumped growth form) along river edges. A tussock creates a microhabitat that can shelter up to 20 species of angiosperms and bryophytes during winter flooding. Along a 7 km stretch of river, Levine found that the higher the species richness of a tussock, the more likely it was to be invaded by the non-native species *Agrostis stolonifera* (creeping bentgrass, Poaceae), *Plantago major* (common plantain, Plantaginaceae), and *Cirsium arvense* (Canada thistle, Asteraceae). This pattern was the result of processes working in opposite directions at different scales. Very small-scale processes—those occurring within a single tussock—were investigated by an experimental manipulation of the number of native species on

individual tussocks and by the introduction of seeds of the invaders into the experimental tussocks. At this small scale, as species richness increased, the germination, survival, and growth of the invaders declined. In contrast, at the larger scale of the 7 km stretch of river, Levine found that seeds of all species (both natives and invaders) moved with the flow of the river, and so both native diversity and frequency of invasion increased downstream. Thus, although native diversity contributed to community resistance to invasion at very small scales, at larger scales other factors that changed with diversity (here, seed supply of invaders) were more important, creating a positive relationship between invader and native diversity.

Levine hypothesized that such scaling up might be occurring in the *Carex* system. At the beginning of the study in 1998, the river chosen for the experiment was one of the few whose banks were not flooded by high-water conditions, and thus it had one of the few stretches of *Carex* tussocks available for study. In the following year, water levels dropped severely, threatening the survival of the *Carex* tussocks on this stretch of river, while they recovered elsewhere in the region. This variation in local flooding is related to a 3- to 7-year climatic cycle, the El Niño Southern Oscillation, which affects weather conditions across the Pacific (see Chapter 16). How this temporal variation in stream flow affected invasion is not known. However, Levine speculated that such landscape-scale variations in flooding, across temporal changes in weather happening at the scale of thousands of kilometers, miht influence regional invasion success by creating greater opportunities for invasion as tussocks became disturbed by flooding followed by drought. Thus, the process of invasion in this system could be influenced by processes occurring over a range of spatial and temporal scales, sometimes acting in opposition to one another.

## Spatial pattern and scale can be analyzed using graphical and statistical methods

Given the complexity of environmental factors and biotic interactions that determine spatial patterns at different scales, how can we hope to describe or quantify these patterns? One measure of variation among sites is β-diversity (see Chapter 12). Beyond such summary measures, ecologists have developed spatially explicit, spatially implicit, and nonspatial methods of quantifying patterns. Spatially explicit methods take specific locations into account in descriptions and analyses of patterns and processes. Nonspatial methods average measurements without regard for the specific location in which they were made. Methods for pattern description that are not spatially explicit—such as ordination and some statistical approaches to quantifying spatial association—are older; we discussed some of those methods in Chapter 12.

A spatially explicit graphical technique, known as **geographic information systems (GIS)**, is a key tool in landscape ecology. GIS is a set of computational techniques that combines spatial information with details about particular locations to depict relationships between these different kinds of information ("layers" in GIS terminology). For example, we might have several sets of information that relate to each location in a landscape (Figure 15.4). One set of information might be the species composition of each patch. For each patch, we might also have information on soil type, nutrient levels,

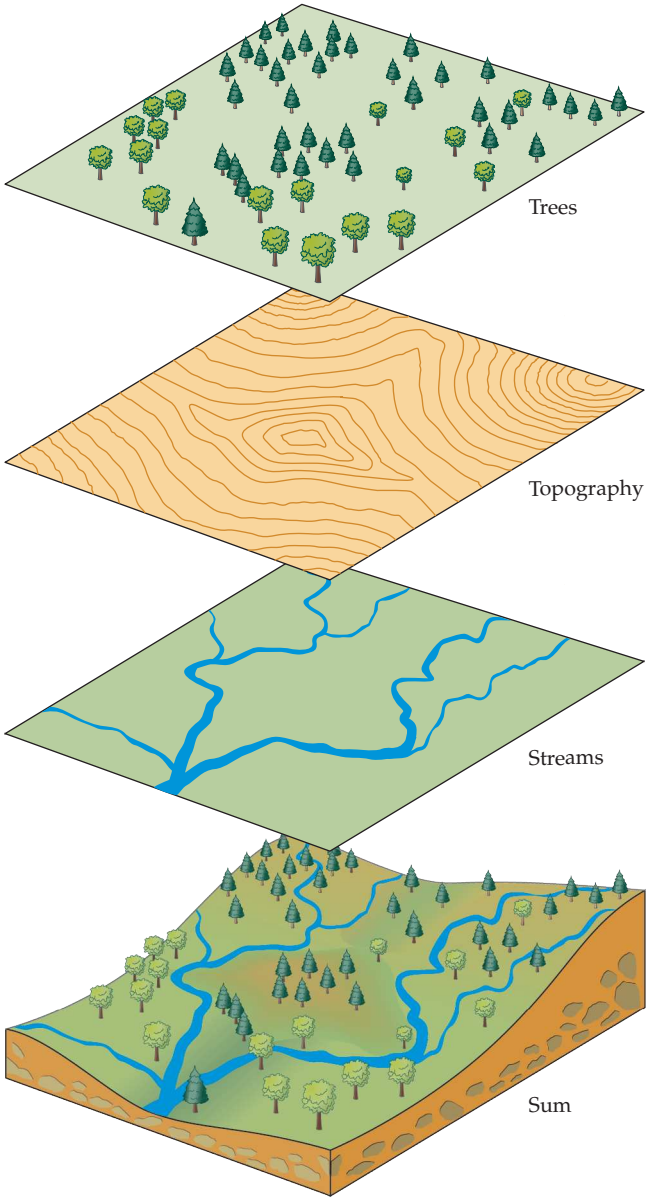



**Figure 15.4** GIS is used to examine the relationships among multiple kinds of information, which are arranged in layers.

water table depth, slope, and the last time the patch was logged or burned. Our information could also include socioeconomic information, such as land ownership, taxes on that piece of land, the number of people residing there, and the legal classification determining how the land can be used. GIS allows us to map these variables singly and in any combination to get a visual sense of potential patterns of correlation among them so that we may pose hypotheses regarding causality. GIS by itself does not allow one to quantify these relationships or statistically test hypotheses about them, although some GIS programs have added features that may do so.

Spatially based relationships can be quantified using spatial statistics. Many advances in the field of spatial statistics have been made since the late 1980s (Legendre and Fortin 1989; Wagner and Fortin 2005; Dale and Fortin 2014). These advances include the use of geostatistics, taken from the field of mining geology, and techniques for studying spatial autocorrelation, taken from the field of geography (Liebhold et al. 1993; Liebhold and Gurevitch 2002; Negrete-Yankelevich and Fox 2015). Older statistical techniques in ecology used spatially implicit or nonspatial approaches to analyze and model spatial variation in ecological data (Pielou 1977; Greig-Smith 1983). These methods were useful in distinguishing certain kinds of spatial patterns but did not take the actual spatial locations of the data points into account and had numerous limitations. The newer approaches have contributed substantially to our current understanding of the spatial aspects of organisms' responses as individuals, populations, and metapopulations (summarized by Legendre and Legendre 1998, Dale 2000, and Bélisle 2005). Curiously, spatial statistics are not often used by ecologists in conjunction with GIS, although we suspect that using them together may often add to our insights. Just as we combine graphical and statistical analyses with other types of data, together GIS and spatial statistics can constitute a remarkably powerful set of tools. The development of species distribution models, using GIS as a key tool, has been a major development in ecological research. Species distribution models have many applications in conservation and in making predictions about the ecological consequences of climate change (see Chapter 16).

We have discussed a number of the issues involved in describing and studying patterns in nature. As you have doubtless concluded, there is no single way to study, or even describe, such patterns (Levin 1992). Viewing similar questions from different vantage points (such as the spatial versus ecological scales discussed above) or at different scales can lead to very different answers. This may be frustrating if you are expecting simplistic answers, but as we have seen, considering information at more than one scale, while difficult, can also lead to new insights and understanding of the underlying biology. We now turn to some approaches to studying the processes responsible for some of the types of patterns we have just examined.

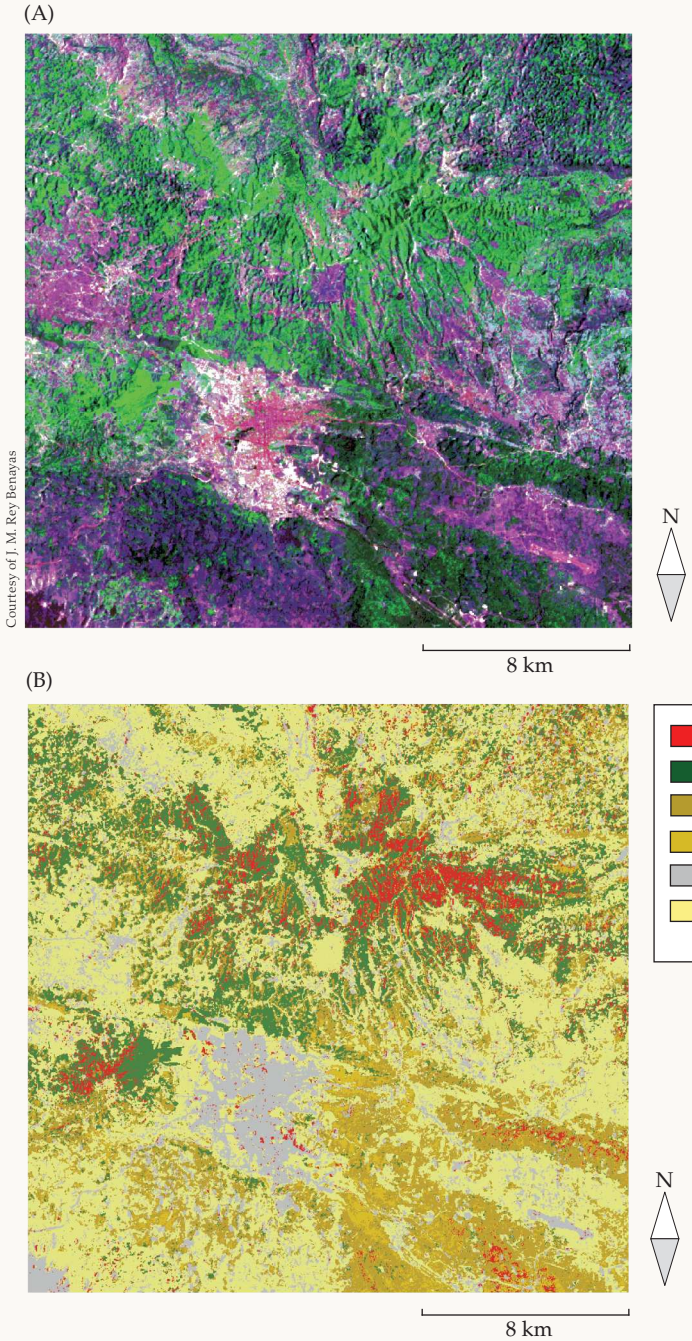
## 15.2 Landscape Ecology Involves Measuring Spatial Patterns and Looking at Their Effects

### Defining patches is a key step in measuring patterns

The basic unit we use to examine spatial patterns is the patch. A **patch** is any specified area, either an area defined arbitrarily or, often, an area that is in some way relatively homogeneous or internally consistent. A patch can be defined by naturally occurring edges (e.g., the boundary between a meadow and a forest), or a 10 × 10 m sampling unit could be considered a patch. However, if we start with such artificially defined patches, we might want to aggregate them into "natural" units. Sometimes this is done based on **compositional similarity**—the extent to which adjacent patches share a similar set of species at similar frequencies. Of course, adjacent patches usually share at least some species, so the cutoff point for when patches should be joined or not is somewhat arbitrary. Often the decision is aimed at ending up with a manageable and interpretable number of patch types.

Remote sensing has become an important tool for sampling large or inaccessible areas. In this case, the minimum patch size is determined by the resolution of the remote sensing technology—the grain. Until recently this had been 30 × 30 m; more recently available technology has provided resolutions of 1 × 1 m and even 1 × 1 cm. Our aggregation of patches is again based on their similarity, but now similarity is determined by spectral quality—the wavelengths of light reflected by the surface—rather than by direct identification of species (Box 15A). This technology has advanced swiftly in recent decades. Until recently these images were limited to satellite-based instruments, but now they can be mounted on airplanes and even drones that can be flown at low altitudes to get very detailed views of the vegetation. Increases in the resolution, number of wavelengths measured, and other technological advances may allow the assignment of individuals to species, or the measurement of traits such as the amounts of nitrogen or phosphorus or the water status of vegetation. This information can be combined with that from LiDAR (light detection and ranging), which measures distances by bouncing pulsed laser light off an object. The results are very detailed and information-rich measures of vegetation either at a

## BOX 15A Differentiating Vegetation Based on Spectral Quality




Courtesy of J. M. Rey Benayas

Remote sensing has been widely used to discriminate and map vegetation cover and community types since the 1960s. Discrimination is based on the different responses of different plant species to the various regions of the electromagnetic spectrum (microwave, visible, infrared, and radar) and to different bands within those regions. Thus, the 520 to 600 nm (green) band within the visible light region matches the green reflectance peak of vegetation, which is useful for assessing plant vigor. The 630 to 690 nm (red) band matches a chlorophyll absorption band that is important for distinguishing vegetation types. The 1550 to 1750 nm (reflected infrared) band indicates the moisture content of soil and vegetation and also provides a good contrast between vegetation types. For example, Luis Cayuela Delgado and associates (2006) classified a tropical montane landscape into developed areas, agricultural and pasture lands, and four forest types based on visible and infrared bands. Additional information can come from the radar wavelengths (0.1 to 70 cm), which provide information on vegetation density, architecture, and flooding. Nowadays, technology can provide images taken with sensors capable of registering electromagnetic signatures with many bands and with very high spatial resolution (pixels 1 × 1 $m^2$), creating the potential to count individual trees (Foody et al 2005). A new technique for measuring vegetation structure is airborne laser scanning (LiDAR), which involves firing short pulses of laser light at the ground and measuring the time it takes for the pulses to return (Lefsky et al. 2002). LiDAR can "see" through leaves and produces a three-dimensional picture of the multiple layers in a forest. That scanning can be done using airplanes or, most recently, through the use of drones.

(A) Composite color image of the region around the city of San Cristóbal de las Casas, a tropical montane region located in the state of Chiapas in southern Mexico, derived from Enhanced Thematic Mapper bands 3 (visible green), 4 (magenta), and 5 (reflected infrared, shown in red). Pixel resolution was 30 × 30 $m^2$. (B) A classification of the same landscape that distinguishes six land cover types. (B from L. Cayuela Delgado et al. 2006. *Inter J Remote Sens* 27: 1951–1971.)

very fine grain in a small area, or at a coarser analysis over a very large area. For example, Greg Asner and colleagues (2017) used airborne laser-guided imaging spectroscopy to measure seven canopy traits in tropical Peruvian forests from the eastern Andes to lowland Amazonia. Because of the small grain of the technology

used, thousands of individual treess could be identified to species level. The result was a very detailed map classified into six forest types, a map that otherwise would have been impossible to create. The resulting map could then be assessed for threats due to resource exploitation, the amounts that were within protected areas, and opportunities for additional protection.

Aggregating patches and deciding on the number of patch types is, therefore, as much an art as a science. Often it depends on the ability of measurement instruments to distinguish between different patch types. Any indirect measure (e.g., reflectance at many different wavelengths) needs to be related to the actual vegetation through a process called **ground-truthing**: an ecologist goes out to the field and samples the vegetation to determine empirically what the remote sensing devices are recording. Two communities with very different species compositions may be indistinguishable in spectral quality. In such a case, if spectral data were the only information being used to classify a large region, one would be forced to classify those two communities as if they were the same community. However, recent advances made in spectral sensors so that they are able to sample a much broader set of wavelengths has meant that it is now possible to distinguish most species with remote sensing devices.

Because landscape ecology is built on the study of patches, it is, inevitably, built on a categorical view of the world. Categorical versus continuous views of nature have been a regular source of controversy in plant ecology, at least since the contrasting views of Clements and Gleason. Thus, nearly 50 years after the Gleasonian continuum became dominant, at least among English-speaking academic ecologists, a categorical framework has proven useful at a different level of study and has emerged to dominate landscape-level studies. However, even when ecologists use a categorical framework, they may recognize that the underlying processes are continuous.

### Patches can be quantified by their sizes, shapes, and spatial arrangement

Once we have defined the nature of the patches and identified them in a particular landscape, we can measure some of their attributes. For individual patches, we can measure sizes and shapes. A typical measurement of shape is the perimeter-to-area ratio, which indicates how much edge a patch has relative to its interior. Edges between patches are important ecologically for many reasons. In particular, individuals and materials moving into or out of a patch must move across those edges.

One approach to quantifying the shapes used in landscape ecology is to calculate the fractal dimension of the landscape elements. A **fractal dimension** is a fractional dimension—for example, something in between a one-dimensional line and a two-dimensional surface, or in between a two-dimensional surface and a three-dimensional area. A classic example is the measurement of a coastline. Imagine that you first measure the coastline on a map; you conclude that the coastline is 500 km long. But you are not satisfied, because you know that there are many small bays and indentations in the coastline that were not taken into account. On a much larger and more detailed map, you now measure the coastline as 750 km, but you are still not satisfied. Walking around the actual coastline in person, you survey the distance and find that it is now measured at 1200 km. Which is correct? A mathematician named Benoit Mandelbrot pointed out that no one measurement is correct: what is correct in this case is the relationship between the scale of measurement and the distance measured. The more convoluted a coastline (or other object), the more the measured length will change as the scale of measurement changes. This relationship is called the fractal dimension of the object. Measuring the fractal dimension is sometimes very important for understanding landscape dynamics. For example, Karen Alofs and colleagues (2014) found that in central Texas, habitat fragmentation led to a decrease in species richness and that this decline happened faster in patches with a greater fractal dimension, because the edges of high-fractal-dimension objects are more convoluted and thus longer. Fractals also can have other very strange and interesting properties, including self-similarity (when the fractal pattern is repeated at many different scales).

Bruce Milne (1992) connected the concepts of fractal dimension and scaling explicitly by using fractal geometry to quantify the landscape features of a New Mexico grassland. He showed that habitat patchiness was fractal and that patchiness and other landscape features were scale-dependent. Scale dependence implies that different species might perceive and respond to the landscape differently. A beetle that lived in grass clumps might perceive a very disconnected landscape, while a grazing pronghorn might perceive that same landscape as one large patch. Some areas of the landscape were very connected both at the level of grass clumps and at larger extents, so in these areas the beetle and the pronghorn might perceive similar landscape patterns. Other areas consisted of isolated patches that might attract species that disperse readily but then settle within a single patch (e.g., small rodents).

As researchers become concerned with going beyond the individual patch, they can examine the arrangement of patches in the landscape. Using a different approach than the one taken by Milne, researchers may ask, Are patches mostly aggregated into like types or not (Figure 15.5)? A checkerboard pattern, for example, represents the maximal interdispersion of two patch types in a landscape. The maximal aggregation of two patch types would have all the patches of each type grouped into two large clumps. If there are more than two patch

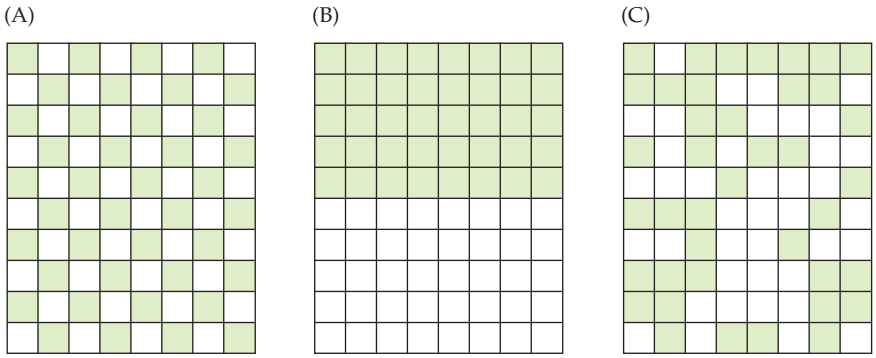



**Figure 15.5** Three different patterns of patch aggregation in a landscape with two patch types. (A) Maximal dispersion of patches. (B) Minimal dispersion of patches. (C) Random dispersion of patches.

types, additional aspects of pattern come into play. For example, in a landscape with three patch types, a patch of type A could be surrounded entirely by patches of type B, or entirely by patches of type C, or by half of each. Clearly, such second-order patterns can get very complicated with a large number of patch types. We can quantify these patterns as the **nearest neighbor probability**, the chance that any two patch types will be adjacent.

The frequency and arrangement of patches can lead to **threshold effects**—abrupt changes in the properties of systems (Turner 2005). When a patch type is rare, the patches are small and isolated, and they act as islands. When a patch type is common, the patches are large and well connected, and they act as a single system. Martha Groom (1998) studied the effects of patch size on pollination rate in the annual *Clarkia concinna* (red ribbons, Onagraceae). These plants are pollinated by various species of bees, butterflies, and bee flies. Plants in patches with ten or fewer plants received many pollinations if the patch was within 16 m of another patch, but the rate of pollination abruptly dropped nearly to zero for more isolated patches (Figure 15.6). Her work showed that dramatic change in how an organism reacts to a landscape can occur over a very narrow range.

## Spatial patterns determine many ecological processes

Accounting for spatial effects can lead to a new understanding of ecological processes. Consider the case of forests in the mountains of western North America, where three tree species are found along an environmental gradient: *Pinus ponderosa* (ponderosa pine, Pinaceae) at lower elevations, *Abies concolor* (white fir, Pinaceae) at middle elevations, and *Pinus contorta* (lodgepole pine, Pinaceae) at higher elevations. All three species grow best at middle elevations, but *A. concolor* grows faster than the other two species and outcompetes them. *P. ponderosa* is more tolerant of the drier conditions at lower elevations, while *P. contorta* is more tolerant of the cold at higher elevations. In order to understand the ecological processes that determine the species' distributions, Dean Urban (2005) created a simulation model in which he could vary both the environmental gradient and seed dispersal distance. In the simulation model, competition, environmental tolerances, and local seed dispersal all combine to reproduce the observed patterns (Figure 15.7). Of course, just because the model reproduced the observed patterns, it does not prove that those processes actually were responsible. How important seed dispersal is in general in determining community and landscape structure is still not known (Levine and Murrell 2003).

## How one analyzes landscape data affects whether the landscape appears to be continuous or discrete

Do the compositions of communities tend to blend into one another, or do they form discrete types? Answers to this question can be influenced by how one analyzes landscape data—by ordination or by classification. Ordination assumes that landscapes consist of a continuum of community types, while classification assumes that they do not. The dominance of ordination as an analytic

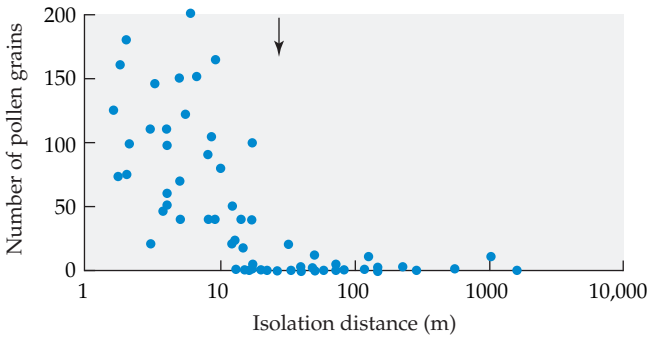



**Figure 15.6** The effect of patch isolation (distance among patches) on the rate of pollination of *Clarkia concinna*. Patches contained ten or fewer plants. The arrow indicates the threshold distance at which movement by pollinators abruptly declines, resulting in the deposition of many fewer pollen grains and a substantially decreased number of seeds produced. (After M. J. Groom. 1998. *Am Nat* 151: 487–496.)

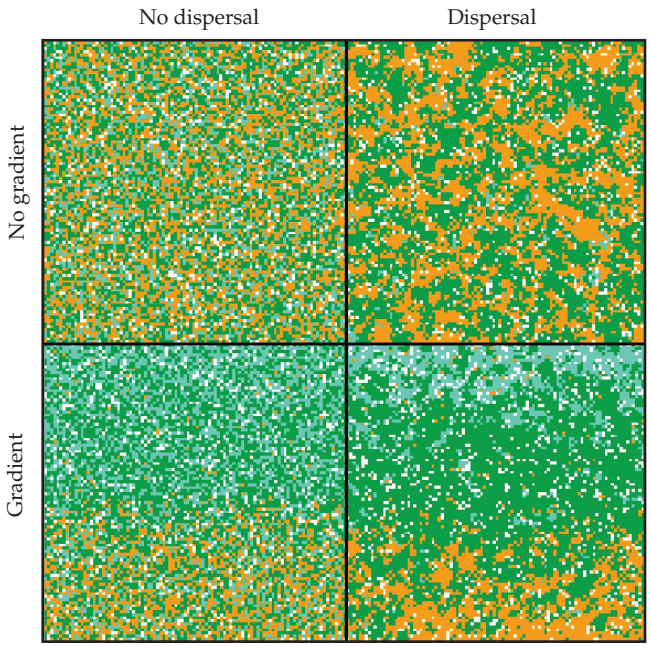



**Figure 15.7** A simulation model illustrating the joint effects of competition, environmental tolerances, and seed dispersal distance on species' distribution patterns. The species are *Pinus ponderosa* (orange), *Abies concolor* (green), and *Pinus contorta* (blue). Each square represents a stand of trees, and the color indicates the dominant species. Each panel is a two-dimensional landscape. In the top two panels, there is no elevational gradient; in the bottom two panels, the gradient runs from hot and dry at the bottom to cold and wet at the top. In the two left-hand panels, there is no seed dispersal, and dominance is determined by competition and environmental tolerances. In the two right-hand panels, seeds disperse to adjacent squares. The results of the simulation after 1000 years are shown. When there is no elevational gradient and no dispersal (upper left-hand panel), the three species are found across the entire landscape. When there is no gradient but there is seed dispersal (upper right-hand panel), both *P. ponderosa* and *A. concolor* form local patches where each dominates, while *P. contorta* is almost absent from the landscape. When there is a gradient (bottom panels), differences in elevational distributions of the species are seen, while the combination of a gradient and seed dispersal produces a pattern that mimics the actual distributions of the species (lower right-hand panel). (After D. L. Urban. 2005. *Ecology* 86: 1996–2006.)

technique among North American academic ecologists is a reflection of the current consensus among those ecologists that landscapes are mostly continuous and are likely to be dominated by one or two very strong gradients. Applied ecologists in North America often use classification methods to categorize vegetation for practical reasons, however (e.g., because of the requirements of regulatory agencies).

Most classification systems and techniques were originally devised by European and Australian ecologists (Mueller-Dombois and Ellenberg 1974). The Zurich-Montpellier, or floristic-sociological, approach was developed by the Swiss ecologist Josias Braun-Blanquet (1932). While ecologists using this approach recognized some continuity, the approach emphasizes discontinuity. Northern Europeans, primarily in Sweden, took an even stronger position toward discontinuity in the early part of the twentieth century. More recently they have blended their methods with the floristic-sociological approach. Russian ecologists participated in many of these same debates. They were most strongly influenced by ecologists in northern Europe and had similar views on discontinuity.

Today, science is becoming increasingly global in scope and outlook, and it is becoming increasingly difficult to assign particular viewpoints to particular geographic areas. For example, in journals such as *Ecology*, *Journal of Ecology*, *Oikos*, *Journal of Vegetation Science*, and *Plant Ecology*, one can find articles representing all traditions being published by plant ecologists from all over the world.

## 15.3 Ecological Processes Occur across Landscapes

### Island biogeography theory

Our exploration of landscape processes starts with **archipelagos**, clusters of islands. Robert MacArthur and Edward O. Wilson's **theory of island biogeography** (1967) was developed to explain patterns of species presence and absence on islands (Table 15.1). That theory was important in the development of much of today's landscape ecology and was a very important influence on the development of quantitative models in all of ecology.

Courtesy of Edward O. Wilson

Edward O. Wilson

Although the initial description of the theory applied it to oceanic islands, many types of communities have island-like properties. An inclusive definition of an **island** is any area that is suitable for the survival and reproduction of a species and is surrounded by unsuitable habitat. A landscape made up of isolated patches of forest consists of a set of islands from the point of view of an herb that lives only in forest understories. Mountaintops act as islands for alpine species. Plants can be specialized for living on particular types of soil or terrain, such as serpentine barrens (see Box 4A)

**TABLE 15.1** Propositions of the equilibrium theory of island biogeography

| Ecological propositions |
|---|
| 1. The rate of immigration of species (that is, the arrival of new species) to an island decreases as the number of species that have arrived on an island increases; the rate reaches zero when all species that could colonize from an available pool of species have done so. |
| 2. The rate of immigration of species to an island decreases with increasing isolation from a pool of potential colonists. |
| 3. The rate of immigration of species increases with increasing island size. |
| 4. The rate of extinction of species established on an island (of a given size) increases with increasing numbers of species. |
| 5. The rate of extinction of species established on an island decreases with increasing size of an island. |
| 6. The number of species on an island will be determined by an equilibrium between rates of immigration and extinction. |
| 7. The rate of species turnover (that is, change in species composition) will be determined by an equilibrium between rates of immigration and extinction. |
| **Evolutionary propositions** |
| 8. In addition to immigration, the number of species on an island can be increased by speciation. |
| 9. Speciation will only be important to an equilibrium in species number in a "radiation zone" found at the outward limits of species capacity for natural dispersal. |
| 10. The distance to a "radiation zone" is taxon specific. |
| 11. An equilibrium in species number reached ecologically is a "quasi-equilibrium" that can be increased over evolutionary time. |

*Source*: D. Sax and S. D. Gaines. 2011. In *The Theory of Ecology*, S. M. Scheiner and M. R. Willig (Eds.). University of Chicago Press: Chicago. Based on R. H. MacArthur and E. O. Wilson. 1967. *The Theory of Island Biogeography*. Princeton University Press: Princeton.

or rock outcrops, which also function as islands at both local and larger spatial scales. Human-dominated landscapes often form inhospitable "oceans" around natural areas. The theory of island biogeography therefore has applications far wider than to oceanic islands.

The basic idea of the theory is quite simple: The number of species on an island is determined by the rate of immigration of new species to the island and the rate of local extinction of species on the island (Figure 15.8). The rate of arrival of new individuals from the mainland determines the immigration rate—the rate at which new species arrive. (This use of *immigration rate* differs from its use in other contexts, where it refers to the rate at which individuals arrive.) As the number of species on the island increases, the immigration rate declines, because an arriving individual is increasingly likely to be a member of a species that is already on the island. The relationship between species number and extinction rate is a bit more complicated. Assume that the island can hold a fixed total number of individuals. As more species take up residence on the island, the potential population size of each becomes smaller because resources (including space) are on average less available for each species. The chance that a species will go locally extinct—disappear from the island—increases as its population size becomes smaller. Thus, the rate of extinction increases as the number of species rises. If we make the further assumption (as MacArthur and Wilson did) that these processes are roughly constant over long periods, then the number of species on the island eventually reaches a steady state, given by the value of the horizontal axis where the immigration and extinction curves intersect. (We are ignoring the possibility of local evolution and speciation.)

A key insight of MacArthur and Wilson was that larger islands can hold more individuals and, consequently, extinction rates on larger islands are lower. Moreover, islands closer to the mainland have higher immigration rates because it is more likely that individuals will arrive there from the mainland (the source of immigrant species). Using their assumption that islands reach an equilibrium number of species, MacArthur and Wilson predicted that the greatest number of species would be found on large islands near mainlands and that the smallest number would be found on small islands far from mainlands (see Figure 15.8C,D). These predictions can be tested in various ways, including by constructing type IV species richness relationships (see Table 12.7).

Real archipelagos, of course, are more complex: among other things, immigrants can arrive not just from the mainland, but also from other islands. A more important limitation of island biogeography theory is that it treats all species as if they were interchangeable. Clearly, this is not the case. Even very similar species have distinctive ecologies. MacArthur and Wilson acknowledged this limitation, although many ecologists subsequently trying to apply their theory appear to have ignored this assumption. The equilibrium assumptions needed to predict species number are unlikely to hold, and empirical studies have provided little support for the theory's predictions. Despite these limitations, the theory is still useful. It provides a starting point for examining patterns and gives us a way of thinking about the movement of individuals and species in a landscape.

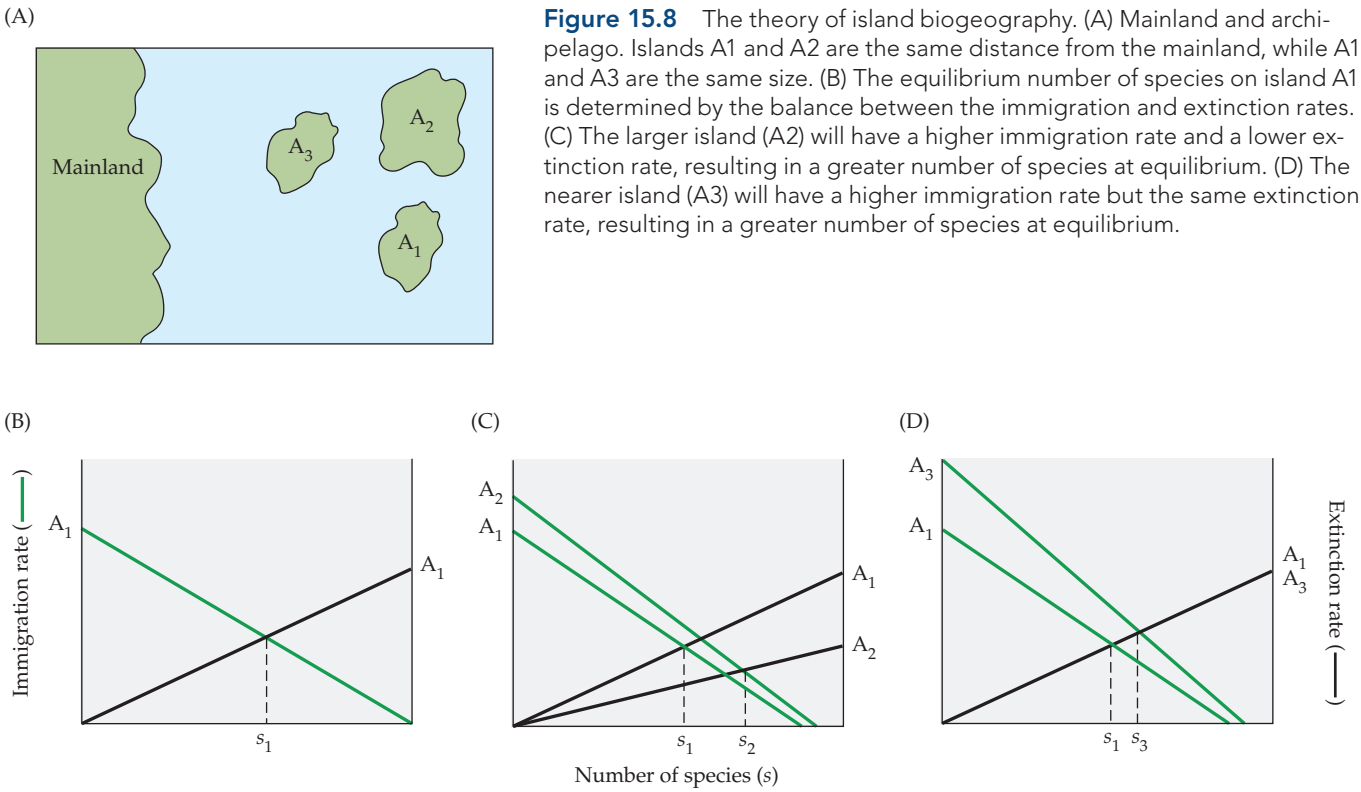



**Figure 15.8** The theory of island biogeography. (A) Mainland and archipelago. Islands A1 and A2 are the same distance from the mainland, while A1 and A3 are the same size. (B) The equilibrium number of species on island A1 is determined by the balance between the immigration and extinction rates. (C) The larger island (A2) will have a higher immigration rate and a lower extinction rate, resulting in a greater number of species at equilibrium. (D) The nearer island (A3) will have a higher immigration rate but the same extinction rate, resulting in a greater number of species at equilibrium.

Robert J. Whittaker and colleagues (2008) made an important addition to the theory that moved it away from its equilibrium assumptions by adding consideration of the geologic history of oceanic islands, the process by which the islands are formed and gradually disappear. Such islands are built by volcanic activity and then gradually erode away. Additional understanding of patterns of species richness on islands comes from accounting for both the age of the island and the potential for new species to arise on the island by speciation (see Table 15.1). Evidence from an analysis of 465 data sets of species richness in island archipelagos, including 108 data sets for plants, was amassed by Kostas Triantis and colleagues (2012). For each archipelago, they examined the shape of the relationship between the number of species and the area of the islands. Those relationships were consistent with that predicted by the MacArthur and Wilson theory of island biogeography. The large number of archipelagos included in this study makes it a powerful test of that theory and shows that despite its equilibrium assumption, it suffices as an approximation.

## Ecologists have debated whether there is a set of rules that determines how communities are put together

What determines the number and identity of the species in a community or landscape? This is a fundamental question in ecology and also is important for our ability to preserve and restore threatened and degraded habitats. Many ecologists have thought about this problem, and there are many approaches and ideas for how to make progress in understanding how ecological systems are put together at the community and landscape scale.

What are referred to as "community assembly rules" are not a single, coherent theory so much as a loose connection of models and theories built to try to explain the identities and abundances of the species in a community (Weiher and Keddy 2001). The roots of some of these models and theories include island biogeography theory, which focuses on species identities, and theories that relate species' traits to environmental conditions. The goal of these models and theories is to do more than simply describe patterns, but to ascribe those patterns to a process or mechanism. The first step has usually been to correlate a pattern of variation among communities with some property hypothesized to be due to differences in the mechanism or process. For example, say there is a set of islands with different sets of plant species. We might hypothesize that competition among those species limits the number on each island (see Chapter 10 and Chapter 14). One way to test this hypothesis is to compare the actual pattern of species occurrences with what might happen if species were simply randomly distributed.

This first step has been fraught with controversy. In the 1970s and 1980s there was a vigorous debate about how to recognize patterns. The problem is what the

pattern should be compared to. For those plants on the islands, how should that random distribution be defined? Say that there are ten islands, each contaning between 4 and 8 species with a mean number of 6.7 species, and a total of 25 species across all of the islands with each one, on average, being found on four islands. When we randomly draw species from the total pool for each theoretical island, do we pick only as many as were found on that island, or do we chose the average number across all islands? Do we limit the number of times that we can pick a species to the number of times that it is found on the islands, or do we use the average number of occurrences? And these choices are only based on occurrences. What if we also want to account for differences in abundances? As you might imagine, different people made different decisions and came to different conclusions. Because this lack of agreement seemed intractable, ecologists used another approach and focused on developing theories for the processes and mechanisms, as discussed in the rest of this section. In Chapter 12 we presented many different types of metrics and approaches for describing communities, even though ecologists still do not fully understand the processes and mechanisms responsible for the development of many of the subsequent patterns.

## Metapopulation theory

The theory of **metapopulations**, originally developed by Richard Levins (1969) and subsequently expanded upon by Ilkka Hanski (1982, 1999), was another important stimulus for landscape ecology. In some respects, metapopulation theory is similar to the theory of island biogeography, but applied to populations rather than sets of species. Begin by thinking of an archipelago, but one far from a continent. The population size of a species on any given island is a result of local population dynamics plus immigration minus emigration (this is the same as the first equation in Chapter 8). A population on a given island may go extinct, but new immigrants from other islands may repopulate that island. The equilibrium between immigration and extinction rates determines the number of surviving populations in the archipelago and their average size, and thus the total size of the metapopulation. Like the theory of island biogeography, metapopulation theory applies, of course, to situations other than oceanic islands. Levins's ideas were undoubtedly influenced by those of the population geneticist Sewall Wright, who envisaged a very similar process as part of his shifting balance theory of evolution (Wright 1931, 1968). We discuss the details of metapopulation models in Box 15B.

Classical metapopulation theory predicts the fraction of patches that are occupied at any time. One important conclusion that comes from extending population dynamics to a landscape is that not all local populations need be self-sustaining. A **sink population** may be continually on the road to extinction, but it may be maintained by constant immigration from **source populations**. And a population may be a source for some periods of time and a sink for others. The idea of source and sink populations is important because it emphasizes that local populations are typically not at equilibrium, even though the larger system (the metapopulation) may be. Thus, testing theories like this requires careful attention to matters of scale. During the 1960s, equilibrium was often the dominant assumption among ecologists. Since the 1980s, the pendulum has swung in the opposite direction, and nonequilibrium is generally assumed, at least at a local level (see Chapter 8). A metapopulation perspective moves the equilibrium assumption up one level to the entire landscape. While individual populations are not at equilibrium, the entire metapopulation is often assumed to be at equilibrium, though this assumption has not been tested.

Metapopulation models have been extended in various ways. Lenore Fahrig (1998) developed the "patchy population model." This model is an **individual-based model**—it consists of simulations in which the properties of individual organisms are explicitly modeled. By comparison, other metapopulation models are **mean field models**—all individuals in the population are assumed to have the same properties, so only mean population parameters are simulated. In Fahrig's model, individuals give birth, move, and die based on individually defined probabilities. Both conventional metapopulation models and individual-based models can be spatially structured, as is the case in Fahrig's model. This model allows population dynamics to occur within each patch, with immigration and emigration rates determined by the sum of individual actions, rather than being predetermined. Resit Akçakaya and colleagues (2004) and others have extended metapopulation models to include a more demographic focus, sometimes combining demographic work with species distribution models (known as "coupled niche demographic models"), where the spatial distribution of a group of populations depends on the factors determining the species distribution. Most of this work focuses on important conservation problems, but more of these studies have been done so far for animal than for plant populations.

Courtesy of Lenore Fahrig

Lenore Fahrig

How well does metapopulation theory work at predicting landscape patterns for plant communities? The

## BOX 15B Metapopulation Models

Metapopulation models have the general form

$$\frac{dp}{dt} = \text{immigration rate} - \text{extinction rate}$$

where $p$ is the percentage of patches occupied by a given species, $0 \le p \le 1$, and $dp/dt$ indicates the rate of change in occupation over time. Typically, we are interested in solving this equation for $p$, which predicts the frequency of occupied patches in a landscape. Obviously, a species' biology sets its immigration and extinction rates.

It is important to realize that in these models, immigration and extinction rates may not be independent of each other, and either or both may depend on the frequency of occupied patches, $p$. It is easy to see how the immigration rate might depend on $p$: the number of seeds arriving in a patch could depend on the number of occupied patches in the landscape. The extinction rate might also depend on $p$ for the same reason: as the number of seeds being dispersed increases (with increasing $p$), there could be a reduced chance of local extinction. This phenomenon is called the **rescue effect** (Brown and Kodric-Brown 1977) because new immigrants "rescue" the local population from extinction.

Ecologists have studied several forms of this model. Under the earliest version of the model (Levins 1969), the immigration rate depends on the fraction of occupied patches, but the extinction rate does not:

$$\frac{dp}{dt} = ip(1-p) - ep$$

That is, the immigration rate is a function of the probability of immigration, $i$, times the fraction of available sources of immigrants, $p$, times the fraction of available patches, $1 - p$. The extinction rate is a function of the probability of extinction, $e$, times the fraction of patches with populations that can go extinct, $p$. If these rates hold constant for a long time, then ultimately the fraction of occupied patches is

$$1 - e/i$$

Incorporating the rescue effect into the model makes the extinction rate depend on $(1 - p)$ as well (Hanski 1982):

$$\frac{dp}{dt} = ip(1-p) - ep(1-p)$$

Under this model, if $i > e$, ultimately all patches will be occupied, and all patches will be empty if $e > i$.

Still another alternative is that immigration could be independent of the proportion of occupied patches. Nicholas Gotelli (1991) refers to occupancy-independent immigration as the "**propagule rain**" because immigration provides a steady "rain" of new individuals to each population. Propagule rain models have as their immigration term $i(1 - p)$. All four possible combinations of occupancy-dependent and occupancy-independent immigration and extinction rates are conceivable and represent the extremes of a more general metapopulation model.

results are mixed. The key issue is whether plant populations have migration and extinction rates high enough for metapopulation processes to be a significant factor in determining species distributions and population sizes. The various versions of metapopulation models make explicit predictions about two aspects of species distributions (Table 15.2). First, if local extinction rates are dependent on the frequency of occupied patches, then those species occurring in a large number of patches will also have large populations when present in a patch—a positive correlation between occupancy and abundance. If local extinction rates are independent of the frequency of occupied patches, the correlation will be zero. Evidence from many studies shows that in most plant species, occupancy and abundance are positively correlated (see Chapter 14) (Brown 1984; Rey Benayas et al. 1999). Thus, at least some metapopulation models are consistent with available data.

The second prediction of metapopulation theory is not so well borne out in plant studies. If both immigration and extinction rates depend on the frequency of occupied patches, then patch occupancy by a species should have a bimodal distribution (Figure 15.9A). In other words, species should tend to be either very common or very rare, although they may switch over time. Hanski (1982) coined the terms *core species* for those species that are common in all communities in a landscape and *satellite species* for the rare ones. This version of the theory is often called the core-satellite hypothesis. Some ecologists occasionally confuse the core-satellite hypothesis with the idea of source and sink populations; they are quite distinct from each other in that the former idea concerns characteristics of entire communities. The bimodal patch occupancy distribution of the core-satellite hypothesis becomes unimodal if either the immigration or extinction rate is occupancy-independent. In that case, the mode is located somewhere in the middle of the distribution (Figure 15.9B)—that is, typical species will be somewhat common, with the exact values depending on immigration and extinction rates.

One series of studies of tallgrass prairie communities in Kansas by Nicholas Gotelli and Daniel Simberloff (1987) and Scott Collins and Susan Glenn (1990, 1991) found evidence of bimodal patch occupancy

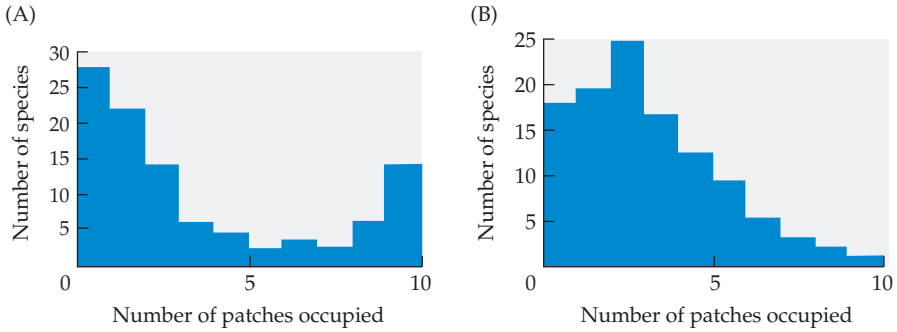



**Figure 15.9** Two possible patch occupancy distributions. (A) Both immigration and extinction rates are dependent on patch occupancy rates. (B) Only one of the rates, or neither, is dependent on patch occupancy rates. (After S. L. Collins and S. M. Glenn. 1991. *Ecology* 72: 654–664. A based on I. Hanski. 1982. *Oikos* 38: 210–221, B based on R. A. Levins. 1969. *Bull Entomol Soc Am* 15: 237–240.)

distributions for surveys done at a scale of 1 km$^2$ or less (Figure 15.10). However, Samuel Scheiner and José Rey Benayas (1994) surveyed 74 landscapes ranging in scale from 0.5 to 740,000 km$^2$ (mean 755 km$^2$) and found no evidence of bimodal distributions. Thus, at medium to large scales, the data for terrestrial plants are not consistent with the hypothesis that metapopulation processes are important, but they may be important at smaller scales of approximately 1 km$^2$ or less.

Why does metapopulation theory do such a poor job of predicting plant species distributions at large spatial scales? The models work only if immigration and extinction rates are high enough to influence local population dynamics. Immigration rates must be large to generate a positive correlation between occupancy and abundance. For most plants, immigration and extinction rates are probably too low for metapopulation models to hold. Long-lived perennials dominate most plant communities. For many of these species, the establishment of new individuals is episodic, perhaps occurring just once a decade or less when the right set of conditions comes along. Metapopulation theory has been successful in predicting abundance patterns for insects and small mammals that have generation times of a year or less and can move substantial distances. For plants, the theory is probably most relevant to populations of early successional plants, weedy plants, and annuals, as was found by Kelsey Martinez and colleagues (2015).

Unfortunately, the term *metapopulation* is often used to refer to any collection of populations, regardless of whether local extinction and recolonization govern the dynamics. As in many situations, if an idea is so broad that it includes everything, then it loses its utility. Metapopulation ideas have contributed and continue to contribute to ecology and its application, so long as the ideas are not diluted to the point of vagueness.

## Demographic processes occur across landscapes

Incorporating consideration of spatial scale and heterogeneity into our thinking has influenced many different areas of contemporary ecology. The demography of plant populations—the study of the components of the growth of populations (see Chapter 8)—has almost always been considered at the level of one individual population or a few populations. Roberto Salguero-Gómez and colleagues (2015) surveyed over 5600 demographic studies

**TABLE 15.2** Predictions about the correlation between the frequency of occupied patches and species abundance and the shape of the frequency distribution from four different versions of metapopulation models

| Model | | Predictions | | |
|---|---|---|---|---|
| Immigration | Extinction | Correlation | Distribution | Source |
| Dependent | Dependent | Positive | Bimodal | Hanski 1982 |
| Dependent | Independent | Zero | Single interior mode | Levins 1969 |
| Independent | Dependent | Positive | Single interior mode | Gotelli 1991 |
| Independent | Independent | Zero | Single interior mode | Gotelli 1991 |

*Note*: The models are classified by the dependence of immigration and extinction rates on the fraction of sites that are occupied (see text for details).

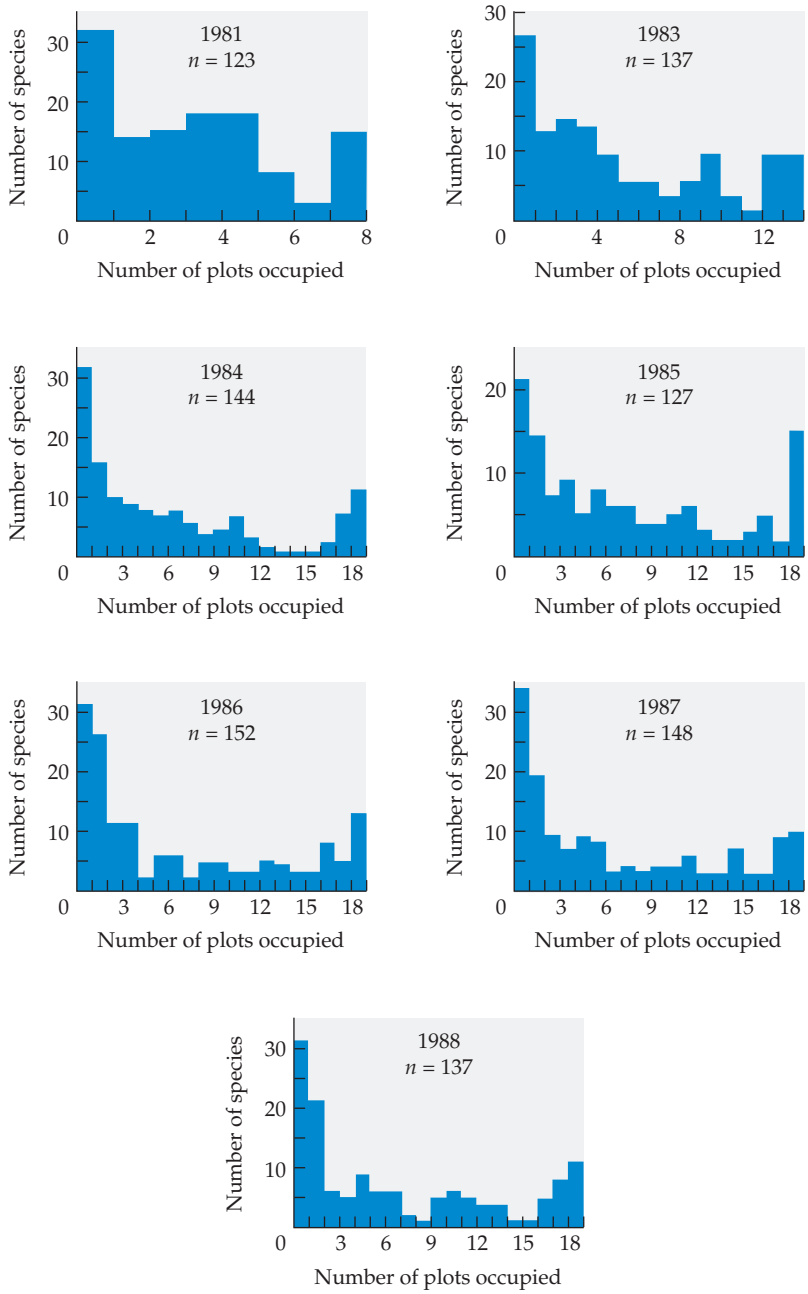



**Figure 15.10** Patch occupancy distributions over 8 years in the Konza Prairie Research Natural Area, Kansas. Eight plots were sampled in 1981, 14 in 1983, and 19 from 1984 to 1988. The *n* for each year is the total number of species among all plots. Each plot consisted of 20 quadrats of 10 $m^2$ a grain of 200 $m^2$, spread across an extent of 3487 ha. (After S. L. Collins and S. M. Glenn. 1991. *Ecology* 72: 654–664. Data for 1986 modified from S. L. Collins and S. M. Glenn. 1990. *Am Nat* 135: 633–648.)

from about 600 plant species and found that three populations per study was the most common number. The study of a single population might be representative of other populations in other sites, but it might not. Since studies of multiple populations typically find considerable demographic variation among them, it is not safe to assume that one or a few populations are representative. In addition, ecologists often want to know how demographic processes vary with spatial scale. For example, if we want to understand how a newly invasive species is spreading across a region, we may need to know how each population is behaving. Are they all growing at the same rate? Are some growing rapidly, while others remain stable? Or, a species might be shifting its range in response to climate change. Some populations in this case might be declining, while others are increasing in a different part of the range.

To emphasize and develop ideas about demographic processes across populations and spatial scales, Jessica Gurevitch and colleagues (2016) introduced the concept of **landscape demography**, the study of the demographic properties of multiple populations and their drivers across spatial scales. Landscape demography brings together the areas of landscape ecology, demography (see Chapter 8), and metapopulation ecology. Such studies may examine how the relationships among populations and their drivers at any one scale influence demographic outcomes at other scales. While there have been only a few studies that truly embrace the concept of landscape demography up to now (Anderson et al. 2009; Fordham et al. 2012; Conlisk et al. 2013), more work in this area may prove useful for both conservation and fundamental science.

## Metacommunity theory

The theory of **metacommunities**, especially as developed by Matthew Leibold and colleagues (Leibold et al. 2004; Leibold 2011), takes the ideas developed for metapopulations and applies them to whole communities by incorporating the types of species interactions that we discussed in Chapter 10 and Chapter 11 into a landscape context. In a similar fashion, it looks at how dispersal interacts with other ecological processes to determine community composition. Metacommunity theory (Table 15.3) grew out of metapopulation theory by relaxing the assumptions of species and patch equality.

One puzzling aspect of many plant communities is that they seem to contain more species than can be

**TABLE 15.3** The propositions of metacommunity theory

1. Dispersal contributes to local community assembly by serving as a source of colonists for species that were previously absent.
2. Dispersal can generate population mass effects between communities that differ in local fitness.
3. Heterogeneity in dispersal among species can allow for novel mechanisms of coexistence in a metacommunity.
4. Metacommunities can have spatial feedbacks due to a complex suite of additional effects beyond those of recurrent dispersal between patches.
5. Within local communities organisms belonging to different species interact via a set of interactions involving both direct and indirect effects.
6. Interactions in local communities are context-dependent in that they are affected by the local environmental abiotic conditions and by the existence of other species in the community.
7. These interactions affect which species can have self-sustaining populations and regulate the abundances of such populations in the absence of dispersal.
8. In the extreme when stabilizing forces are weak and fitness equalizing forces are strong, stochastic demographic dynamics may affect patterns of distribution and abundance.
9. Local extinctions are important.

*Source*: M. A. Leibold. 2011. The metacommunity concept and its theoretical underpinnings. In *The Theory of Ecology*, S. M. Scheiner and M. R. Willig (Eds.) pp. 163–184. University of Chicago Press: Chicago.

explained by processes that are occurring just within the community itself. One explanation for this greater diversity is that although diversity cannot be maintained within a single community and species are continually going locally extinct, those species continue to recolonize the community from elsewhere in the landscape. Previously, we described this source-sink dynamic for populations of a single species; that same dynamic can be part of the processes for the whole community. And just as a population within a community may not be at equilibrium, the same holds for the entire community.

Metacommunity processes can be more complex than those of any one metapopulation because the species can differ in their dispersal capabilities and responses to variation in the environment. For example, a poor competitor might have a high dispersal rate but be maintained in the landscape by a fugitive strategy; this could occur with species that are specialized for colonizing disturbed habitats (see Chapter 13). Even without disturbances, changes in the mix of species in a community can alter the outcome of competition. In the hand game "rock paper scissors," each player chooses one option; scissors cuts paper, paper covers rock, and rock breaks scissors. No single strategy always wins because the winner is dependent on the choice of the other person. Competitive outcomes are context-dependent. So even if one species is being outcompeted in one community, it may thrive in another community. (Of course, this will not hold in the case of a strict competitive hierarchy; see Chapter 10.) Variation among communities in other types of species interactions—the presence of different pollinators, herbivores, or pathogens—can have similar effects. The result is a highly dynamic landscape with diversity maintained at the level of the landscape as a whole, rather than within any single community.

The processes that determine plant community composition can vary depending on the spatial scale being looked at, the ecology of the species, and the environment. Various models have been developed to describe metacommunity dynamics that depend on the balance between the amount of connectivity among communities, based on rates of dispersal, and the extent to which environmental variation among those communities favors different species (Figure 15.11). If the environment is relatively similar among communities, and if the species are similar in their effects on each other, then community patterns will match the predictions of neutral models in which species' identities and traits do not affect community membership (Hubbell 2001; see Chapter 10 and Chapter 19). In a study of coastal vegetation in northeast Brazil, over 95% of the variation in species composition among closely spaced plots could not be attributed to species' identities (Silva et al. 2015). But this pattern can be modified: patch dynamic models based on low dispersal rates and low environmental heterogeneity are similar to neutral models, but they do not assume that the species are similar to each other. Instead, they are based on trade-offs among the species in colonizing and competitive abilities. An example of a patch dynamic pattern was found in scrub and grassland vegetation in Israel (May et al. 2013). In yet another approach, models of species sorting (see Chapter 19) assume that the local environment is what primarily determines community composition; this pattern has been found for coastal dune vegetation in Denmark (Brunbjerg et al. 2012) and trees in southern India (Das et al. 2017). Finally, if dispersal rates are very high, then the communities are dominated by what are termed mass effects, which have been found for forest communities across Wisconsin (Waller et al. 2018).

Most studies of metacommunity patterns focus on numbers of species, but metacommunity theory can also be used with other aspects of biodiversity, such as phylogenetic and functional diversity (see Chapter 12). Marko Spasojevic and colleagues (2014) looked at vegetation in

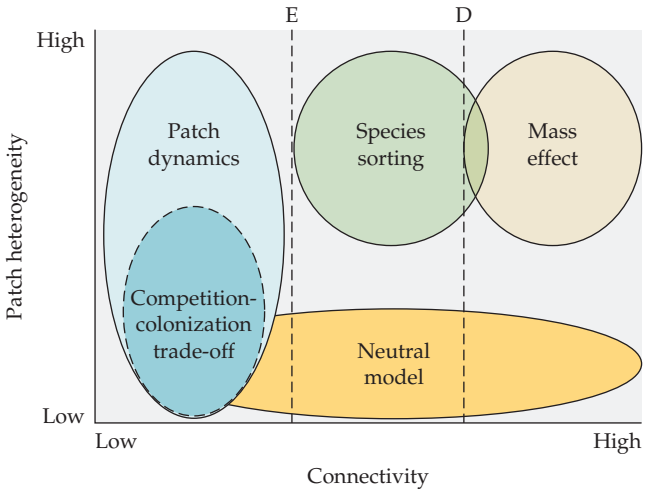



**Figure 15.11** Four metacommunity models in relation to the amount of heterogeneity among patches and connectivity. Patch heterogeneity is the degree to which different species are favored in different patches. Connectivity is determined by the rate of immigration (events/time), which can be compared with two other rates, the local rate at which species go extinct (E) and the rate at which individuals die (D). A separate oval has been drawn for patch dynamic models that focus on competition-colonization models in the absence of patch heterogeneity because most patch dynamic models are of this type. (After M. A. Leibold. 2011. In *The Theory of Ecology*, S. M. Scheiner and M. R. Willig [Eds.], pp. 163–184. University of Chicago Press: Chicago.)

17 alpine tundra patches in the Colorado Rocky Mountains. They considered two key environmental factors—temperature and elevation—and two factors that affect immigration rates—distance from other mountaintops and patch area. Functional diversity was characterized for each species by measures of plant height and four leaf traits: specific leaf area, leaf area, stomatal conductance, and chlorophyll content. As in the studies in Denmark and India mentioned above, they found that the species sorting model best fit the pattern of variation within and among patches, indicating that environmental factors were most important for determining community composition.

## 15.4 Ecological Processes at the Level of Landscapes Are Important for Plant Conservation

Besides being entirely destroyed, natural habitats can be degraded by many factors, including elimination of critical species; damage to the soil and changes in ecosystem processes (such as by erosion or nitrogen deposition); overgrazing and trampling by cattle, sheep, and goats; removal of water or other resources for human consumption; and changes in disturbance regimes. We consider some of these factors at more local scales in this chapter, and we return to larger-scale biodiversity loss in Chapter 19.

### Fragmentation of landscapes is a major threat to biodiversity

One of the greatest threats to many natural habitats, outside of sheer destruction, is the fragmentation of entire systems into smaller parts (Figure 15.12). Habitat fragmentation is a landscape-scale process in which contiguous natural habitat is segmented into pieces, typically by human activities. Fragmented habitats can be affected in many ways, depending on the size of the intact fragments, the kind of ecological community, and especially what is present between the fragments. Shrubby successional vegetation between forest reserves will have a smaller effect on the forest fragments than a large highway or an extensive urban area. Some of the effects of fragmentation are due to edge effects and reduced connectivity between fragments (see below). The effects of fragmentation per se are intertwined with how that fragmentation typically reduces the total remaining area of natural habitat (Curtis 1956). Lenore Fahrig (2003, 2019) showed that many of the negative effects of fragmentation may actually be due to this reduction; a particularly important effect is decreases in biodiversity. Small pieces

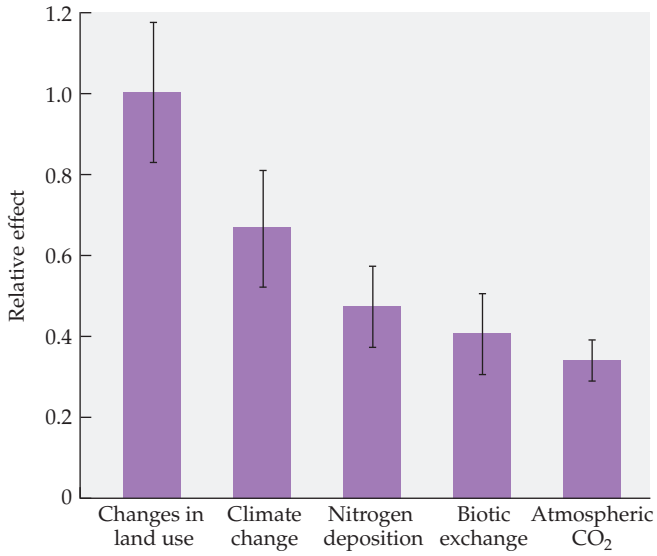



**Figure 15.12** Average estimated relative effects of the most important current threats to global biodiversity, projected for the next 100 years. Land use changes (including habitat destruction and deforestation) are predicted to have the greatest harmful effects on biodiversity, followed by climate change due largely to indirect effects of rising atmospheric $CO_2$ levels, nitrogen deposition, the incursion of invasive exotic species and other biotic exchanges, and direct effects of rising atmospheric $CO_2$ levels. (After O. E. Sala et al. 2000. *Science* 287: 1770–1774.)

of remaining habitat can provide a refuge for many species, and preserving small, fragmented habitat areas may be critical for conservation of species and communities.

Different kinds of human encroachment on natural landscapes can cause different effects. In China, large new cities have been built in areas that were previously rural, and some existing small towns and cities are rapidly expanding into rural and natural areas, with clusters of large apartment buildings and millions of people. In the United States, expansion of suburban and exurban fringes outside of cities and the building of homes in rural areas at the human-wildland interface are the most rapidly growing changes in land use (Figure 15.13; Hansen et al. 2005). As the densities of people increase in these areas, native species diversity often declines, and the prevalence of non-native and invasive species increases greatly.

Fragmentation can sometimes affect communities in unexpected ways. Mary Cadenasso and Steward Pickett (2001) quantified the movement of seeds from outside of several forests into the forests' interiors. Wind-borne seeds, including those of invasive non-native species, entered the forests in substantial numbers, an important but not unexpected finding. But they also found that long-undisturbed edges with denser, "branchier" vegetation formed far better barriers to seeds than did recently cut edges. Old edges not only blocked far more seeds from entering the forests, but also kept those seeds that made it into the forests much closer to the edge, better protecting the interior. These results imply that recent increases in forest fragmentation, which are creating many new edges, will exacerbate the spread of invasive species into the interior of these habitats.

Roads are one of the most influential and widespread factors that diminish and fragment natural habitats, whether they are protected areas or not. While building and improving roads offers many economic

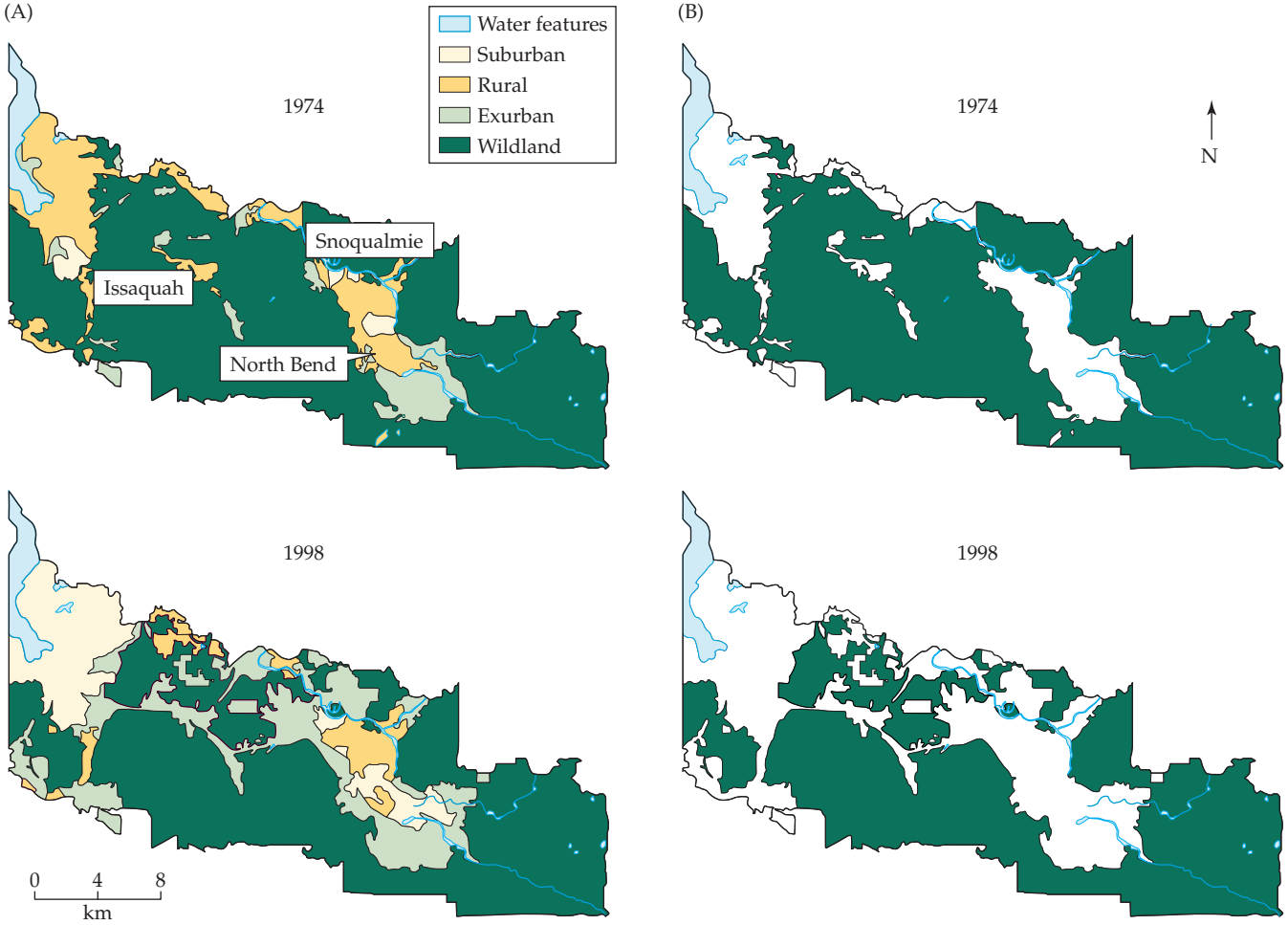



**Figure 15.13** (A) Changes in land use in the east urban fringe of Seattle, Washington, from 1974 to 1988. (B) The decline in forest area as a result of those changes. (From L. Robinson et al. 2005. *Landsc Urban Plann* 71: 51–72.)

benefits and access to health and education resources, they also have many negative effects. The effects of roads include deforestation, changes in the movement and killing of animals (including pollinators and seed dispersers), the spread of pathogens, changes in fire regimes, increases in wind speed, emission of pollution from cars, and the movement of invasive species into natural areas. Roads encourage human settlement and the extraction of resources (mining, oil extraction, hunting, lumbering and collecting wood). Roadless wilderness areas are rapidly disappearing globally, and efforts to conserve those remaining are far from sufficient for protecting biodiversity (Ibisch et al. 2016). Countries in Africa, the Middle East, Southeast Asia, and South Asia are all experiencing or planning great expansions in their road networks.

Fragmentation is occurring worldwide, from the tundra to the tropics (Figure 15.14). David Skole and Compton Tucker (1993) quantified changes in forest fragmentation in the Brazilian Amazon over a 10-year period (1978–1988) using satellite images. By the end of that period, almost three times more forested area consisted of forest fragments—patches of forest less than 100 km$^2$, surrounded by cleared land—than at the beginning, and the total area of forest was also greatly decreasing. The size and shape of cleared patches also changed during that period due to transformations in the economy and human demographics. Clearings changed from smaller and more irregular patches to larger and more continuous areas, with negative implications for the viability of the forest ecosystems (Peralta and Mather 2000).

Our understanding of the consequences of fragmentation does not simply depend on theory; both observational and experimental data exist to document its effects. One of the largest, oldest, and most important ecological experiments on forest fragmentation is the Biological Dynamics of Forest Fragments Project (BDFFP), an ambitious long-term experiment that was begun in the early 1980s and is still under way in Brazil. This collaborative study—involving the Smithsonian Institution of the U.S. and Brazil's Instituto Nacional de Pesquisas da Amazônia (National Institute for Research in the Amazon)—aims to examine the effects of forest fragmentation due to the widespread replacement of tropical forest by cropland and pastureland in Brazil (Bierregaard et al. 1992; Laurance et al. 2018). The study site consists of a series of forest fragments left standing amid initially cleared land (Figure 15.15). The project was designed with replicate plots; there are four plots of 1 ha, three plots of 10 ha, and two plots of 100 ha spread over a total area of about 1000 km$^2$. Plots in the nearby continuous forest serve as controls. The plots have been monitored for changes in the composition and abundance of plant and animal species repeatedly over the decades since the project began.

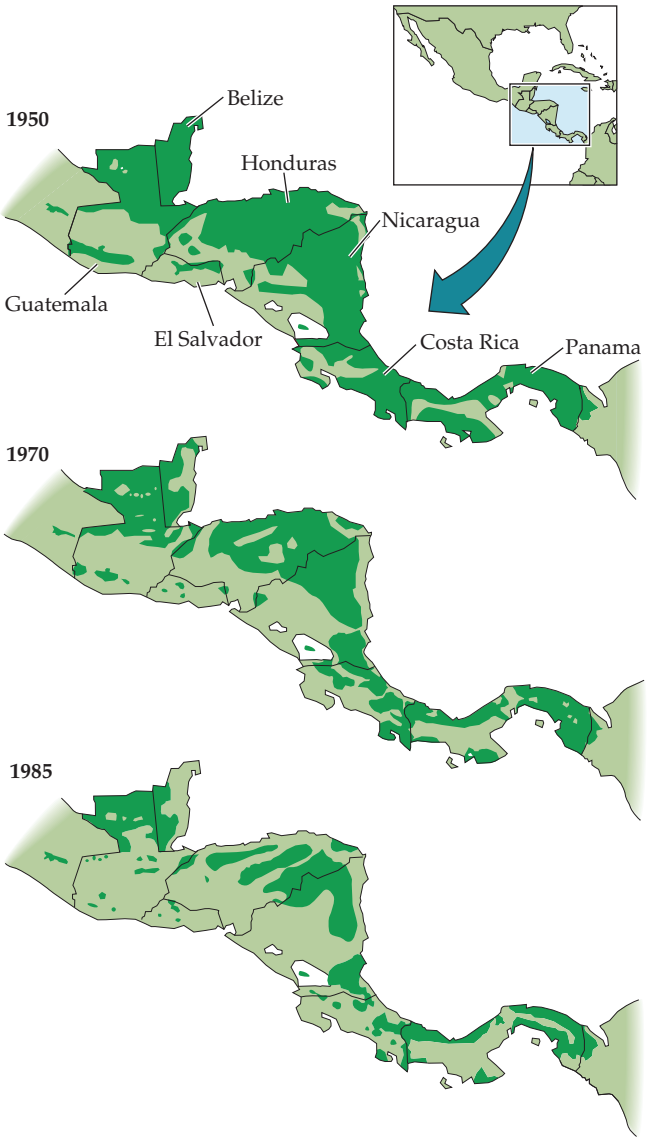



**Figure 15.14** Forest fragmentation in Central America between 1950 and 1985. Deforestation and fragmentation have increased in the years since 1985, although some forests are now protected in national parks, most notably in Costa Rica. Not only has the total area of intact forest declined dramatically, but the remaining forest fragments are small, often isolated from one another, and may experience large edge effects. Similar patterns are repeated for forests in many other parts of the world, and over a range of scales. (Maps after P. Utting. 1993. *Trees, People and Power: Social Dimensions of Deforestation and Forest Protection in Central America*. UNRISD and Earthscan. London.)

In the BDFFP, mortality rates of trees declined with increasing fragment size, but not in a linear fashion (Figure 15.16). Rare species are more likely to be missing

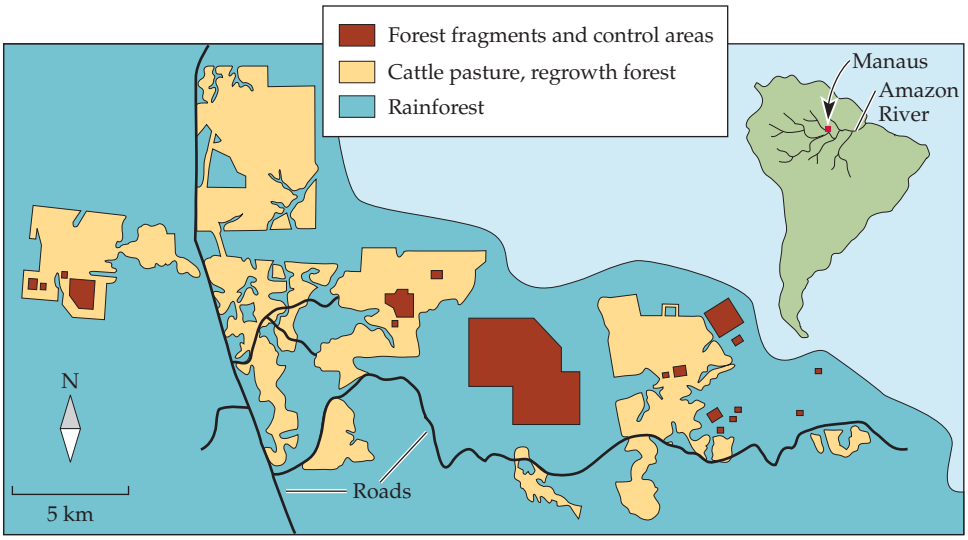



**Figure 15.15** The Biological Dynamics of Forest Fragments Project near Manaus, Brazil, in the Amazon basin. (After W. F. Laurance et al. 1998. *Ecology* 79: 2032–2040.)

from smaller forest fragments compared with 100 ha fragments. The difference is likely due to greater mortality rates in the smallest fragments (Figure 15.17A).

Working in the BDFFP, Emilio Bruna and Madan Oli (2005) found that the processes controlling population dynamics depended on fragment area. They collected 6 years of demographic data on the understory herb *Heliconia acuminata* (Heliconiaceae) and calculated its population growth rate ($\lambda$; see Chapter 8). The population growth rate for *H. acuminata* was 1.05 in the continuous forest, but lower (1.00) in both the 10 ha and 1 ha fragments, but for different reasons. In the 10 ha fragments, the lower growth rate was due to lower seed production. In the 1 ha fragments, the lower growth rate was caused by plants growing more slowly as well as producing fewer seeds.

Habitat fragmentation can change not only the sizes, but also the genetic diversity of local populations. Small populations—those in small habitat fragments isolated from other populations of their species—are likely to lose genetic diversity through the process of genetic drift (see Chapter 9). In turn, reduced genetic diversity (as well as small population size itself) may increase the chance of extinction because the population will be unable to evolve in response to changes in the environment. Because the process of fragmentation may itself cause environmental change, these effects become even more important. For instance, in small fragments of the BDFFP, individuals of the Brazilian canopy tree *Corythophora alta* (Lecythidaceae) are all very similar genetically, in contrast to a lower level of genetic relatedness in large fragments (Hamilton 1999). Because seed and pollen movement are limited, inbreeding is likely to increase, leading to further loss of genetic diversity.

Some of the long-term results of the BDFFP study reveal that the plant species most likely to suffer from fragmentation are those that are most affected by the

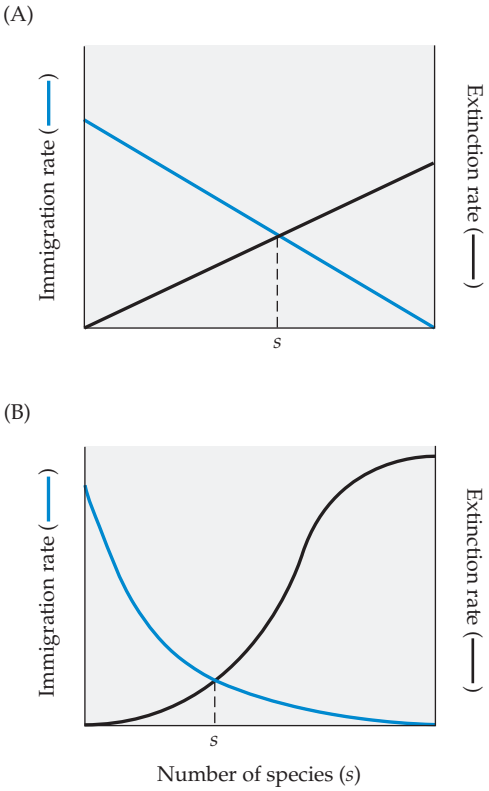



**Figure 15.16** How changing the dependency of immigration and extinction rates on species density affects the equilibrium number of species on an island or in a forest fragment. (A) Linear relationships, as predicted by the simplest version of island biogeography theory. (B) Nonlinear relationships. The shapes of the relationships depend on the biology of the species, such as how the strength of competition depends on species density. Depending on the exact shapes of the immigration and extinction curves, the equilibrium species number will go up or down.

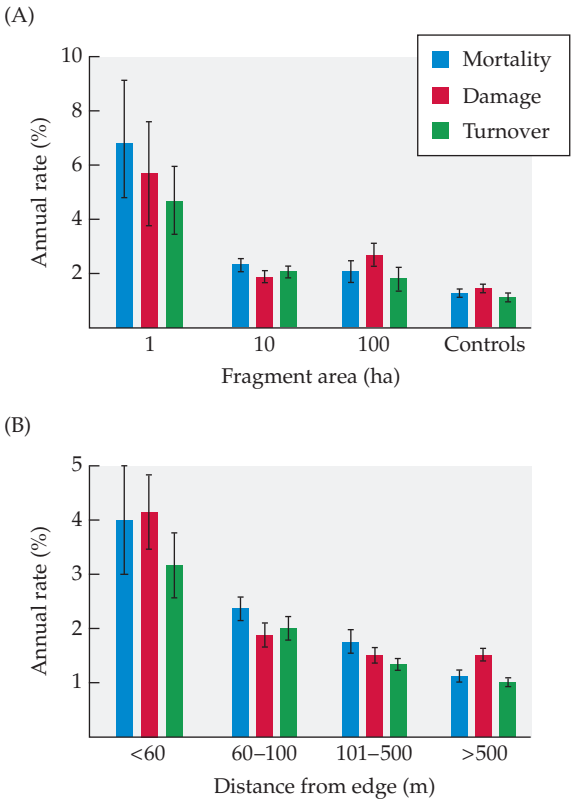



**Figure 15.17** Annualized rates (mean ± 1 standard error) of forest dynamic parameters for all trees in the fragments of the Biological Dynamics of Forest Fragments Project. Turnover is the rate at which trees are being replaced. (A) Effects of fragment area (size of fragment). (B) Effects of distance from the edge of the fragment. (After W. F. Laurance et al. 1998. *Ecology* 79: 2032–2040.)

conditions at forest edges, those that rely on animals for pollination or dispersal of seeds (when those animals themselves have become rare), and plants that are most vulnerable to disturbance. Forest fragments exhibited much more rapid change and less stability of population numbers and species composition, with strong and persistent effects of flooding, cold spells, wind, and drought. Marked changes overall include greater forest productivity (probably in response to higher atmospheric $CO_2$) and large increases in lianas (woody vines), particularly but not only at forest edges (Laurance et al. 2018).

## Key landscape characteristics are edges, connectivity, and nestedness

**Edge effects**—systematic differences between the perimeters of habitat patches and the interiors of those patches—can also be important to plants. Wind speed, for example, is greater at the edge of a forest than in the interior, and plants at the edge are exposed to more desiccation. If winds are strong enough, trees on the edges may have more limbs broken off or may have a greater chance of being knocked down. The effect of wind has been particularly important in the BDFFP plots. Near fragment edges, the rates of tree damage and mortality have been three times higher than inside intact forests (Figure 15.17B), leading to much higher rates of change in tree density (Figure 15.18A), species turnover (Figure 15.18B), and species loss (Figure 15.18C). Other edge effects include differences in available light (and therefore in the magnitude of diurnal temperature changes) and humidity between edge and interior and, in some cases, differences in $CO_2$ concentrations. Another effect of edges is forest drying. Moist air from transpiration above a forest is cooled by the evaporation of the water (see Chapter 3), and the air pressure above the forest decreases. The air over the nearby clearings is warmed by sunlight, and its pressure increases. This difference in air pressure can draw moist air out of the forest into the surrounding clearings and push dry air over the forest, where it can pull moisture out of the plants and soil. There may also be differences in the presence or activity of herbivores, pollinators, and seed dispersers between edge and interior. Julieta Benítez-Malvido and Aurora Lemus-Albor (2005) found that along forest edges in Chiapas, Mexico, the area of plant leaves damaged by pathogens (1.85%) was three times greater than in the forest interior (0.57%).

Other key aspects of landscape geometry include the distances between patches and the degree to which patches are connected to one another. **Connectivity** is the degree to which a landscape facilitates or impedes the movement of organisms (Tischendorf and Fahrig 2000). It is a function of both the physical arrangement of patches and the behaviors of organisms in response to that pattern. Different species may judge the connectivity of a landscape very differently. Much attention has been devoted to the idea of connecting habitat patches with "corridors," mainly in the context of animal conservation. Dispersal among habitat patches, even if it occurs only occasionally, can be an important means by which populations are preserved in the long run, as isolated populations are more likely to go extinct.

Much less attention has been given to connectivity in plant ecology, probably because most attention in conservation as a whole has been paid to animals, especially vertebrates. One exception is the work of Janez Pirnat (2000), who used GIS and vegetation sampling to study connectivity between forest fragments in Ljubljana, Slovenia. Here, forest patches existed as urban forests and parks, as suburban forests around the fringe of the metropolitan region, and as connecting forest corridors and fragments. Road building and development disrupted these connections, threatening the dispersal of forest plant species, particularly endangering the maintenance of diversity in the urban forests, which are population

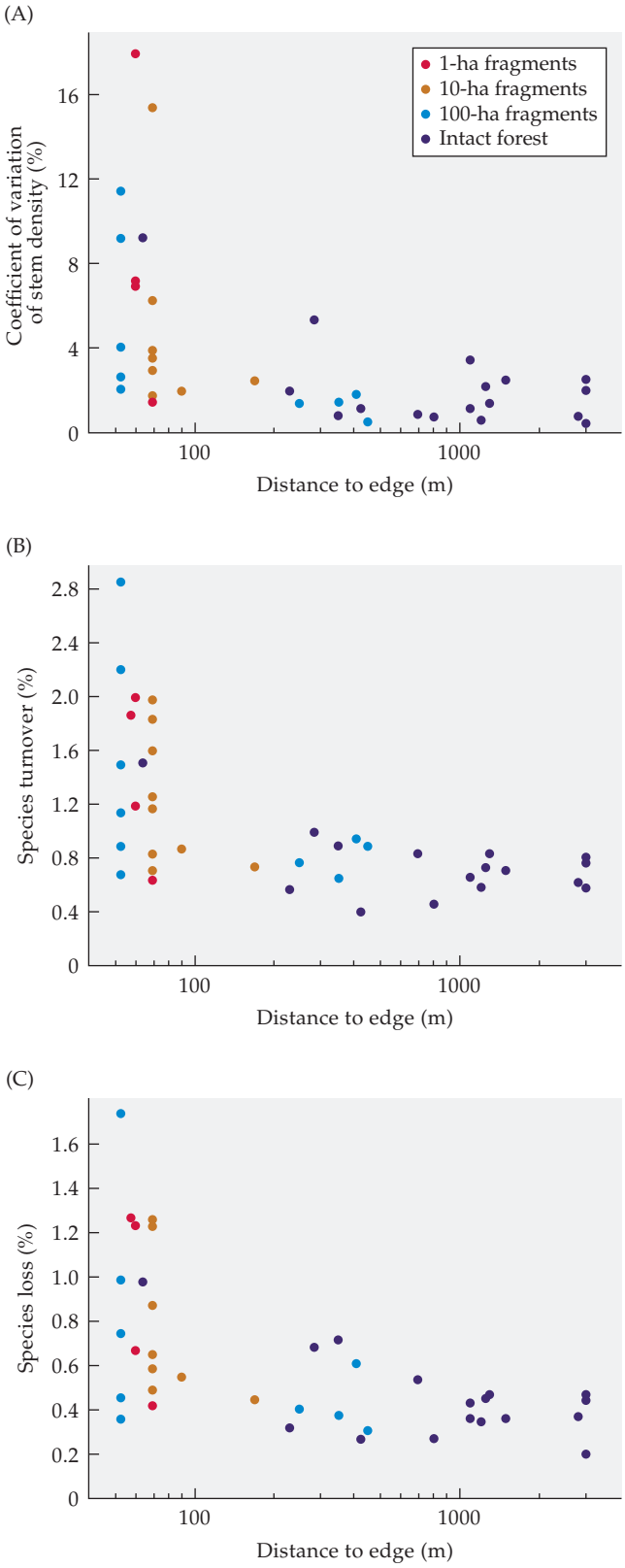



**Figure 15.18** Effects of distance from the edge of the fragment of the Biological Dynamics of Forest Fragments Project on (A) the temporal variation in tree density, (B) the annual percentage of species turnover, and (C) the annual percentage of species loss. (A after W. F. Laurance et al. 2018. *Biol Rev* 93: 223–247; B, C after W. F. Laurance et al. 2006. *Proc Natl Acad Sci USA* 103: 19010-19014. © 2006 National Academy of Sciences, USA.)

sinks for various tree species. Similarly, in many landscapes affected by natural or anthropogenic fragmentation, from large lava flows or fires to suburban development, the availability of foci of dispersal—patches of plants that can serve as refuges, reservoirs, and sources of seeds for dispersal into the sink regions—can be of critical importance in the maintenance and recovery of plant populations and thus play a valuable role in conservation.

The **nestedness** of a set of communities—the degree to which communities tend to be subsets of other communities (Table 15.4)—is an important landscape-level issue in conservation biology. In a perfectly non-nested set of communities, each community is composed of a unique set of species; for example, the sets of species growing in a forest and in a neighboring lake do not overlap. By contrast, in a perfectly nested set of communities, the smallest patches have only species contained in the larger patches. In a perfectly nested system, one could thus preserve all the species diversity by preserving just the largest patch, while in a perfectly non-nested system one would need to preserve every patch.

A simple application of island biogeography theory predicts that the small patches in a landscape will contain random samples of the entire set of species, with different species appearing in different patches. But fragmented landscapes have a tendency to become more nested over time: small patches often contain the same few species because those species are the hardiest or the best dispersers. Animal communities tend to be highly nested (Wright et al. 1998), although this is less true of plant communities. A major reason is that local environmental effects play a large role in determining plant community composition, and patches tend to differ in important environmental parameters. In the BDFFP, new immigrants into the fragments were mainly individuals of species from the disturbed areas, rather than the forest-interior species previously found there. Thus, there was a profound shift in the species composition of the fragments, which was not simply a random sampling of the previous species.

The BDFFP experiment examined the effects of fragmenting a landscape that contained a well-developed forest, the effects of community disassembly. Fragmentation can also affect the processes of community assembly and succession (see Chapter 13). Two long-term studies of the effect of patch size and connectivity on community assembly are the Kansas Fragmentation Experiment (Holt et al. 1995) and the Savanna River Site Corridor Experiment in South Carolina (Damschen

**TABLE 15.4** Two examples of nested landscapes

| (1) | Site | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Species | A | B | C | D | E | F | G | H | I |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| 3 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 |
| 4 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 |
| 5 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 7 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

| (2) | Site | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Species | A | B | C | D | E | F | G | H | I |
| 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 |
| 2 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 |
| 3 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 |
| 4 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 1 |
| 6 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 0 |
| 7 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 0 |
| 8 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

et al. 2006). The Kansas experiment was begun in 1984 in what had been a farm field. It consists of patches of varying spacing and sizes (up to 0.5 ha), with mowing between the patches to keep them distinct. The Savanna River experiment was begun in 2000 and consists of 1 ha patches of cleared areas within a pine tree plantation; these patches were either isolated or connected to others by smaller corriders of cleared area. The effects of fragmentation on plant community composition in these studies were compared with those of the BDFFP study and another in which mature eucalyptus forest patches were left intact and the surrounding area cleared—the Wog Wog Experiment in New South Wales, Australia, which was begun in 1987 and included patches up to 3 ha in size (Farmilo et al. 2013). The analysis by Cathy Collins and collaborators from all four studies (2017) found greater divergence in community composition between the most- and least-fragmented habitats in the two studies where communities were disassembling—BDFFP and Wog Wog—than in the two studies where communities were assembling—Kansas and Savanna River. β-diversity (see Chapter 12) within a fragmentation treatment did not change over time at any of the studies and was similar to β-diversity in nearby intact vegetation. In all four studies, compositional changes over time were due mostly to species turnover, although nestedness also played a role in the Kansas and Wog Wog studies.

## Ecological theory can help guide reserve design

Landscape ecology is very important to conservation efforts (Burgman et al. 2005). For example, the idea of using corridors to connect nature reserves comes directly from landscape ecology. While the goals of conservation efforts are quite varied (e.g., some efforts are aimed at protecting a single species, while others are aimed at preserving communities), landscape considerations are important to all of them. Many new nature reserves try to save the largest possible number of species (or communities) while juggling the realities of private land ownership, developmental pressures, and public use of lands. Older nature reserves, such as many national parks around the world, were developed mainly to protect scenery and often were areas that were not easy to farm or were relatively inaccessible for residential settlement. Often when a new reserve is proposed, there is not sufficient time and money to do a detailed analysis of the population ecology of many species, or even the most vulnerable species. Instead, reserve designers have looked to landscape ecology theory to provide simple guidelines that can be applied in the hope of saving the largest number of species.

This goal led (in the 1980s) to a dispute among ecologists called the SLOSS (single large or several small) debate. The debate centered on whether (given a fixed total area) it is better to create one very large reserve or many small reserves. You may ask, Why not always create very large reserves? On the one hand, larger reserves have many benefits, including reduced edge effects (discussed above) and larger population sizes. On the other hand, a single site may be vulnerable to random environmental (temporal) variation, or a major catastrophe, which could potentially eliminate a single population of a species existing in a single site. The simplest version of island biogeography theory (in which immigration and extinction rates are a linear function of area) predicts that the total number of species contained in the entire collection of fragments should not depend on the sizes of individual fragments as long as the total area is the same.

But this simple theory ignores some important biology. Immigration and extinction rates depend on the identities of the species involved and the nature of the habitat. Some species are better dispersers and colonizers than others. Some fragments might have richer soils than others. Extinction rates may be higher in more crowded communities because interactions among species intensify competition. These factors result in nonlinear immigration and extinction curves (see Figure 15.16). Island biogeography theory, which favors large reserves, does not consider environmental variation. Thus, there were reasonable arguments on both sides in the debate. Interestingly, the SLOSS debate was never resolved in an intellectual sense; rather, it was ended by the realization that in real life, the alternatives of single large versus several small reserves almost never present themselves—the point being argued was moot. The political and economic reality is that assembling a single large reserve can be expensive and very difficult, if not impossible. Buying up many small, disconnected parcels of land is sometimes much easier. The consensus reached was the one you might expect: preserve as much as you can. In any case, the SLOSS debate brought issues of landscape ecology to the fore in the planning of nature reserves.

## Summary

- A plant community exists as part of the set of communities that make up a landscape.
- A landscape need not be large relative to human perception. A landscape can be very small if the organisms are small.
- Processes that occur within communities can affect landscape patterns, and processes that occur across landscapes can affect community structure.
- The size of the communities and the distances among them—aspects of scale—are critical to understanding landscape patterns. Scale can be considered based on spatial distance or by how different the communities are in their environment.
- Scale has four aspects: sampling unit, grain, extent, and focus.
- Spatial patterns can be measured using a variety of on-the-ground and remote sensing methods.
- Three different sets of theories have been developed to explain landscape patterns: island biogeography theory, metapopulation theory, and metacommunity theory.
- Movement—whether of individuals or of materials such as nutrients—is a key ecological process in landscapes.
- An understanding of landscape dynamics and their consequences is important in conservation biology.
- Nature reserves can be designed to maximize the number of species preserved.
- Landscape fragmentation leads to decreases in connectivity among patches and increases in habitat edges.
- These changes in landscape structure can have large effects on community structure and composition, including the loss of species and decreases in the genetic diversity of populations.

Courtesy of S. Scheiner

16

# Climate, Plants, and Climate Change

If you were to travel any considerable distance, either across a continent or across the globe, you would experience different climates and see strikingly different plant communities (see Chapter 18). In this chapter, we consider the underlying climatic factors that control not just vegetation patterns but all life on Earth. We look at the physical factors determining seasons, intermediate-scale cycles of climate change lasting years to decades, and much longer-term cycles. We review the role of plants in shaping climates, and then discuss how humans are changing Earth's climate, as well as some of the impacts of these changes.

Above: A retreating glacier in Alaska, U.S.A.

## 16.1 There Are Important Differences between Climate and Weather

Any attempt to determine the causes of large-scale vegetation patterns must start with climate. **Climate** refers to the long-term averages and statistical distribution of the weather at any scale from local to global (e.g., London generally has cool temperatures and a lot of rain; the northern coniferous forests of Siberia have a short growing season and a long, cold winter), while **weather** refers to the immediate or short-term conditions at local scales (e.g., temperatures reached a record high on a certain date in a particular place; this week or this month it might be unusually warm, windy, and dry in New York City; Figure 16.1). Although weather has profound effects on the immediate function, growth, and survival of plants, it is climate that determines the general type of vegetation in an area (e.g., broad-leaved deciduous forest versus desert) and influences large-scale patterns of diversity. In this chapter, we describe the patterns of climatic variation around the world (with special attention to North America), explain the mechanisms responsible for those patterns, consider the role of plants as they respond to and influence climate, and look at how humans are influencing global climate.

The two primary components of a region's climate are temperature and moisture. First we consider long-term mean (average) temperature and precipitation and the effects of these average conditions on vegetation patterns. Variation around these means is also important, and we look at two aspects of this variation: predictable changes (e.g., it is usually warmer in the summer than in the winter in temperate and boreal environments) and departure from average conditions (e.g., in a particular year it may be unusually dry in the Amazon basin). The time scale of variation is also important, with temperature and precipitation varying on scales from a single day, a week, a season, a year, a decade, to cycles that take hundreds of thousands of years to complete. We investigate these patterns of variation, consider the underlying explanations for variation in precipitation and temperature patterns among different parts of the globe, and summarize how these climatic differences shape plant communities. We have already seen why the sky is blue and sunsets are often red (see Box 3B). We will explain in this chapter why tropical rainforests are warm and wet, why we have four seasons in Europe and North America with cold winters and warm summers, why the southern California beaches are renowned for their dry, sunny weather, why deserts exist where they are, and why polar regions are cold. You will learn why winds traveling toward the equator appear to veer to the west, and you will have practiced thinking in three dimensions and at global scales. We hope that you find these things exciting and satisfying to understand.

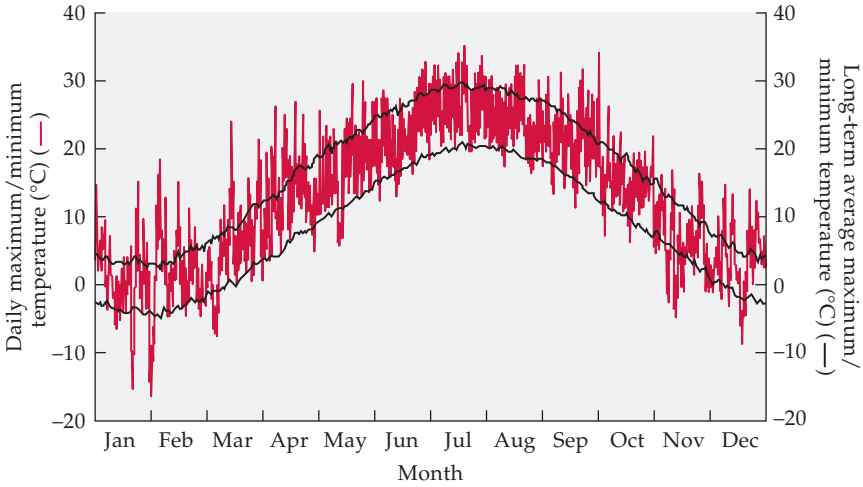



**Figure 16.1** Weather in New York City's Central Park in 2019 is shown by the red lines graphing daily maximum and minimum temperatures; climate is shown by the average daily maximum and minimum temperatures from 1869–2019 (black lines). (Adaptation of graph appearing in https://www.climatestations.com/wp-content/uploads/2020/01/ny2019.gif. Data from https://w2.weather.gov/climate/index.php?wfo=okx).

## 16.2 The Kinetic Energy of Molecules Determines Heat and Temperature

If you sit outdoors on a sunny day, you can feel the sun's radiant energy warming your skin. Radiant energy from the sun, and variation in how much of it reaches different parts of the Earth's surface, is the primary cause of variation in temperature (and much more). Radiant energy is the energy of the electromagnetic spectrum; as it heats molecules, it is transformed into kinetic energy. **Heat** is a measure of the total kinetic energy—the energy of motion—of the molecules in a substance, while **temperature** is a measure of the average kinetic energy of those molecules. The same terms and the same concepts apply at global scales as at the scale of a leaf (see Figure 3.25). Radiant energy from the sun heats objects (such as plants, the soil, ice, or the oceans) directly. This energy is mostly in the shortwave part of the electromagnetic spectrum (wavelengths < 1000 nm) dominated by visible light (wavelengths of 400–700 nm; see Figure 2.3A). As objects are heated by sunlight, they emit longwave, infrared radiant energy (wavelengths > 700 nm, IR or LW

radiant energy) in proportion to $K^4$, the fourth power of their absolute temperature in degrees Kelvin, becoming a secondary source of heating for their surroundings. The warming of physical objects by sunlight and the subsequent emittance of that heat energy as longwave radiation (that is, its storage and later release) creates lags in temperature change, which we explore in more detail below. The transport of that heat energy from one place to another by atmospheric and oceanic currents by convection is also a key element in global climates.

The temperature at any spot on the Earth's surface is determined primarily by the amount of radiant energy it receives from the sun. That amount is determined, in turn, largely by the angle of the Earth's surface at that spot relative to the sun's rays, as well as by atmospheric conditions. Imagine a spot at sea level on the equator at noon during the vernal or autumnal equinox, when the sun is directly overhead. At the top of the atmosphere at that spot, the Earth receives approximately 341 $Wm^{-2}$ $year^{-1}$, or 2 langleys (ly) of energy per minute (1 ly = 1 calorie/$cm^2$), as high-energy, short wave (mostly visible and near-infrared), radiation. If we follow the energy striking the top of the Earth's atmosphere, we can trace what happens to that energy (Figure 16.2). Almost a third of it returns to space, reflected by the Earth's surface (23 $Wm^{-2}$ $year^{-1}$, about 7%) and by clouds (54 $Wm^{-2}$ $year^{-1}$) and scattered by the atmosphere (25 $Wm^{-2}$ $year^{-1}$). Almost half (161 $Wm^{-2}$ $year^{-1}$) is absorbed by the Earth's surface. The rest of the energy entering the atmosphere is absorbed by clouds (14 $Wm^{-2}$ $year^{-1}$) and by the greenhouse gases and dust (65 $Wm^{-2}$ $year^{-1}$) in the atmosphere.

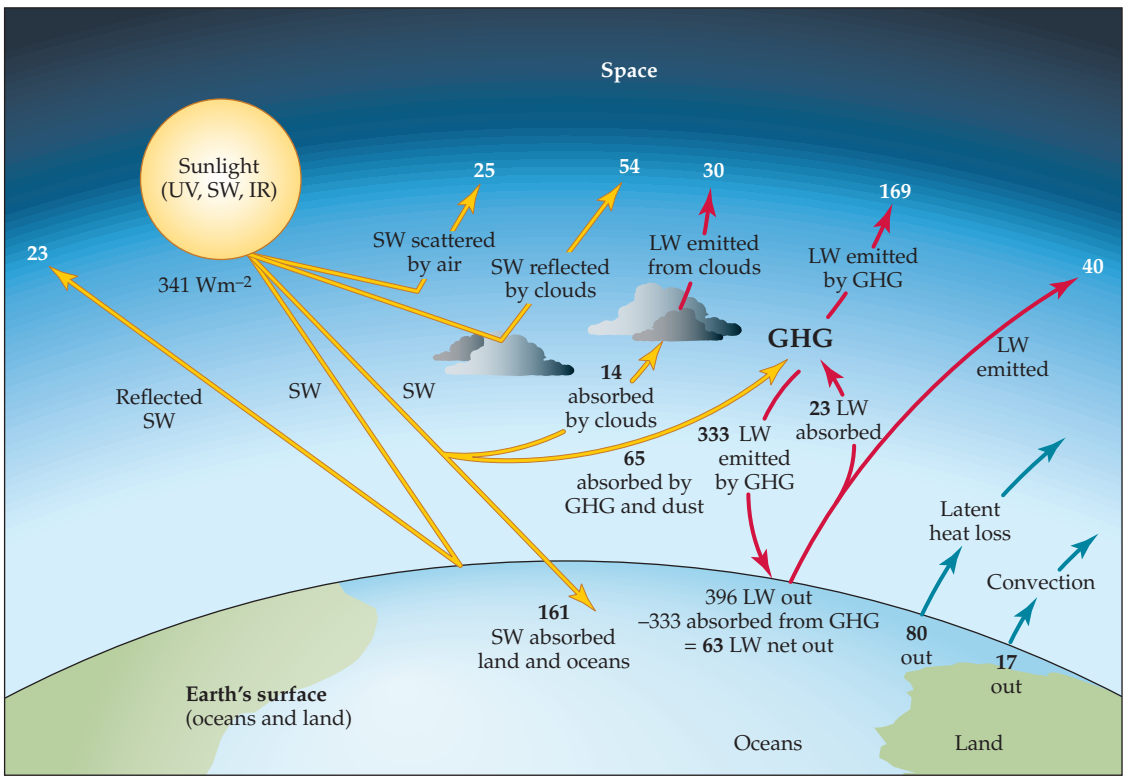



**Figure 16.2** The radiant energy balance of the Earth. The energy reaching the top of the atmosphere is 341 $Wm^{-2}$/year and much of the energy is in the visible (short wave, SW) part of the spectrum. This energy is reflected by the Earth's surface and by clouds and scattered by the atmosphere, so almost a third returns to space without being absorbed. Of the remainder, almost half is absorbed by the surface and a quarter by clouds and the atmosphere, warming both. Longwave radiant energy (LW or infrared, IR) is emitted by the warmed surface, heating the greenhouse gases (GHG) in the atmosphere further as well as being lost to space. Convection and latent heat exchange also transfer heat energy from the surface to the atmosphere. The total amount of energy absorbed by the surface, atmosphere, and clouds (as shortwave radiation) and the total amount that leaves the planet (as longwave radiant energy) are in balance with each other. The total amount of energy entering the top of the atmosphere (as shortwave radiation) is equal to the sum of the outgoing shortwave radiation (from reflection and scattering) plus the outgoing longwave radiation. (After K. E. Trenberth et al. 2009. *Bull Amer Meteor Soc* 90: 311–324; M. C. MacCracken. 1985. In *Projecting the Climatic Effects of Increasing Carbon Dioxide*. M. C. MacCracken and F. M. Luther [Eds.], pp. 1–23. U.S. Department of Energy. Er-0237. Washington, D.C.)

The Earth's surface emits energy in the longwave, infrared part of the spectrum, radiating a net difference of 63 $Wm^{-2}$ $year^{-1}$, transferring an additional 80 $Wm^{-2}$ $year^{-1}$ as latent heat loss and 17 $Wm^{-2}$ $year^{-1}$ by convection to the atmosphere and clouds, and emitting 40 $Wm^{-2}$ $year^{-1}$ to space. The **greenhouse effect** results when **greenhouse gases** in the atmosphere absorb some of the infrared radiant energy emitted by the Earth's surface, radiating 333 $Wm^{-2}$ $year^{-1}$ back to the Earth's surface while 169 $Wm^{-2}$ $year^{-1}$ is emitted to space (Figure 16.3). Secondary heating of the Earth's surface then occurs by this longwave radiant energy from the greenhouse gases in the atmosphere, which were warmed by emission of infrared radiant energy from Earth's surface.

This reabsorption and reradiation of solar energy through the greenhouse effect is a natural aspect of Earth's atmosphere, and it is critical for life on Earth. Without the greenhouse effect, the average temperature of the Earth's surface would be bitterly cold everywhere, at approximately –21°C (a very deep freeze), rather than its actual balmy average temperature of 14°C. Secondary heating by longwave energy from the atmosphere accounts for about two-thirds of the total radiant energy received at the surface. The total outgoing longwave radiation from the Earth's surface, the atmosphere, and clouds that returns to space is about 69% of the energy received initially as shortwave radiant energy from the sun.

One way to picture how the greenhouse effect works is by thinking about the analogy of an automobile parked outside on a sunny day. Like all analogies and models, this one is imperfect, but makes some of the major aspects of the greenhouse effect more tangible. Visible, shortwave solar energy—light—passes through the windshield and windows into the vehicle, warming the seats, dashboard, and everything inside. As they absorb this radiant energy, they warm up and radiate in the infrared (IR) part of the spectrum. However, the IR wavelengths cannot leave the car through the glass, because although glass is transparent to visible light, it is opaque to IR wavelengths. That is why we use glass in windows—we can see through it, but it keeps heat inside (although only partially). The inside of the vehicle keeps getting warmer from the incoming sunlight, and the IR radiant energy is trapped, further warming the inside (Figure 16.4). This happens even if the windows are slightly open, because the convective heat loss (see Chapter 3) is insufficient to keep up with the greenhouse effect of the energy trapped by the glass. (That is why children and pets in a parked vehicle in the sun can easily overheat and die within minutes, even with the windows open a bit.)

In the Earth's thermal balance, incoming energy must eventually be balanced by outgoing energy. Temporary imbalances will result in the Earth—or any other object, living or nonliving—warming or cooling, but eventually a new equilibrium will be reached as the Earth's emission of longwave radiation changes according to its new temperature. Thus, the

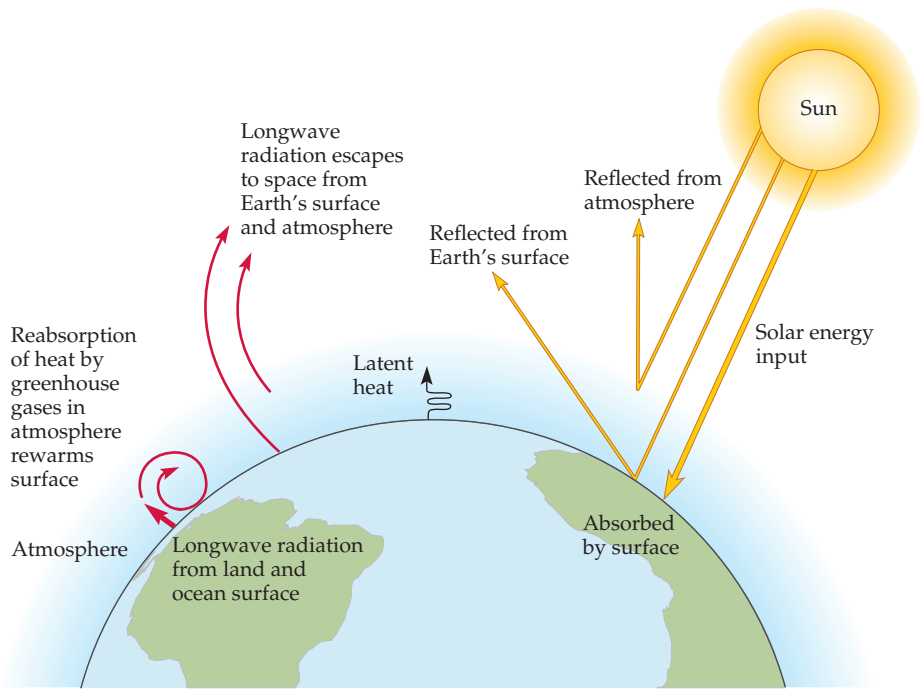



**Figure 16.3** The greenhouse effect. About 70% of the incoming shortwave radiant energy from the sun is absorbed by the Earth's surface, clouds, and atmosphere; the remainder is scattered and reflected back to space by the surface and the atmosphere. About half of the absorbed energy heats the Earth's surface. Some of that energy is radiated by the surface as longwave (primarily infrared) radiant energy. Most of that radiated energy is absorbed by atmospheric greenhouse gases, which then radiate it back to further warm the Earth's surface. A small portion of the longwave energy is radiated by the atmosphere and escapes to space. The other half of the energy absorbed by the Earth's surface is used in the evaporation of water, which adds water vapor to the atmosphere and initially cools the surface. When that water vapor condenses to form clouds, the latent heat energy removed from the Earth's surface is transferred to the atmosphere, driving the creation of storms and releasing precipitation.

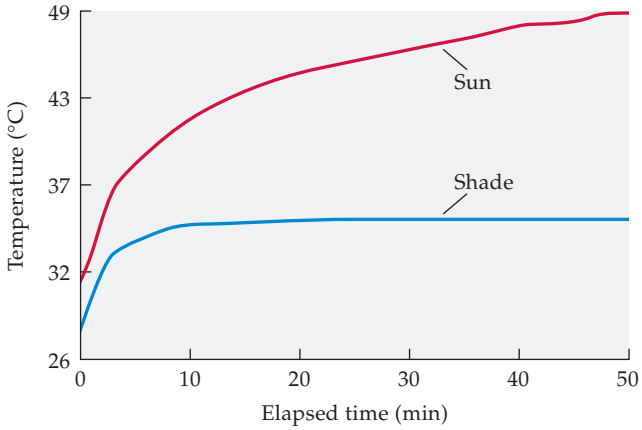



**Figure 16.4** A car parked outdoors in the sun offers an analogy to the greenhouse effect in common everyday life. Imagine that the car has its windows closed or slightly open, with outdoor temperatures at 27°C. The sunlight, consisting largely of visible, shortwave radiant energy, enters through the windshield and windows of the car. The light is able to pass through the glass because it is transparent to visible light, and the interior of the car absorbs the incoming radiant energy. As the seats, dashboard, etc., heat up, they increasingly radiate heat in the longwave, infrared (IR) part of the spectrum, but IR energy cannot go through glass (which is largely opaque to IR) and the heat energy gets trapped in the car. The interior of the car continues to heat up as long as sunlight keeps coming in. Even with the windows slightly open, the interior will heat rapidly in the sunlight to dangerously hot temperatures. The glass is analogous to the greenhouse gases in the atmosphere, which are transparent to shortwave radiant energy but absorb in the longwave part of the electromagnetic spectrum, trapping heat on Earth's surface and in the atmosphere.

shortwave radiant energy absorbed by the atmosphere and the Earth's surface (161+14+65; see Figure 16.2) is balanced by the longwave energy (169+30+40, with rounding error) emitted. Although the natural properties of Earth's atmosphere create the greenhouse effect, humans have been changing those properties and altering the natural greenhouse effect. The most popular usage of the term *greenhouse effect* refers to the rise in the temperatures at Earth's surface resulting from the increases in absorption and emission of longwave radiant energy by atmospheric $CO_2$ and other greenhouse gases caused by human activities. It is this *increase* in the greenhouse effect that is leading to global climate change, as we discuss later in this chapter.

## The sun's angle is the main factor determining the radiant energy received at Earth's surface

The combination of shortwave and longwave radiation determines the ambient temperature at the Earth's surface. The energy balance we just examined describes a spot at sea level on the equator at noon during the vernal or autumnal equinox with maximum solar energy input. As you travel away from that spot, several factors act to determine the amount of incoming radiation and the subsequent temperature at Earth's surface. Most important is the angle of the sun, which determines the maximum amount of incoming shortwave radiation. Over the course of the day, the Earth rotates on its axis, and the sun's rays hit the surface at different angles. The solar **angle of incidence** is the angle that a ray of sunlight makes with a line perpendicular to the surface. The smaller the angle of incidence—the closer to 0° from the horizontal and the lower the sun appears in the sky—the larger the area over which a given amount of solar energy is spread. Consequently, when the incoming sunlight is at a low angle, any given square meter of surface receives less energy than it would if the sun were at a more direct angle overhead, closer to 90° from horizontal to that perpendicular plane (Figure 16.5). In addition, at lower angles the incoming solar energy must travel a longer distance through the atmosphere, and so more energy is absorbed by the atmosphere and less reaches the Earth's surface.

As one moves north or south of the equator (at the equinox), the angle of incidence, even at noon, decreases. The actual angle depends on the time of year and how far one is from the equator. The Earth's axis has a constant tilt (over the short term, but see below) relative to the plane in which it revolves around the sun. Seasons are created by differences in the angle of incidence during the year-long progression of the Earth around the sun, which determines the amount of incoming solar energy at any spot on Earth's surface (Figure 16.6A). The Northern and Southern Hemispheres are tilted toward or away from the sun at the opposite extremes of the Earth's trajectory around the sun, at the solstices in December and June. At the equinoxes, in September and March, the Earth's axis is tilted sideways to the plane on which it revolves around the sun. At these times, neither pole is tilted toward or away from the sun, and day and night length are approximately equal to each other, and equal everywhere across the globe. (*Equi-nox* means "equal night.") At the equinoxes, the Earth is approximately halfway through its trajectory between the perihelion and aphelion of its revolution around the sun (explored in more detail below). When the sun is closer to being directly overhead in the Northern or Southern Hemisphere, it is summer in that hemisphere. In the opposite hemisphere, low solar angles at that time result in decreased radiant energy input, and it is winter. This change in solar angle with the seasons is greatest at the poles and least at the equator. As a result, there are large changes in radiant energy input over the course of a year at high latitudes. In contrast,

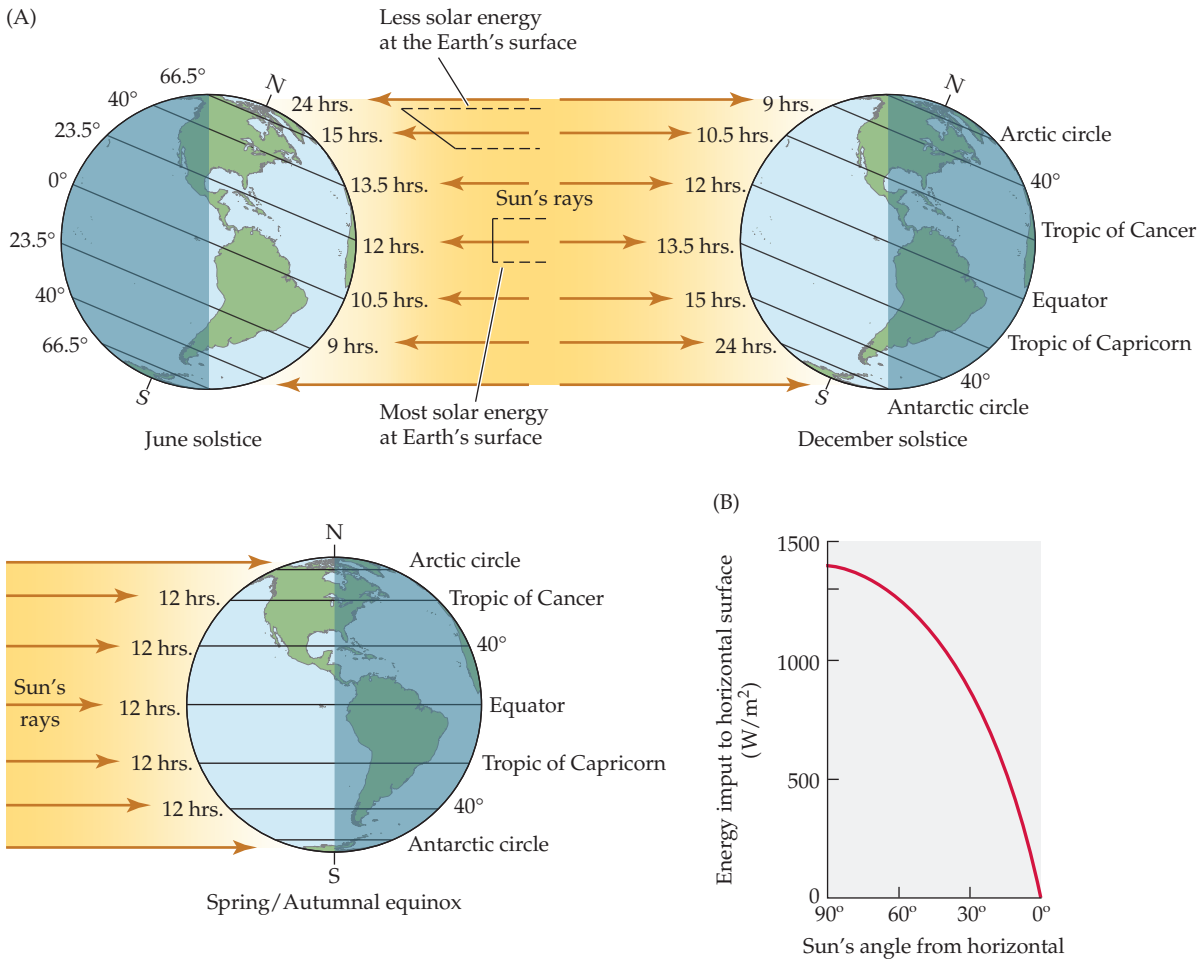
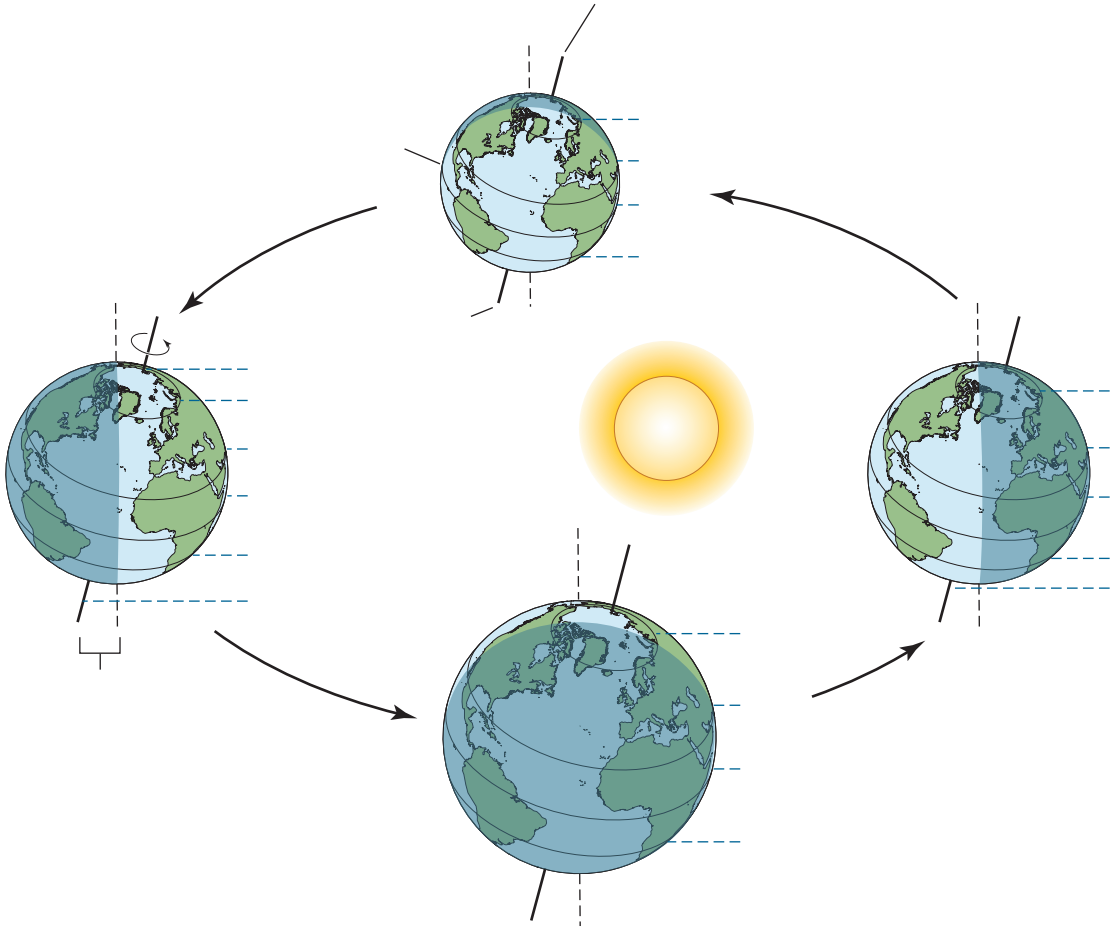
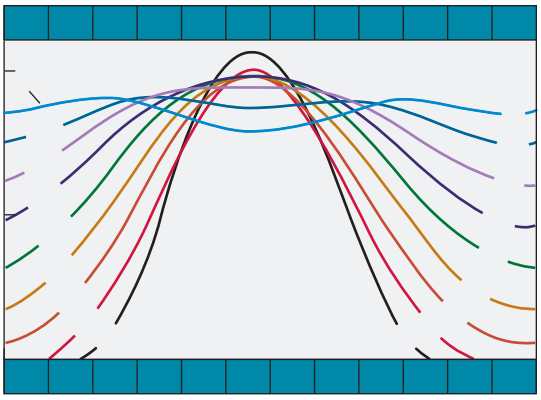



**Figure 16.5** (A) Differences in the angle and duration of sunlight reaching different latitudes on Earth's surface at the June Solstice (summer in Northern Hemisphere, winter in Southern Hemisphere), December Solstice (summer in Southern Hemisphere, winter in Northern Hemisphere), and at the equinoxes in September and April. The angle of the sun determines the amount of energy received at each part of Earth's surface; steeper angles spread the same amount of energy over larger areas, reducing the amount of energy received at any one spot. At the equinoxes the length of the day and length of the night are equal at any one spot on Earth, and also daylength is the same everywhere on Earth from the poles to the equator. (B) The amount of energy received from the sun at Earth's surface depends on the angle of the sun from the horizontal. At 0° the sun would be visible at the horizon; at 90° it is directly overhead and shines perpendicular to the surface.

equatorial regions have nearly constant solar energy inputs throughout the year (Figure 16.6B). These effects combine to produce distinct patterns of seasonal temperature variation across the globe (Figure 16.7).

Another important aspect of climate is **temperature lag**—delay in the effects of warming and cooling due to the storage of energy in the atmosphere, ground, and ocean. For example, it is usually warmer in the afternoon than at noon, even though the sun is lower in the sky and incoming solar radiation is less than at noon, and warmer in late afternoon than when the sun is at a comparable angle in the morning with equal radiant energy input. The coldest month of winter in the Northern Hemisphere is usually not December (when the sun is lowest and days are shortest), but January, when solar radiant energy input has increased; conversely, the hottest summer months in temperate regions are July and August, not June (when the days are longest and the sun most directly overhead). The lag in temperature is caused by heat storage and results in the gradual increase in temperatures as the summer solstice approaches, and the ground, ocean (or large lakes), and atmosphere absorb and store energy, and the release of heat (as longwave radiant energy) by those bodies as one passes the summer solstice later in summer and into autumn (and similarly with day and night). As the winter solstice nears, the water, land and air cool and solar radiation decreases. As the Earth passes the





Winter

0°

43°

**December solstice**

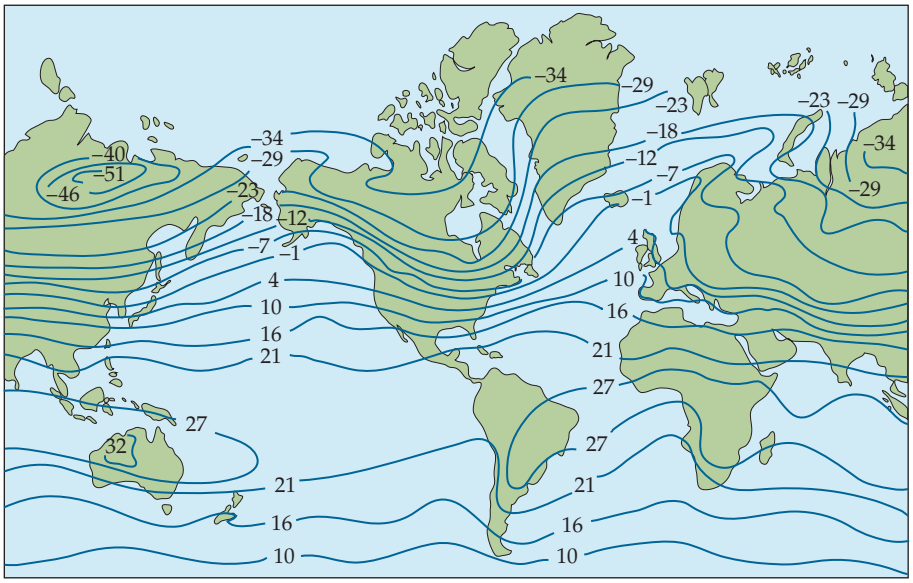
66.5°

90°

23.5°

Summer

(A)




(B)

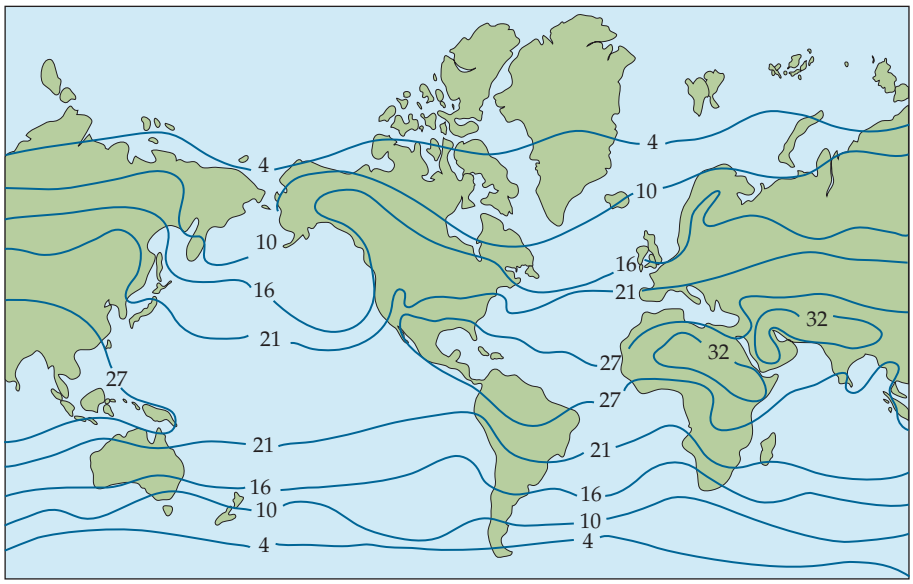



**Figure 16.7** Mean temperatures (°C) across the globe in (A) January and (B) July. (After G. R. Rumney. 1968. *Climatology and the World's Climates*. Macmillan: New York.)

winter solstice, solar energy input increases, but the water and land absorb this energy, and it is the coldest time of the year. Eventually, as land and water continue to absorb more energy and release it as long wave radiant energy, the Earth warms again in spring.

Picture a place somewhere at a temperate latitude at the winter solstice. Incoming solar radiation is at a minimum. As the Earth revolves around the sun and the year progresses, day length increases and the sun is more directly overhead at midday. The ground absorbs solar energy and begins to warm. The ground has a certain **specific heat**—the amount of heat energy that is required to warm 1 g by 1°C. Water has a higher specific heat than the ground, so the ocean takes longer to warm up than nearby land. Air has a far lower specific heat than either land or water and warms up and cools down more readily than they do. (The specific heat of soil is in part due to the water in the soil; clay soils warm up more slowly in spring and cool more slowly in winter than do sandy soils, for that reason; see Chapter 4.) As this imaginary spot on the Earth passes the summer solstice and incoming radiation begins to decline, the longwave output from the now-warmed ground and nearby bodies of water continues to warm the air, keeping air temperatures relatively high. Eventually, as solar input declines, the reverse pattern occurs: air, ground, and finally water become cooler. A high specific heat results in slow cooling, just as it causes slow warming, and cools and warms less completely. Coastal environments, influenced by the adjacent bodies of water, consequently cool and warm more slowly and less than those in the interior of continents, so they experience more moderate temperatures. The water ameliorates hot summer temperatures as cooler offshore breezes convectively reduce heat on land. In autumn, the water releases heat gradually, slowing temperature declines on land by providing warming breezes. These breezes may feel cold to a person, but they are warm relative to the ground and the air over it. One implication of these lags is that times of maximum or minimum solar radiation are not the times of highest or lowest temperatures, either daily or seasonally.

The magnitude of daily variation in temperature—the range from nighttime to daytime temperatures—depends in part on local atmospheric conditions, particularly the amount of moisture in the air. At night, there is no incoming shortwave radiation. We do not immediately freeze after sunset, however, because the ground underfoot and the air surrounding us hold heat

and release it as infrared radiant energy all night. Moist air can hold more heat than dry air, as we will see below. A place like Tokyo, adjacent to the Pacific Ocean with typically humid air, usually has nighttime temperatures that are only a little cooler than daytime temperatures. In contrast, Tucson, Arizona, with very dry air, may have dramatic swings, with nighttime temperatures 20°C or more cooler than those during the day. As altitude increases, the density of the air decreases. As a result, areas at high elevations with fewer air molecules (particularly, fewer molecules of greenhouse gases), such as mountaintops, also tend to have large swings in temperature from day to night. Plants must be adapted to tolerate the variation in temperature, as well as the average conditions, where they live. Interestingly, the **troposphere**, the lowest layer of the atmosphere in which living organisms reside, is not equally deep everywhere. At the equator, the troposphere extends about 17 km up, but at the poles it is only 9 km from Earth's surface to the top of the troposphere.

Elevation and latitude jointly affect climate. Equatorial regions receive nearly constant amounts of solar radiation throughout the year and do not have either cold winters or hot summers. But high-elevation regions at the equator, such as the Ecuadorian Andes in South America and Kilimanjaro in Africa, can experience enormous temperature swings over each 24-hour period, approaching those experienced between summer and winter at mid-latitudes, although obviously they also receive much more radiant energy during the day.

## There are long-term cycles in Earth's path around the sun that affect radiant energy at Earth's surface

Over long time scales, there are three kinds of changes in the Earth's orbit around the sun that influence climate, collectively called the **Milankovitch cycles** (or **Croll-Milankovitch cycles**; Figure 16.8). Milutin Milanković was a Serbian mathematician and climatologist whose life's work was modeling the correlation between incoming solar radiation, cycles in Earth's path around the sun, and ice ages in Earth's past, based on mathematics and physics (Milanković 1941). His discoveries were made in the early twentieth century but were not accepted for over 50 years. Support for Milancović's work came in 1976 from researchers reconstructing

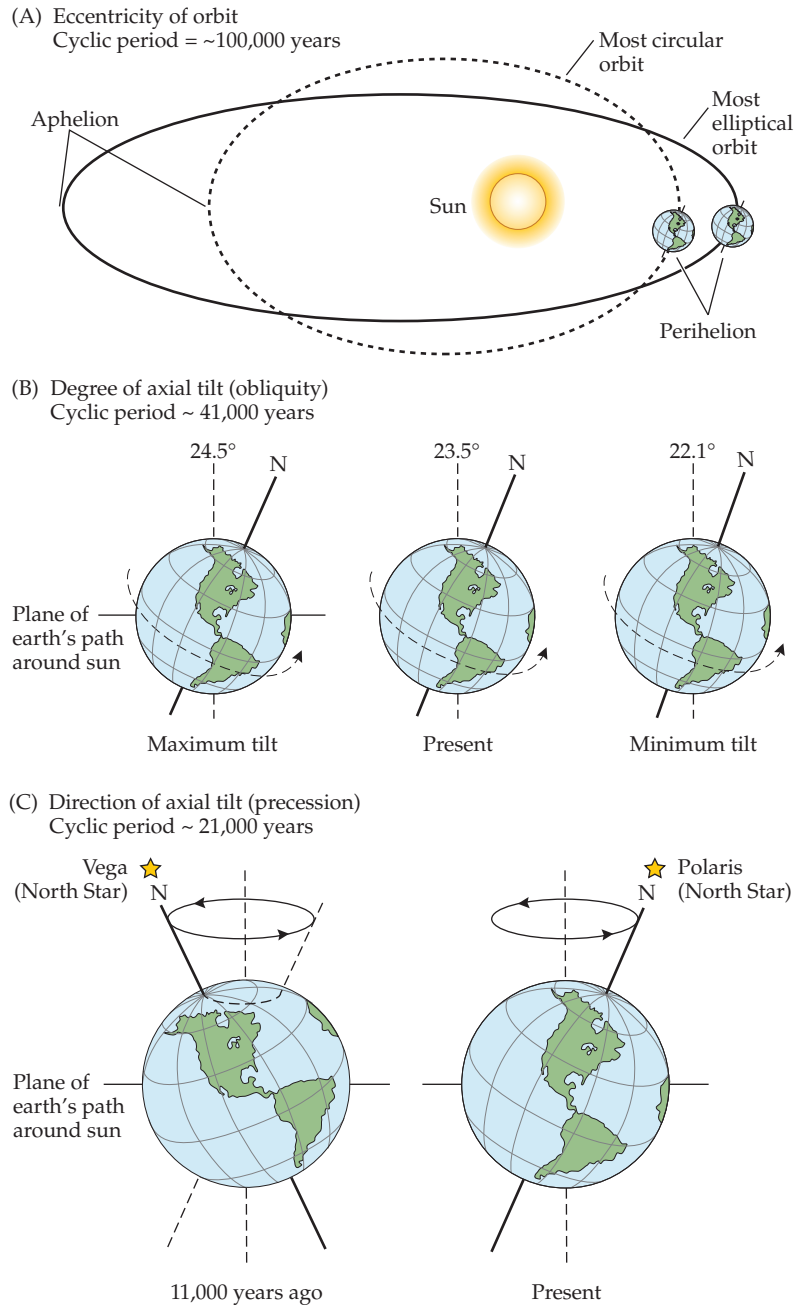



**Figure 16.8** Milutin Milanković, a Serbian astrophysicist, building on the ideas of James Croll, a Scottsman who rose from very humble origins to become a scientist, posed a theory that related cycles in Earth's orbit around the sun to long-term climate change, in particular to ice ages. These cycles, sometimes called Milankovitch cycles, include three aspects: periodic changes in (A) the eccentricity (ellipticity) of the Earth's orbit, (B) the degree of axial tilt (obliquity), and (C) the direction of axial tilt (precession). The Earth is farthest from the sun at the aphelion, and closest to the sun at the perihelion. The solstices do not occur exactly at the aphelion and perihelion, but rather whent the sun appears to be directly over the Tropic of Cancer or Capricorn; likewise the equinoxes occur when the sun appears to be perpendicular to a horizontal plane at the equator. The position of the Earth at different points in each of these cycles affects the distribution of solar radiation received at different latitudes on Earth's surface at different times of the year. These cycles interact to determine long-term climatic patterns on Earth. (After D. M. Gates. 1993. *Climate Change and Its Biological Consequences*. Oxford University Press/Sinauer: Sunderland, MA.)

past climates using deep-sea sediment cores (Hays et al. 1976). It was not until 1999 that his ideas were completely confirmed (Kerr 1999; Rial 1999). The three cycles determine the amount of sunlight reaching Earth's surface, seasonal differences at different latitudes, and the advance and retreat of Earth's ice ages. Atmospheric $CO_2$ in Earth's past has lagged climate changes caused by these cycles by a short period of time (hundreds of years). As the atmosphere warms (at first due to greater solar radiation reaching Earth's surface), this warming increases the $CO_2$ (as well as water vapor) in the air, further increasing the warming (due to the greenhouse effect), creating a positive feedback. Only in the past 150 years has $CO_2$ preceded climate warming, because humans (rather than natural processes) have released great quantities of $CO_2$ into the atmosphere.

The longest of the orbital cycles is an approximately 100,000-year cycle of change in the shape of the Earth's orbit around the sun, called **eccentricity**, from nearly circular to a more elliptical trajectory. Currently the orbit is almost circular with the sun near the center of the circle, so the difference between the Earth's nearest point to the sun—its **perihelion**—and its farthest point from the sun—its **aphelion**—is only 3.5%. When the orbit becomes more elliptical over the course of this cycle of eccentricity, and the sun is closer to one side of the ellipse than the other, the difference between the perihelion and the aphelion may be as great as 30%. (The Earth reaches the point farthest from the sun—the current aphelion—sometime in July, about a month past the solstice, and reaches the perihelion in January.) Such changes in orbital shape affect the seasonal amounts of incoming solar radiation at temperate and polar latitudes, which in turn affect the range of seasonal variation, alter Earth's climate, and influence the advance and retreat of Earth's glaciers. What may be surprising is that these profound changes in the climate are due to relatively small changes in the total amount of radiant energy, only about a 0.2% change in total solar insolation at the extremes of the three cycles. However, there is a far greater variation in the amount of incoming solar radiant energy in mid-latitudes, of up to 25%. It is the change in seasonality, particularly in the Northern Hemisphere, that is most influential in precipitating ice ages rather than the average total insolation reaching the top of Earth's atmosphere.

The next longest pattern of variation is a 41,000-year cycle in the degree of tilt of the Earth's axis, called **obliquity**, from 22.1° to 24.5°. Currently the tilt is in between, at 23.5°. These changes in the axial tilt will greatly affect the range of seasonal variation because changes in the angle of the sun will change the amount of solar energy input at any given spot at temperate and polar latitudes. When the axial tilt is less, the incoming solar radiation is more even, between tropical and polar environments and also between summer and winter. Less tilt leads to the accumulation of glaciers and therefore to the advance of an ice age. That is in part because warmer winter air holds more moisture, producing more snow and ultimately the growth of glaciers, and cooler summers melt less of the accumulated ice each year. Feedback effects "snowball" the accumulation of glacial ice. Greater tilt leads to more extreme seasonality in the Northern Hemisphere, in particular to warmer summers that combine with feedbacks to the greenhouse effect and lead to melting glaciers and the end of ice ages.

Finally, there is a 23,000-year wobble of the Earth's axis, called **precession**. The North Pole is currently pointing to Polaris, which we call the North Star; 11,000 years or so ago, the North Pole pointed to a different star, Vega. This wobble is similar to what you would see on a top that is beginning to slow down and its stem no longer points straight up but is beginning to circle around. The result of this wobble for Earth's climate is a change in whether the Northern or the Southern Hemisphere is pointing toward the sun during the time of the Earth's nearest approach to the sun. At the opposite point of this precession, the timing of summer and winter is opposite what it is currently, and halfway through the cycle, the equinoxes occur when the solstices now occur. The Earth is now a little bit closer to the sun at the winter solstice in the Northern Hemisphere (the summer solstice in the Southern Hemisphere) than at the opposite point in its orbit. Changes in the direction of the axial tilt will, like the two other cycles, also affect the intensity of seasonal variation. Northern Hemisphere winters are currently milder than Southern Hemisphere winters at the same latitude. At the opposite end of this cycle, when the North Pole is pointed toward the star Vega, Northern Hemisphere winters will coincide with the point when the Earth is farthest from the sun, and summer in the Northern Hemisphere will be when the Earth is closest to the sun, making for colder winters and hotter summers. Because there is far more land mass in the Northern Hemisphere than in the Southern Hemisphere, this also has implications for the absorption and movement of heat around the globe.

What causes the Milankovitch cycles? The changes in eccentricity of Earth's orbit are due to the gravitational tugs of the other planets in our solar system; obliquity and precession are affected by the gravity from the pull of the sun and moon on Earth's equatorial bulge. These three cycles act to modify seasonality, and they can work together or in opposition to magnify or minimize seasonal variation (Figure 16.9). While the effects may be small, they are sufficient to cause large-scale and long-term changes in the Earth's climate and are likely responsible for the cycles of glaciation that the

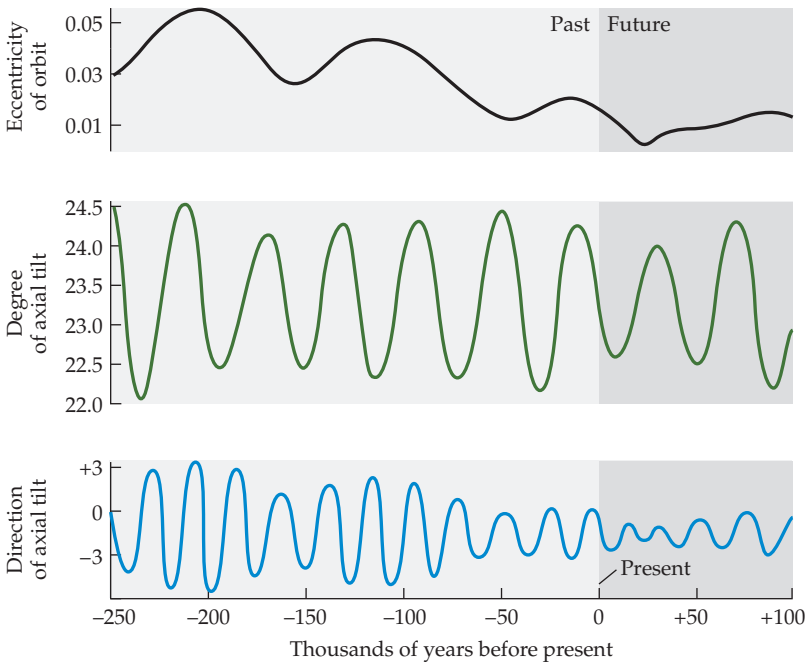



**Figure 16.9** Past and predicted Croll-Milankovitch cycles. Sometimes the cycles can reinforce one another, as when, for example, both ellipticity and the degree of axial tilt are large. At other times they can cancel one another. (After D. M. Gates. 1993. *Climate Change and Its Biological Consequences*. Oxford University Press/Sinauer: Sunderland, MA.)

Earth has experienced over the last million years (see Figure 17.10). We look at the most recent of these glaciations and its aftermath in Chapter 17. These cycles have occurred several to many times since the origin of humans as a species approximately 300,000 years ago.

Other factors have smaller influences on the solar radiation reaching Earth. One of these is the 11-year sunspot cycle. Over this cycle, the amount of solar radiation reaching the Earth varies by up to 0.1% of full sunlight. A 22- to 23-year double sunspot cycle is also known and is related to reversals of the solar magnetic field. These cycles probably do not directly affect plant growth or local temperatures. Rather, they affect atmospheric circulation and **precipitation** patterns, resulting in indirect effects on plants due to complex and subtle changes in weather patterns.

Recently there has been extensive debate about global warming, an overall rise in global temperatures. Part of the debate stems from a genuine scientific question: Is the recent warming of the climate likely to be caused by human addition of greenhouse gases to the atmosphere, or does it reflect natural cyclic changes and perhaps some random variation in the weather? Without considering the actual data, either hypothesis (or both) could be plausible. The warmer temperatures we are experiencing could be part of a long-term trend, or they might be a short-term blip that will soon be reversed. If this warming is part of a trend, to what extent is it due to human actions? Certainly, it is very difficult to separate natural variation in weather and climate from anthropogenic (human-caused) changes. In any case, the observed warming may be a harbinger of an increasingly warm climate. We examine this issue in greater depth later in this chapter.

## 16.3 Precipitation Patterns Vary across the Earth

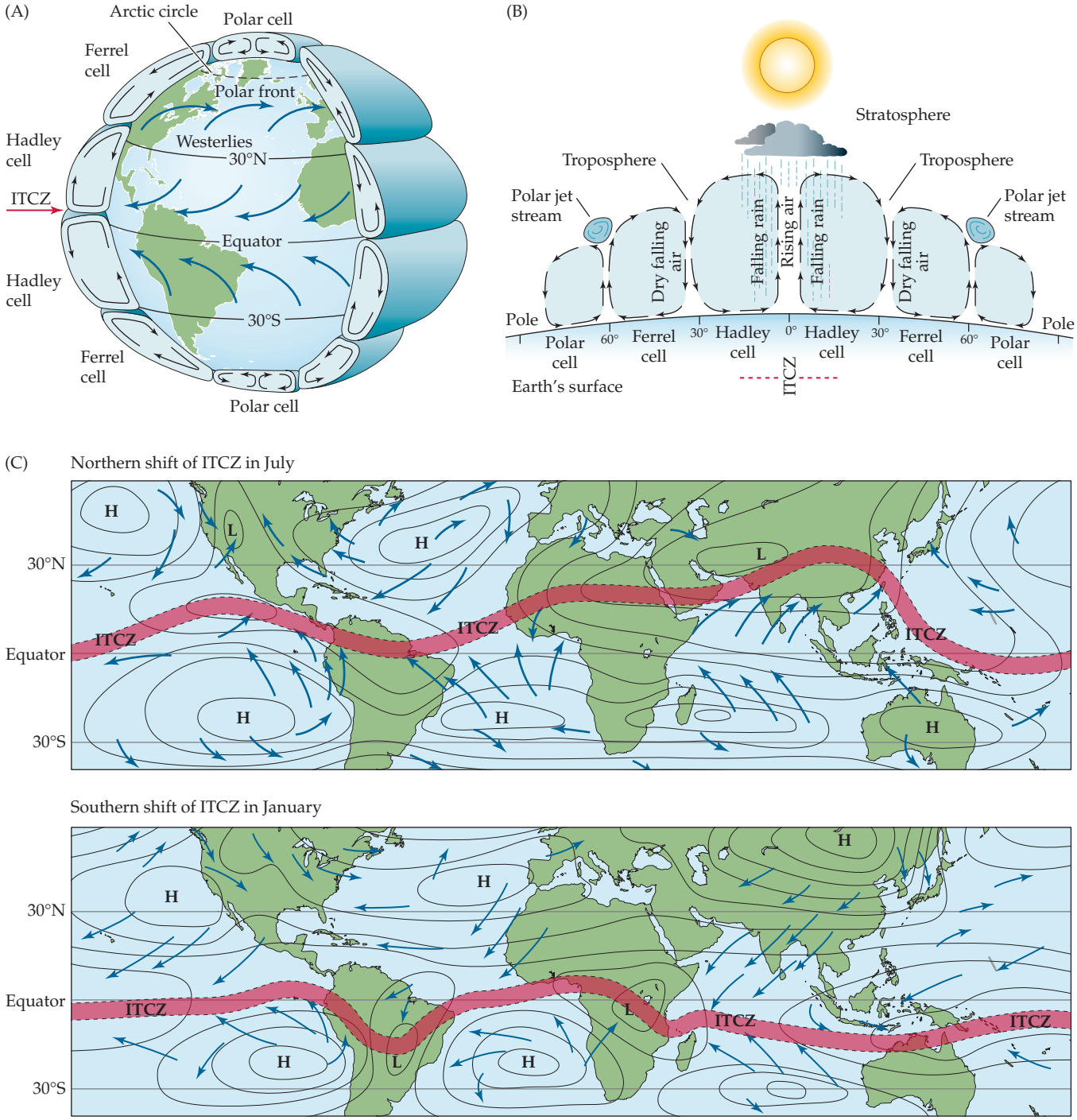
Besides temperature, the other major aspect of climate is water: **precipitation**, the amount and pattern of rainfall and snowfall, and **relative humidity**, the water vapor in the air as a percentage of the amount in fully saturated air at the same temperature. Precipitation has a major determining effect on the type and amount of vegetation found in a region. Humidity levels are also important ecologically. Fog (which occurs only at saturating humidity) can be a critical source of moisture for plants in certain habitats, and humid air contains both more moisture and more heat than dry air at a given temperature. Water molecules in the atmosphere are very effective at absorbing longwave radiation and have a high specific heat. Therefore, humid air acts as a reservoir of stored heat. Clouds also absorb, radiate, and reflect incoming solar energy. A number of factors determine how precipitation and water vapor are distributed over the globe and during the year.

### Global patterns are determined by air moving in three dimensions at huge spatial scales

Enormous cells of air rise over the tropics, moving moisture and energy up from Earth's surface to the top of

the troposphere (the layer of the atmosphere closest to Earth's surface), then they traverse great distances away from the tropics at the boundary between the troposphere and the stratosphere, fall back down to Earth's surface at roughly 30°N and 30°S, and then rush across the surface back toward the tropics. These giant conveyor belts are the **Hadley cells** (Figure 16.10A,B). This pattern of air movement results in large amounts of precipitation and tropical forests surrounding the equator, and dry air and deserts at the Tropics of Cancer and Capricorn at 30° north and south of the equator. What causes these patterns?

Warm air is less dense than cold air and therefore rises. That is why hot air balloons go up and why it is warmer near the ceiling of a room than at the floor. As explained previously, solar radiation is greatest in the tropics, warming the air there. This warming creates a pattern of rising air in a belt around the Earth in the tropics. The air begins to cool as it rises to higher altitudes



above Earth's surface because air pressure decreases with altitude. As a given packet of air rises, there is less atmosphere above it pushing down on it. The rising warm air expands, because gases at lower pressures expand, filling up a greater volume the farther it moves from the surface. But the expanding air also cools because kinetic energy is used to push the air molecules farther apart.

The warm tropical air masses pick up moisture from the land surface, both from transpiration by plants and by evaporation from the ground, and in large amounts evaporating from the oceans. When water molecules are warmer, they have the kinetic energy to remain in the air as water vapor; when they cool, they become liquid water and precipitate out of the air. Although we often say that warmer air 'holds' more moisture, this shorthand is not really accurate; it is the energy of the water molecules that determines whether water is in the vapor or liquid phase. Warm air can therefore contain more water vapor than cooler air can. You can see this for yourself by observing moisture condensing on a cold beverage glass on a warm, humid day. As the water vapor in the air near the surface of the glass cools, it condenses as liquid on the cool surface of the glass. You can also see this condensation happening in the natural world, particularly over mountains, as clouds form as air is forced upward over the mountains and cools. Rising air cools at the **adiabatic lapse rate** when no heat is exchanged with surrounding parcels of air. Air saturated with moisture cools at 5.4°C/1000 m, the saturated (or moist) adiabatic lapse rate, while dry air cools more rapidly, at the dry adiabatic lapse rate of 9.8°C/1000 m. As moist air rises and cools, the condensing water releases heat to the air—the opposite of what happens when a surface is cooled by evaporating water; the latent heat gain (see Chapter 3) slows down the rate of cooling.

The air rising over the tropics in the Hadley cells cools as it rises, reaching the point, called the **dew point**, at which water molecules in the air have lost the kinetic energy to remain a gas and they condense to a liquid state. The water vapor condenses and forms clouds and precipitates in liquid form, falling as rain over the tropics (see Figure 16.10B). This is why tropical rainforests exist: with warmth and light from high incoming solar radiation, and supplied with plentiful rainfall by the Hadley cells, high NPP is possible all year.

The continual addition of new air rising from below means that this rising mass of air in the Hadley cells has to go somewhere; it can't just fall back down directly, because the air below it is pushing upward. The rising air is forced to move away from the tropics at the top of the troposphere, moving either northward or southward. As this air moves horizontally away from the equatorial regions where Earth is most directly warmed by the sun, it cools further because it is now receiving less radiant energy. By the time it has traveled to about 30° latitude (still at the top of the troposphere), it has cooled sufficiently to become denser and begin sinking. The falling air warms adiabatically as it falls and is subject to the pressure of the air above it, reversing the process that occurred as it rose in the tropics. Any liquid water encountered is absorbed by the warm, falling dry air, which releases little of it, creating intensely arid regions where the air in the Hadley cells rejoins Earth's surface. This is why many of the major deserts of the world occur in two bands, at 30°N latitude (including the Arabian, Chihuahuan, and Sonoran Deserts, the Thar Desert of western India, the great deserts of Iran and Afghanistan, and the Sahara) and 30°S latitude (including the Atacama Desert of Chile, the Kalahari Desert of southern Africa, and the Great Sandy and Great Victoria Deserts of Australia).

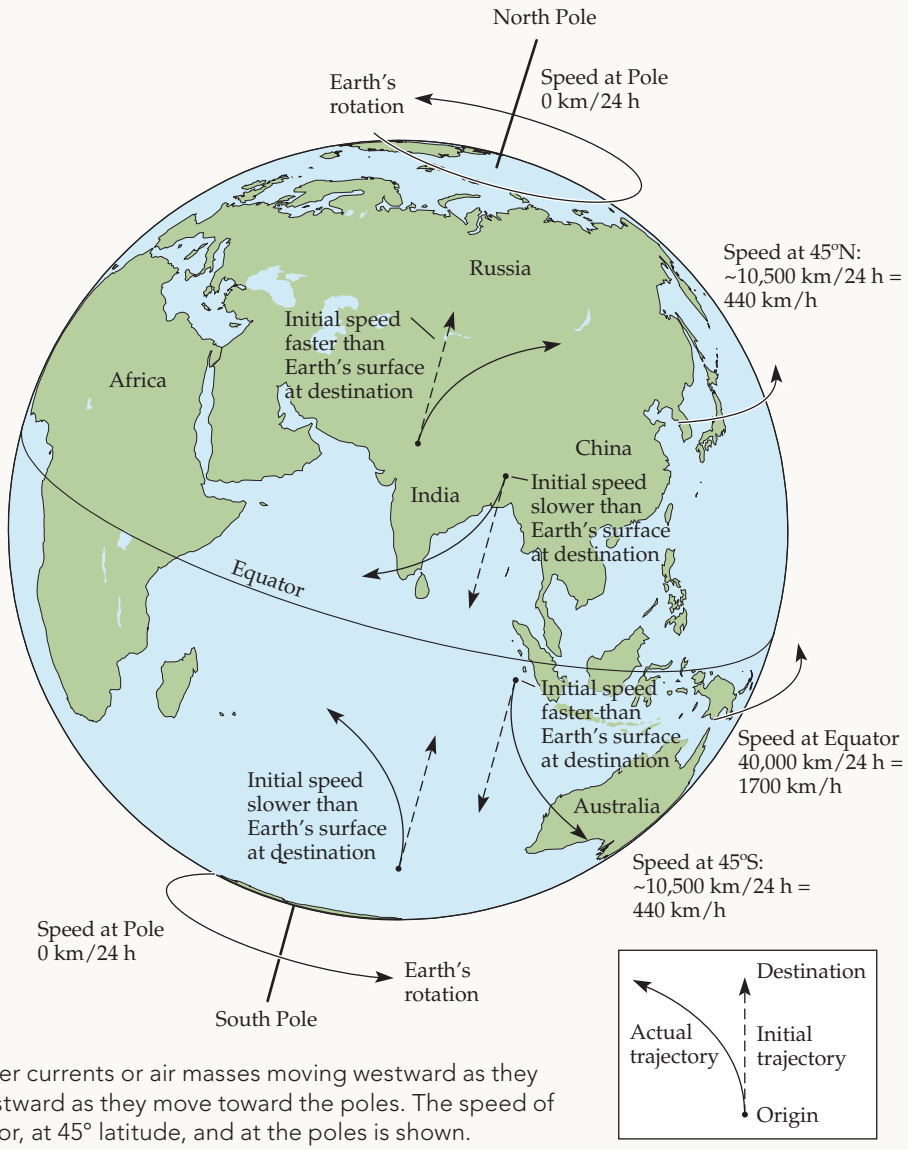
◀ **Figure 16.10** (A) The three dimensional global circulation of air, heat, and moisture is driven by incoming solar radiation over the tropics, causing rising air in the Hadley cells. The rising air cools, releasing rainfall, travels away from the tropics at the top of the troposphere high in the atmosphere, falls again at about 30 degrees N and S, warming and drying the ground below, then travels along Earth's surface back towards the tropics. The air from the two Hadley cells meets at the Intertropical Convergence Zone (the ITCZ). Notice that the part of the atmosphere closest to Earth's surface, the troposphere, is thickest and highest in the tropics and much shallower towards the poles. The Hadley cells drive the pattern of air movement of the Ferrel cells to the north or south, somewhat like interlocking gears. The Polar cells are driven by dense, falling cold air over the poles. The air traveling along the surface curves due to the Coriolis effect (see Box 16A) creating the westerlies north and south of 30 degrees and the trade winds from the equator to 30 degrees north and south. This diagram is shown at one of the equinoxes; at other times of the year the ITCZ follows the pattern of direct incoming solar radiation (see Figure 16.6). (B) The pattern of rising and falling air in Hadley, Ferrel, and Polar cells. The rising air of the two Hadley cells causes high precipitation over the tropics, travels across the boundary between the troposphere and the stratosphere (the tropopause) and falls back to Earth's surface, warming and drying, resulting in deserts at around 30 degrees N and S of the equator, joining again back at the ITCZ. The ITCZ and the Hadley, Ferrel, and Polar cells move northward in June and southward in December. The polar jet streams are fast-moving rivers of air high up in the atmosphere that move from west to east in the north and south separating arctic air from temperate regions, and whose movement influences climates and weather at higher latitudes. (C) General global patterns of wind circulation in July (above) and January (below), showing the movement if the ITCZ northward in July and southward in January. (C After F. K. Lutgens and E. J. Tarbuck. 2001. *The Atmosphere*, 8th ed. Prentice Hall: Upper Saddle River, NJ.)

Finally, once the air nears the ground, it flows along the surface toward the tropics, replacing the air rising there. The Coriolis effect (Box 16A) causes air to move westward relative to Earth's surface (appearing to veer to the right in the Northern Hemisphere and to the left in the Southern Hemisphere as it moves toward the topics). The **trade winds** are the result of this air moving toward the tropics and westward in both hemispheres, from northeast to southwest in the Northern Hemisphere and from southeast to northwest in the Southern Hemisphere.

## BOX 16A The Coriolis Effect

The **Coriolis effect** results in the general clockwise pattern of circulation in the oceans and atmosphere in the Northern Hemisphere but counterclockwise rotation in the Southern Hemisphere. Understanding the Coriolis effect is fundamental to understanding many patterns in global climate and weather, including the reasons for the general patterns of atmospheric and oceanic circulation. The explanation starts with the Earth rotating toward the east. (That's why the sun appears to rise in the east.)

One way of summarizing the Coriolis effect is this: Air masses or water currents moving straight toward the equator appear to curve westward. Those moving straight toward the poles appear to curve eastward. This apparent curve relative to the surface of the Earth occurs because, odd as it sounds at first, the surface of the Earth does not rotate eastward at the same speed everywhere. If you stand on the equator, in one day you will have traveled much farther going all around the planet than if you were standing on the Arctic Circle—this must be so, if you are to arrive back around the Earth in 24 hours, because there is a lot farther to go at the equator than at the Arctic Circle. An object, or an air mass, moving toward the equator from somewhere in the temperate zone in either the Northern or Southern Hemisphere initially has the same rotational speed as the ground where it originated; as Earth rotates eastward on its axis, the object moves along with the Earth at the same speed as the ground surface. As this object moves toward the equator, it retains its momentum moving eastward (in the direction of Earth's rotation), but as it gets closer to the equator, it is moving at a speed relatively *slower than* the surface beneath it. So, it lags relative to the surface and appears to curve to the west, because the ground has moved faster eastward than it has. As an object moves in the opposite direction, away from the equator toward one of the poles, its momentum carries it eastward faster than the ground beneath it, and it appears to curve eastward




The Coriolis effect results in water currents or air masses moving westward as they move toward the equator, or eastward as they move toward the poles. The speed of the Earth's rotation at the equator, at 45° latitude, and at the poles is shown.

The air rushing toward the tropics and westward along Earth's surface from the northern and southern Hadley cells meets up at the intertropical convergence zone (ITCZ), whose location depends on the time of year (Figure 16.10C). At the equinoxes in September and March, the ITCZ is near the equator, with the sun directly overhead, propelling the rising air of the Hadley cells. At the June solstice the ITCZ has moved to approximately 23.5°N, the Tropic of Cancer, where the sun is now directly overhead, and at the December solstice the ITCZ has moved to approximately 23.5°S, the Tropic of Capricorn, with the sun directly overhead (see Figure 16.10C). The Hadley cells move north of the equator in Northern Hemisphere summer, and south of the equator in Southern Hemisphere summer (they are shown in Figure 16.10A, B at one of the equinoxes). Rainfall follows the ITCZ, causing rainy seasons where it is situated and dry seasons where it has moved away, as it moves north and south (see Chapter 18).

The Hadley cells propel another giant pair of air cells, the weaker Ferrel cells (see Figure 16.10A, B), much like interlocking gears. The Ferrel cells are believed to be driven by the other two cells between which they are sandwiched. These air masses fall with the Hadley cells at around 30°N and 30°S, move along Earth's surface toward the poles and eastward relative to the surface, and rise at around 60° latitude. This pattern creates the prevailing **westerlies** (winds moving from west to east) in North America, western Europe, East Asia, and from 30° to 60° in the Southern Hemisphere, south of most land masses. People living on the East Coast of the U.S. often get their weather from the Midwest, for instance, as these winds move from northwest to southeast. Early travelers in sailing ships going between Europe and North America took advantage of these patterns, allowing the trade winds to carry their ships southwestward across the Atlantic on their journey out, reaching the Caribbean, and returning to Europe with the westerlies farther north. These large-scale patterns of moving air masses have critical influences on world climates because they convey moisture and energy over great distances, influencing conditions in the lands they cross. It is difficult to draw these in a flat plane; picture them more as giant tubes spanning Earth's surface rather than as circles, with the sides reaching up into the higher atmosphere and then down to Earth's surface.

The third type of giant atmospheric cells are the polar cells. Dense, cold air sinks at the poles, travels along the surface moving west, and rises at about 60° latitude; polar regions also tend to be very dry. Between the various cells, high in the atmosphere and running perpendicular to the rising and falling air masses are fast-moving rivers of air called the **jet streams**. The polar jet streams are located between the polar and Ferrel cells in the Northern and Southern Hemispheres and flow from west to east at about 10 km in altitude. These roaring rivers of air define the polar front where cold air moving toward the pole meets warm air moving toward the tropics, and they are where frequent massive exchanges

as it outruns the ground surface beneath it. You can try this for yourself: take a bowl (an approximate half sphere) and use a straight edge to try to draw a straight line up from the base (the South Pole) to the rim (the equator) while rotating the bowl eastward (to your right); the line will curve "westward" to the left. The opposite will happen if you start your line at the rim and head toward the base.

The conventional way the Coriolis effect is explained requires remembering more rules: objects moving toward or away from the equator in the Northern Hemisphere appear to be deflected toward their right, and in the Southern Hemisphere toward their left. Winds beginning in Paris and headed due north would move to the northeast relative to the Earth's surface. Winds headed due south would move to the southwest relative to the Earth's surface. The reverse would be true in the Southern Hemisphere: an air mass headed due north from Sydney would move to the northwest, and air headed due south would move over the ocean to the southeast. Moving air masses or ocean currents in the Southern Hemisphere moving north or south appear to be deflected to their left, in the Northern Hemisphere, they appear to be deflected to their right. The Earth rotates counterclockwise if you are looking down at it from the North Pole, or clockwise if you are looking up at it from the South Pole (but it moves eastward regardless of how you are looking at it, even if you are standing on your head or looking at the Earth from space).

The Coriolis effect means that air masses being moved by the Hadley cells from 30° latitude toward the equator are deflected westward—toward the southwest in the Northern Hemisphere and toward the northwest in the Southern Hemisphere. These are the trade winds. Air masses moving from 30° latitude toward the poles by the Ferrel cells appear to be deflected eastward—toward the northeast in the Northern Hemisphere and toward the southeast in the Southern Hemisphere. These are the westerlies. In turn, these wind patterns are the major reason that oceanic currents move clockwise in the Northern Hemisphere and counterclockwise in the Southern Hemisphere.

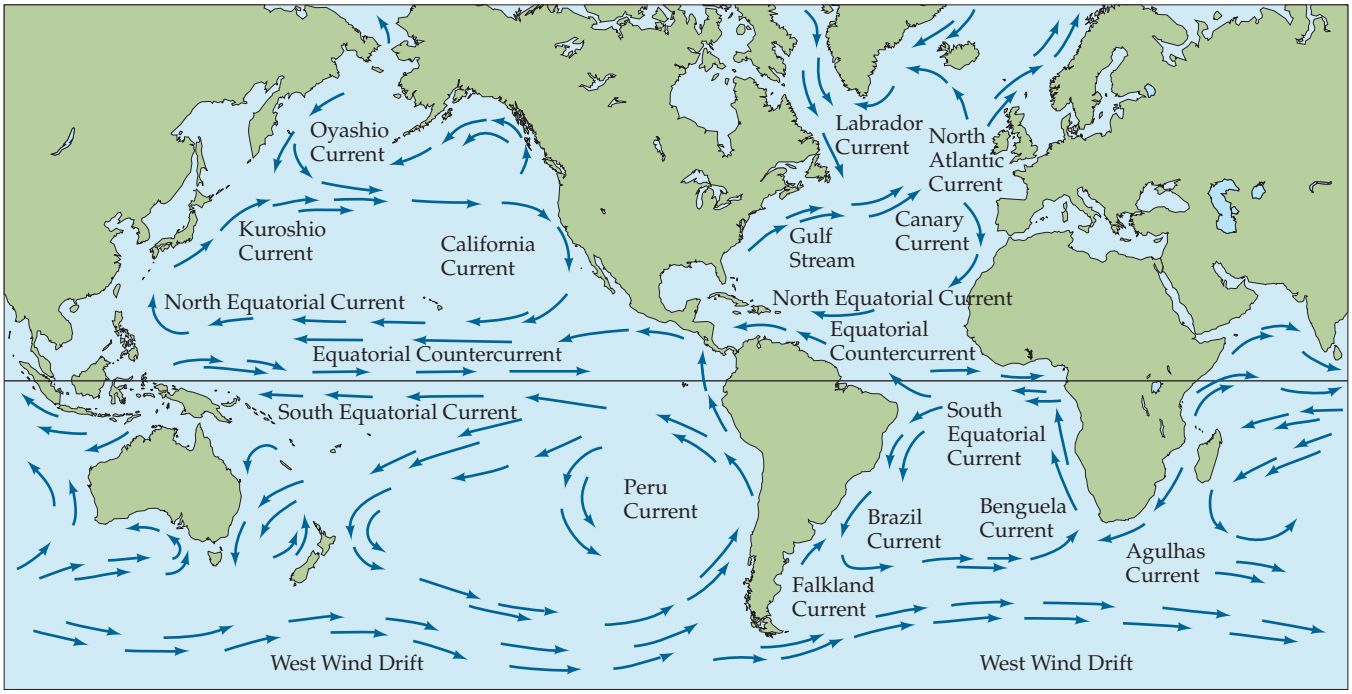
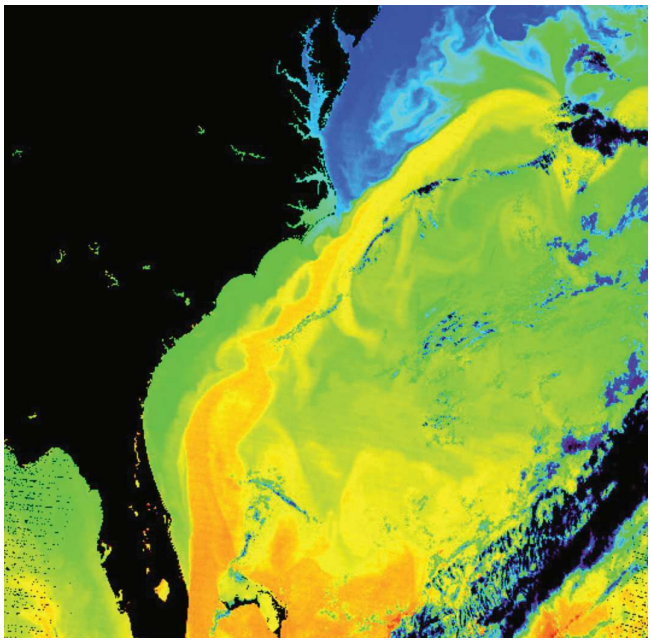



**Figure 16.11** The world's major oceanic currents.

of energy occur in the form of storms at mid-latitudes. The locations of the polar jet streams, and the associated polar fronts, change with the seasons as well as shifting within seasons and are a primary factor determining short-term weather events like severe snowstorms at mid-latitudes. The subtropical jet stream between the Hadley and the Ferrel cells is weaker.

The movement of air masses at the Earth's surface is a major factor in determining surface oceanic currents (Figure 16.11), with the winds pushing surface ocean water, which in turn plays a major role in terrestrial climates. Because of the high specific heat of water, these currents are even more effective conveyors of heat energy than the moving air masses. They bring water warmed or cooled far away to the coasts of all of the continents north of Antarctica and to most islands, adding or removing heat energy wherever they travel, and thus have a major influence on local temperature and moisture regimes. These ocean currents act like rivers of water distinct from surrounding ocean water, with their own temperatures, salinities, and nutrient concentrations. Like the winds pushing them, these currents move westward when moving toward the equator and eastward when moving away from the equator (the Coriolis effect again).

Although air masses can rise above Earth's surface, ocean currents cannot, so when they are blocked by continents, ocean currents are forced to go either north or south. This creates **gyres**, enormous circles of water that move clockwise in the Northern Hemisphere and counterclockwise in the Southern Hemisphere in temperate latitudes, as water is directed westward when moving toward the equator and eastward when moving toward the poles. Major gyres exist in the North and South Atlantic, North and South Pacific, and Indian Oceans. Smaller gyres exist toward the poles, with water likewise moving northeast toward the poles and southwest back toward the temperate regions. The gyres carry heat energy to and away from the areas they traverse. The Gulf Stream in the North Atlantic, for example, brings tropical waters northeastward (Figure 16.12), warming the island of Bermuda and resulting in a far warmer climate in Ireland, Great Britain, and other parts of northern and western Europe than they would have based on their latitudes alone. The Humboldt or Peru Current brings frigid Antarctic waters north along the western coast of South America, causing a cooling of the air masses above them. When this cool air meets the warm land surfaces on shore, it heats up, absorbing moisture and drying the land. The California current brings cold water southward from Alaska along the west coast of North America and winds push the water westward away from land, drawing cold, deep waters to the surface (see Figure 16.11). The cold water offshore creates foggy conditions in summer when it meets the cool land surface in northern California, and very dry, sunny summers when it reaches the much warmer land surface in southern California (which also experiences falling air from the northern Hadley cell). We consider other effects of these currents when we look at the El Niño Southern Oscillation and related phenomena later in this chapter.



Courtesy of MODIS Ocean Group NASA/GSFC SST product by U. Miami

**Figure 16.12** NASA satellite image of the warm waters of the Gulf Stream running up the eastern seaboard of the United States and eastward into the Atlantic Ocean. The Gulf Stream shows up as a winding rope of orange and yellow (indicating warm waters) against the cooler green and blue waters.

Surface oceanic currents are not the only conveyor belts of water and energy affecting regional and continental climates. Huge, three-dimensional currents in the world's oceans slowly move water, nutrients, and heat over great distances. **Thermohaline circulation** (THC) is oceanic circulation driven by differences in the density of water, and it moves water both along the surface of the oceans across deep stretches in the ocean, and between the surface and ocean depths (Broecker 1991; Ramstorf 2000). Differences in the density of water are caused by two factors: cold water is denser than warm water, and salty water is denser than fresh water. Cold, dense water sinks deep to the ocean bottom in specific places, flows along at mid-depth, and then rises again elsewhere (Figure 16.13). In the Atlantic Ocean, as the surface water sinks in the Arctic and moves southward at depth, surface water currents are driven toward those areas to replace the sinking water. These surface portions of the THC contribute to the North Atlantic Current and add to the wind-driven Gulf Stream. The climatic effects and long-term history of the THC are the subject of research and controversy, and its future is more uncertain still. However, it is clear that it has been very strong at some periods in the history of the Earth and far

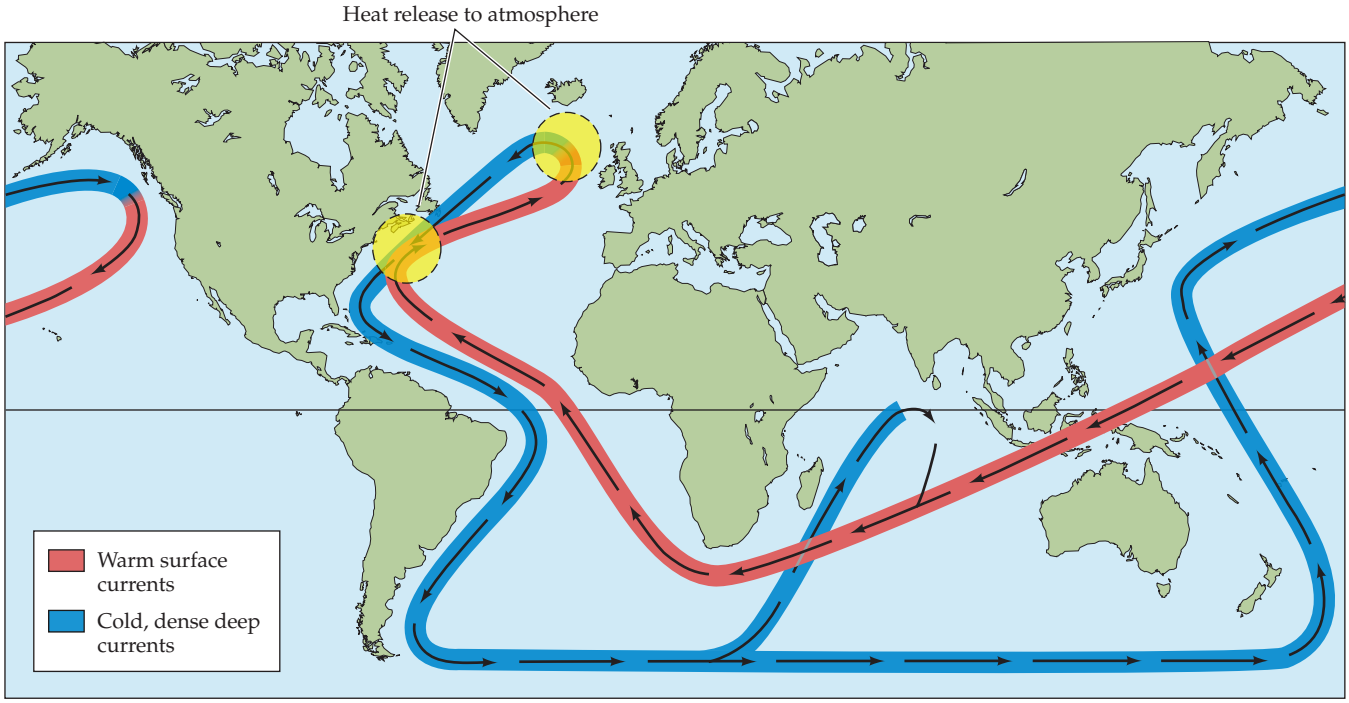



**Figure 16.13** Global thermohaline circulation. Surface currents are illustrated in red, and circulation of deep oceanic waters in blue. The two major spots where cold, dense water sinks to great depths are shown as yellow circles. (After NASA Goddard Space Flight Center (GSFC). Greenbelt, MD; https://www.nasa.gov/centers/goddard/images/content/95552main_0625oceanbehavior_conveyor.jpg)

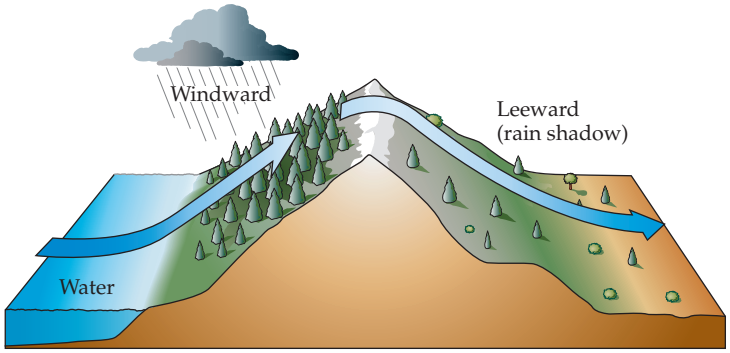



**Figure 16.14** Rain shadows are formed when warm, moist air is forced upward and over a mountain or mountain range, depositing high levels of precipitation on the windward side. The rain shadow is created where dry air descends on the leeward side of the mountain.

weaker at other times, perhaps both responding to and contributing to large changes in regional climates.

## Continental-scale movement of air and water explain regional differences in snow and rain

Precipitation patterns at continental and regional/landscape extents are strongly affected by several factors in addition to those just described, particularly the distribution of mountain ranges, proximity to large bodies of water, and differences in temperature between the land and nearby bodies of water. When air moving over the Earth's surface meets a mountain range, it has nowhere to go but up. This rising air cools adiabatically and releases precipitation. The windward sides of mountains are generally areas with high precipitation. Eventually the air is pushed over the tops of the mountains and falls down the other side. The falling air warms again and (like the descending air from the Hadley cells at 30° latitude) picks up and holds moisture. Thus, the leeward sides of mountains tend to be areas of low precipitation. These dry areas are known as **rain shadows** (Figure 16.14). This pushing of air over mountains also contributes to mountaintops being windy places.

Rain shadows occur in many parts of the world and at various scales. One of the driest areas of the world, the Gobi desert of central Asia, is in the rain shadow of the Himalayas. The pampas, the great mid-latitude grasslands of Argentina, are in the rain shadow east of the Andes caused by the prevailing westerlies. Farther north in South America, air masses moving to the northwest across the Atlantic, laden with moisture, cross the Amazon basin and are forced to rise when they reach the Andes at the eastern edge of the mountains. This pattern results in very high rainfall on the eastern side of the Andes at latitudes north of 30°S and contributes to the extreme aridity in the Atacama Desert on the western side of the Andes. The Atacama Desert is further affected by the falling air of the southern Hadley cell and the drying effect of the offshore Humboldt current as described earlier.

Even a single mountain can cast a rain shadow. For example, Mauna Kea on the island of Hawaii rises 4200 m above sea level. The eastern face of the mountain receives as much as 750 cm of rain a year. Just 50 km away on the western face, less than 50 cm of rain falls in a year (Figure 16.15). This pattern is repeated

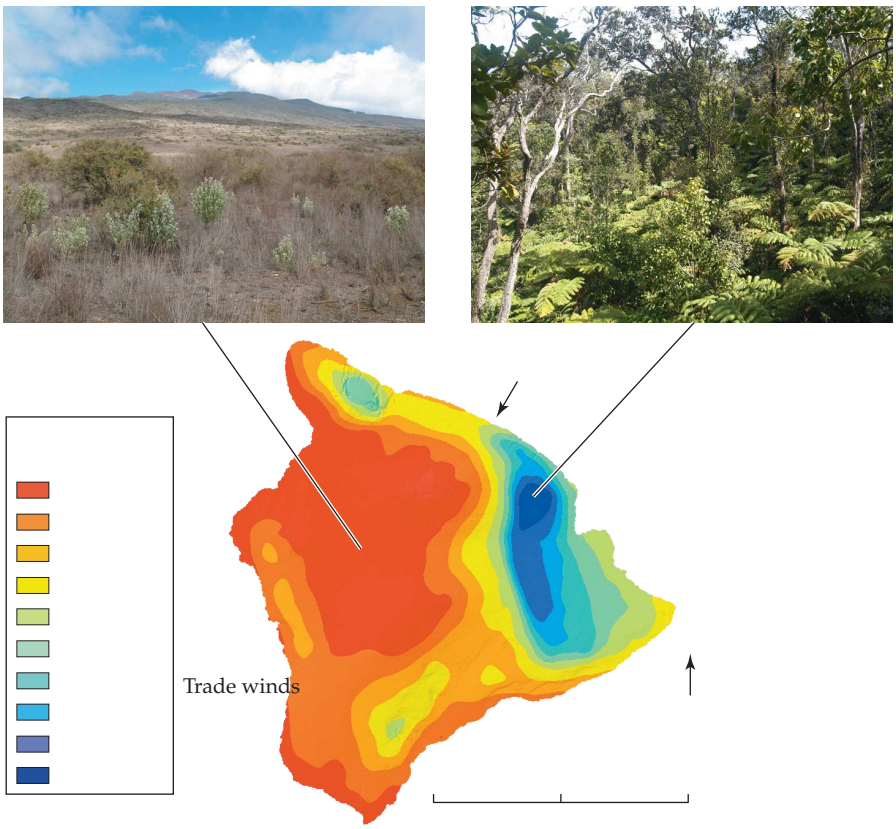



**Figure 16.15** Precipitation patterns on the on the island of Hawaii (the Big Island). Contours show millimeters of rain per year. The prevailing winds are the trade winds, from the northeast. Mountains on all of

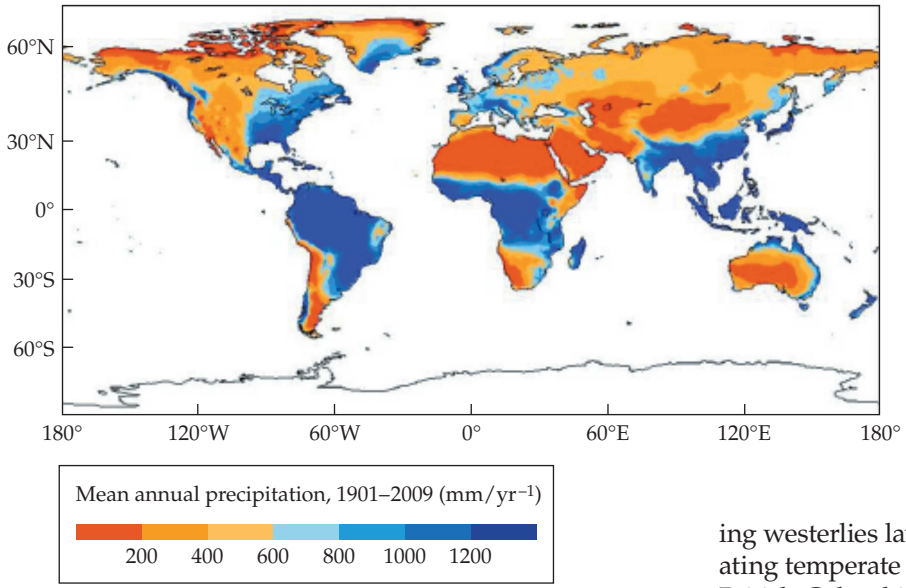



**Figure 16.16** Global patterns of mean annual precipitation from 1901–2009 in mm/year, ranging from under 200 to over 1200 mm/year. (After J. Huang and F. Ji. 2012. *Atmos Chem Phys* 12: 5391–5398.)

on islands in the Caribbean with sufficiently high mountains; for example, the island of Puerto Rico has a rainforest on one side of its mountains and semiarid conditions on the other side.

Land masses can be affected by their proximity to oceans or other large bodies of water in several ways. Air passing over the water can pick up moisture and deposit it on land as rain or snow, or the land can tend to lose moisture to the body of water, depending on relative air and water temperatures. One example of the first phenomenon is the "lake effect" around the Great Lakes of North America: winter storms tend to deposit much more snow as one moves eastward through this region, because the moving air masses gain moisture from the lakes. That is why Buffalo, New York (at the eastern edge of Lake Erie), is well known locally for its heavy winter snow, while nearby Toronto, Ontario (on the north shore of Lake Ontario, and thus out of the way of most storms off the lakes), receives much less snow.

At a larger spatial scale, when air masses cooled by cold oceanic currents travel to lower latitudes and meet warm land masses, the air warms as it passes over the land, absorbing moisture and preventing precipitation. If this air veers offshore, it can even tend to lose moisture to the ocean, compounding the drying effect on the land. This phenomenon occurs to some degree on the western coasts of all continents, all of which have cool currents offshore. As we have seen, the coastal Atacama Desert in Chile is affected by the cold Humboldt Current, contributing to its exceptional aridity. Some areas in that desert have been rain-free for stretches of 10 years, although just offshore there are cool, foggy air masses. Coastal southern California also has limited rainfall (though its vegetation is chaparral rather than desert) because the warm land surface next to the cold California Current prevents oceanic air masses from delivering rain during much of the year. In winter, the land cools enough to allow rainfall there some of the time. Farther north, however, the land is much cooler, and the cool, moist prevailing westerlies lavish large amounts of precipitation, creating temperate rainforests in Oregon, Washington, and British Columbia (see Chapter 18). Even there, much of the rain falls in winter, when the ground is coldest (discussed further in the next section).

Air circulation patterns and geographic features interact to produce continental-scale patterns of precipitation (Figure 16.16). In North America, for example, the major mountain ranges (the Appalachians, Rockies, and Sierra Nevada) run in a north-south direction. The primary direction of airflow at mid-latitudes, as we have just seen, is west to east. These factors combine to produce distinct patterns of precipitation across the North American continent.

Along the West Coast, from northern California to southern Alaska, there is a region of high precipitation where moisture-laden air coming in from the Pacific runs into mountain ranges (and the ground is cool enough to allow precipitation, at least in the cooler months). To the east of these mountains, the climate is semiarid or desert. The intermountain or Great Basin desert of Nevada and western Utah, for example, is in the rain shadow of California's Sierra Nevada (Figure 16.17). A second area of relatively high precipitation occurs on the western side of the Rocky Mountains. To the east of the Rockies there is another rain shadow. Precipitation levels gradually increase toward the Atlantic seaboard. The increased moisture in the Midwest and east comes primarily from moist air moving north from the Gulf of Mexico (Figure 16.18). The Appalachian Mountains are not high enough to create a rain shadow.

## Seasonal variation in precipitation is an important component of climate

Variation in precipitation, as in temperature, occurs on several different time scales. On a scale of days, precipitation varies as weather systems move across the globe. These weather systems are caused by the movement of high and low pressure areas. Precipitation usually

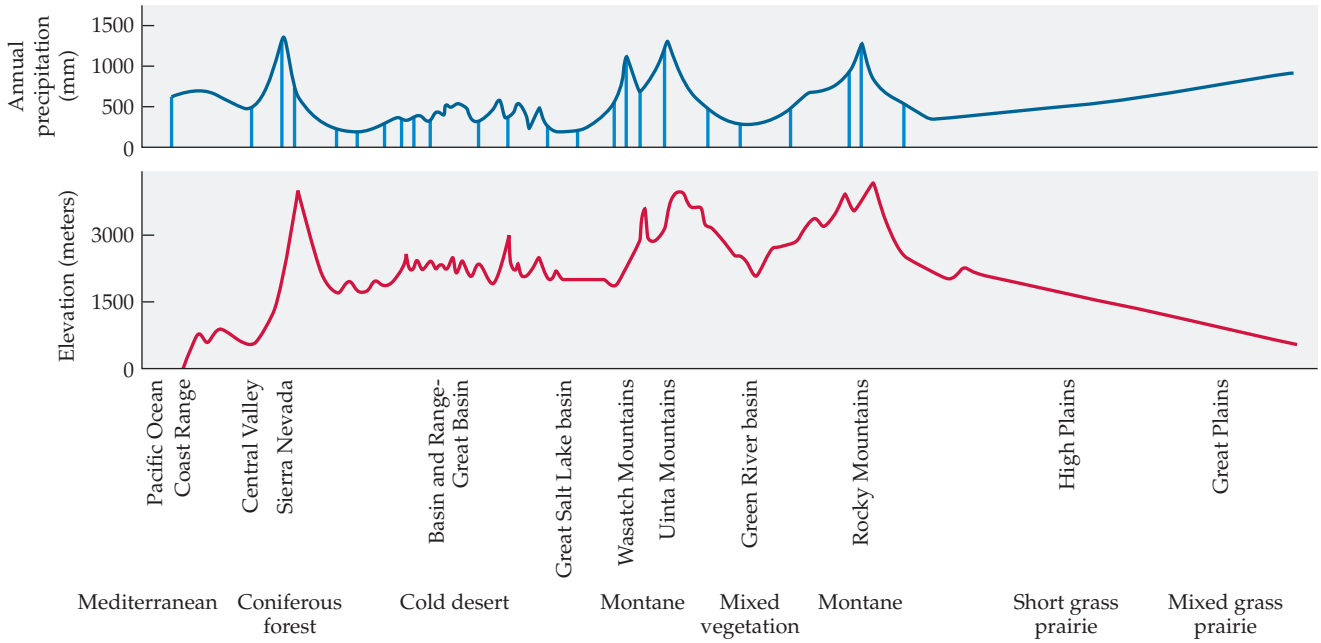



**Figure 16.17** Transect showing the patterns of elevation in m and mean annual precipitation in mm/year across western North America at 40°N latitude (from the coast of northern California, through Nevada and Utah, and ending in mid-continent on the Great Plains). (After R. W. Bailey. 1941. *Climate and Man. Agricultural Yearbook*. USDA: Washington, DC; additional data R. D. Pieper. 2005. In *Grasslands of the World*. J.M. Suttie et al., [Eds.] FAO United Nations: Rome; *Annual Precipitation 1961–1990*. PRISM OSU/WRCC. Western Regional Climate Center, Desert Research Institute: Nevada.)

occurs at the leading edge of high-pressure systems. On average, these systems take several days to move through an area, creating alternating wet/dry cycles.

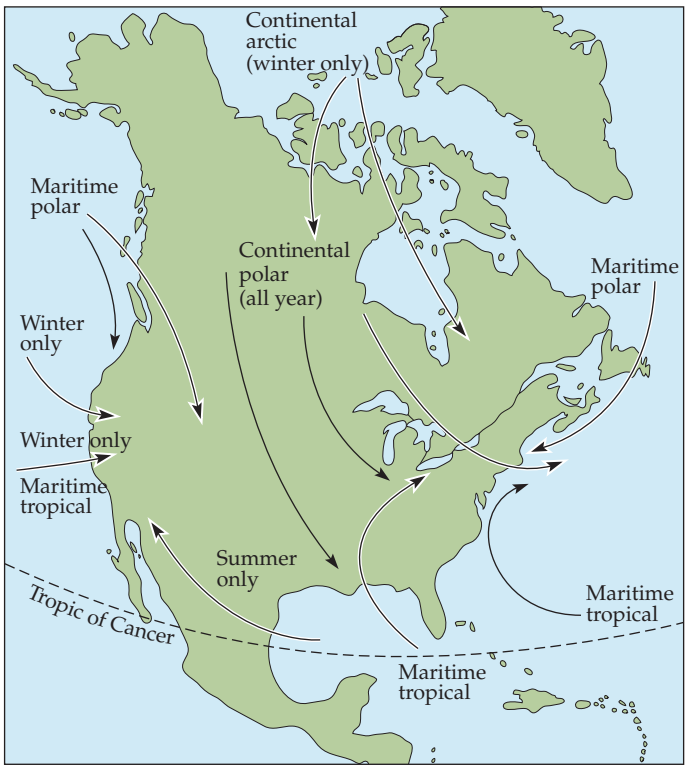
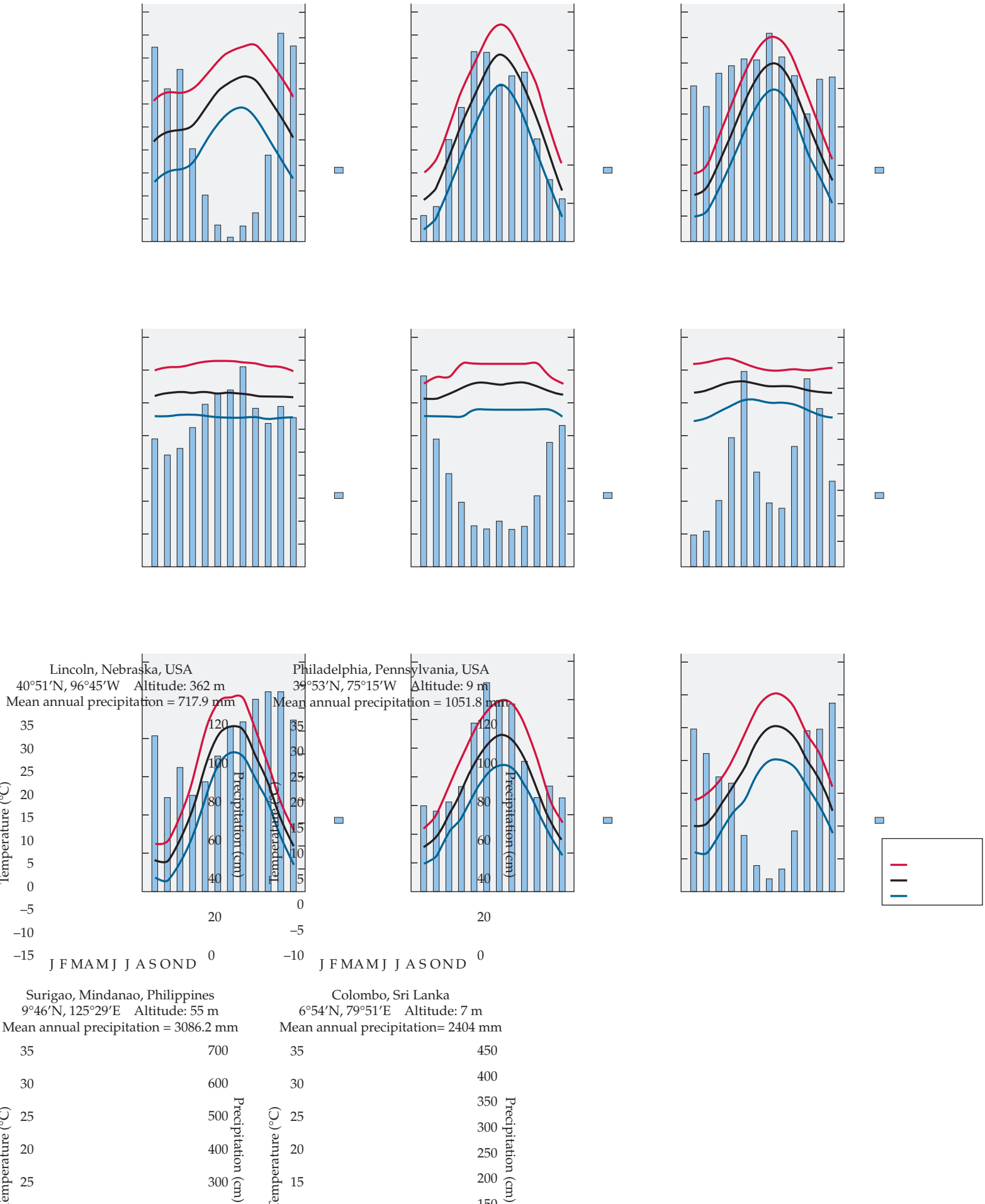


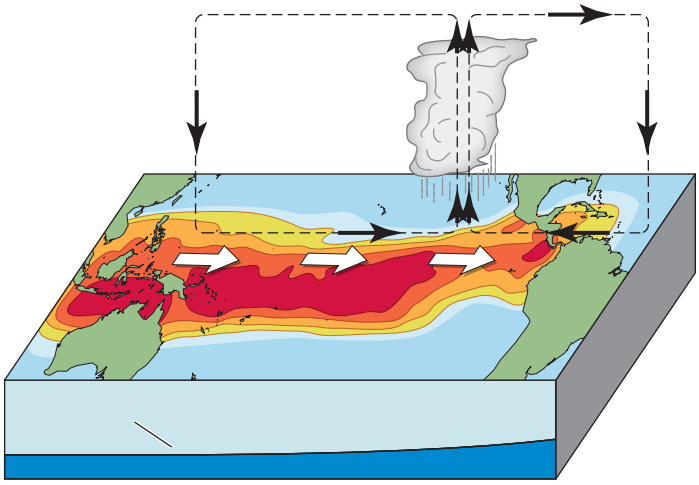

**Figure 16.18** Principal air masses and their direction of movement at different seasons in North America. (After G. R. Rumney. 1968. *Climatology and the World's Climates*. Macmillan: New York.)

Weekly and seasonal variation in larger-scale phenomena, such as the position of the jet stream, causes variation on larger spatial and temporal scales.

The most important scales of variation in precipitation are seasonal and yearly. Annual variation in temperature helps to drive variation in precipitation. Several different kinds of annual patterns of precipitation exist. Along western continental margins there are distinct seasonal patterns, with rainy winters and dry summers (Figure 16.19; Eureka, California and Palermo, Sicily). Where temperatures are warm, this pattern is often referred to as a **mediterranean** climate, as it is typical of areas around the Mediterranean Sea. In North America, mediterranean climates are found in coastal southern and central California; other regions that exhibit this pattern include Western Australia, southern Chile, and the southwestern tip of Africa. **Continental** climates are found in the interior of large continents at mid-latitudes, as in the North American Midwest and central Russia. Precipitation here is also highly seasonal, with the greatest amount in the summer, although there is generally some precipitation year round (see Figure 16.19; Lincoln, Nebraska and Munich, Germany).

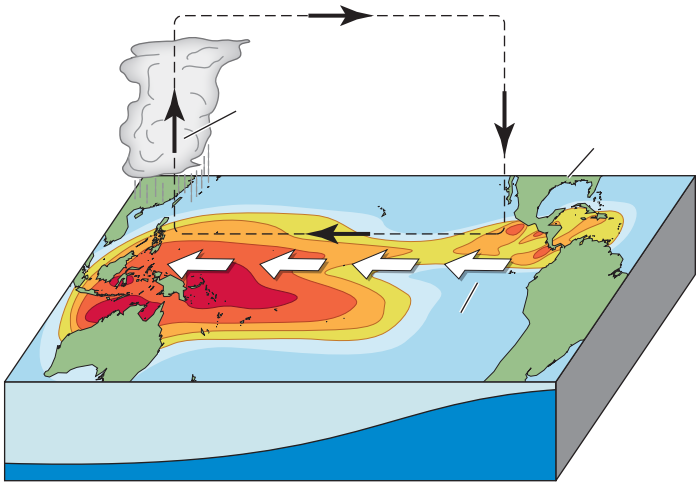

Lincoln, Nebraska, USA
40°51′N, 96°45′W Altitude: 362 m
Mean annual precipitation = 717.9 mm
Philadelphia, Pennsylvania, USA
39°53′N, 75°15′W Altitude: 9 m
Mean annual precipitation = 1051.8 mm
Temperature (°C)
Precipitation (cm)
J F M A M J J A S O N D
Surigao, Mindanao, Philippines
9°46′N, 125°29′E Altitude: 55 m
Mean annual precipitation = 3086.2 mm
Colombo, Sri Lanka
6°54′N, 79°51′E Altitude: 7 m
Mean annual precipitation= 2404 mm

**Maritime** climates are often found along the eastern edges of continents. Here precipitation is more uniform throughout the year, although it comes as rain in the summer and snow in the winter (see Figure 16.19, Philadelpia, Pennsylvania, and to some extent, Göteborg, Sweden). In equatorial regions, temperatures are high throughout the year. Precipitation in the tropics may show one of three patterns: relatively constant precipitation throughout the year, a distinct wet season and dry season, or two wet/dry cycles in a year, depending on the movement of the intertropical convergence zone (see Figure 16.19; El Caraño, Colombia in the Amazon Basin, Surigao, Mindanao, Philippines, an island east of the Asian continent, and Colombo, Sri Lanka, in South Asia; see also Figure 18.14).



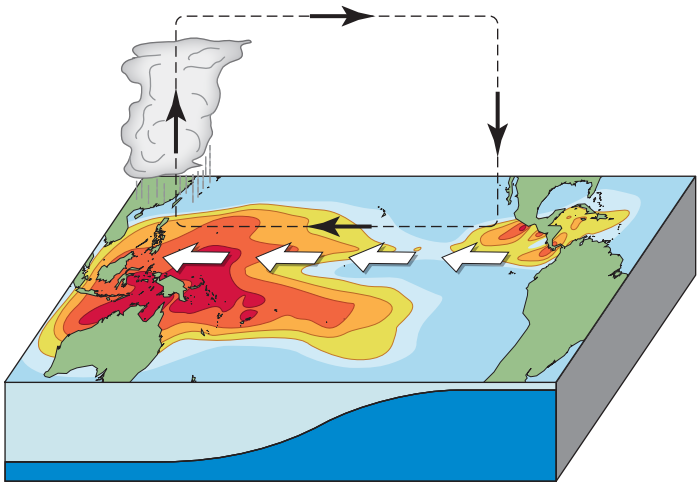
## The El Niño Southern Oscillation affects rainfall at large spatial scales and intermediate time scales

Besides these weekly and yearly cycles of precipitation, other cycles also exist. Of these, the most well known is a 3-to-7-year cycle called the **El Niño Southern Oscillation** (ENSO). Narrowly defined, the ENSO determines weather patterns in a band from 20°N to 20°S across the Pacific Ocean. However, the ENSO is closely interconnected with other global atmospheric and oceanic effects. Together, ENSO events and their associated effects drive some of the most dramatic global fluctuations in weather in some parts of the world, although the strength of the events, and the magnitude of their effects, can vary substantially.



ENSO effects occur over a very large geographic scale, influencing temperatures and rainfall from Africa to Australia, across the Pacific, and all the way across North and South America, affecting by some estimates about 75% of the globe (among the few places that are not much affected are northern Europe and northern Asia, although parts of southern Asia can be strongly affected). Remarkably, although climate scientists had been investigating these cycles for over 100 years, they did not began to fully integrate their findings into the modern understanding of the ENSO until the record El Niño year of 1982–1983. The study of the ENSO and the role of the Pacific Ocean in this cycle has transformed our understanding of world weather patterns. Climatologists are still refining their understanding of all of its effects, which may also be changing due to global warming (discussed below).

Niño conditions



El Niño (Figure 16.20A; literally "the boy") is named for the Spanish term for the infant Jesus

◀ **Figure 16.20** Pacific Ocean surface temperatures, major ocean and air currents, areas of rising and falling air and precipitation patterns (A) during an El Niño year (B) during a neutral or normal year, and (C) during a La Niña year. The thermocline separating surface and deeper oceanic waters flattens during an El Niño, but rises to the east in normal and La Niña years with upwelling at the western coast of South America. Events are described in more detail in the text. Reds indicate warmest surface water temperatures, oranges and yellows are somewhat cooler, and blues are cool surface waters. Terrestrial areas are shown in green, with South America to the right and North America above it; Australia is to the left of the figures. (From NOAA/National Weather Service/Climate Prediction Center. 2005. El Niño- and La Niña-related winds, equatorial walker circulation, and subsurface ocean structure. Accessed via: https://www.gfdl.noaa.gov/el-nino-media/.)

because its effects begin to be observed off the coast of Peru around Christmas. El Niño is only one-half of the ENSO; it alternates with La Niña ("the girl") years, during which conditions are reversed.

The heart of El Niño is a cycle of tropical Pacific oceanic and atmospheric conditions. When El Niño is not present, strong trade winds in both hemispheres blow west across the Pacific, pushing oceanic currents along in the same direction (a "normal" year; Figure 16.20B). The water being pushed away from the western coasts of North and South America is replaced with very cold, nutrient-rich water dragged up from the ocean bottom. This upwelling of cold bottom water in the eastern Pacific supports extremely productive and diverse marine life as well as affecting global climates. The trade winds push warm surface waters into the far western Pacific north of Australia. Warm water results in increased evaporation, bringing rainfall to the islands of the western Pacific and to eastern Australia and Asia.

Preceding an El Niño, the trade winds slacken, and warm surface water flows east, resulting in higher sea levels in the eastern Pacific along the coast of South America and lower sea levels in the western Pacific along the eastern coasts of Australia and southeast Asia. This movement of warm water from west to east forces the thermocline to flatten out and sink deeper in the ocean, triggering the El Niño phase of ENSO. When an El Niño year commences, surface temperatures in the central Pacific Ocean become as much as 5°C warmer than normal (see Figure 16.20A). As upwelling in the eastern Pacific ceases, there are huge negative repercussions for fish populations, marine mammals, seabirds, and other organisms that are ultimately dependent on the nutrient-rich bottom water. This cold water never makes it to the surface, warm surface water now extends across the tropical Pacific from east to west, and there are drastic decreases in rainfall in the western and central Pacific. As an example, the very severe El Niño year of 1997–1998 (Figure 16.21) resulted in extreme drought in Australia, Indonesia, and southern Africa.

During an El Niño period, the atmospheric pressure in the eastern Pacific is low, while it is high in the western Pacific. These differences in air pressure force surface air to move from the areas of higher pressure in the west to lower pressure in the east, eastward across the Pacific Ocean. Across the tropical Pacific Ocean, the usually strong east-to-west wind is therefore suppressed, and the westward movement of oceanic currents is also suppressed by the air moving east. This atmospheric pattern reverses during a La Niña year. The reversal in atmospheric pressure between the east and

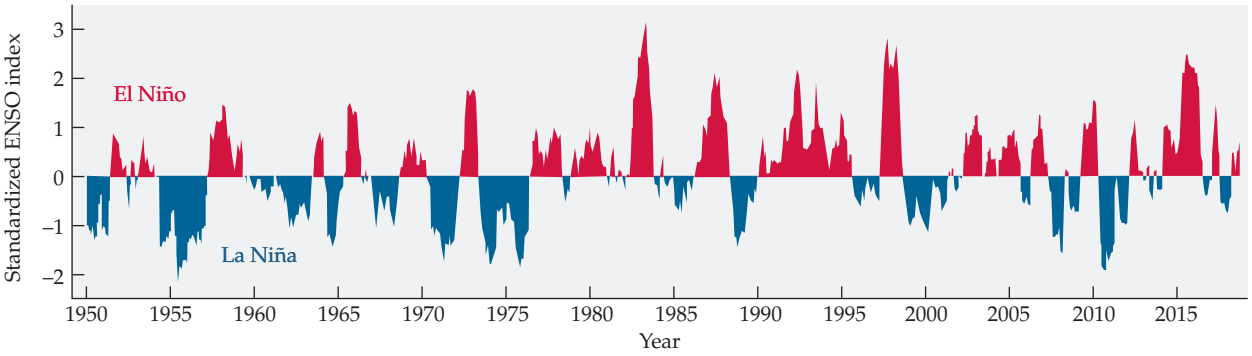



**Figure 16.21** Historical record of intensity of El Niño (above the midline of the graph) and La Niña (below the midline) events from 1950 to 2018, based on a standardized multivariate index that includes ocean surface temperatures, surface air temperatures, wind characteristics, and other variables. (After NOAA/ESRL/PSD. 2018. Accessed via: https://www.esrl.noaa.gov/psd/enso/dashboard.lanina.html.)

west Pacific is called the Southern Oscillation. The link between the oceanic and atmospheric cycles is the combined El Niño Southern Oscillation, the ENSO cycle. In an El Niño, warm, moist air rises above the eastern Pacific, causing heavy rainfall in ordinarily dry regions such northern Chile and northward to the coasts of Peru and Ecuador and to southern California. This rain can result in severe flooding and landslides in these areas. Finally, a complex change in high-altitude winds 12 km over the Pacific forces a jet of air to move eastward, past Central America, and then across the Atlantic to Africa, depressing hurricane activity in the Atlantic and causing other weather disturbances. These changes in upper-level atmospheric winds also force changes in the jet stream, spreading El Niño's effects to other parts of the globe. In El Niño years, the jet stream is forced southward across North America, bringing flooding to the southeastern U.S. but warm, dry weather to the U.S. Pacific Northwest and southwestern Canada and drought and wildfires to the U.S. Southwest.

Eventually, these conditions reverse. During La Niña years (Figure 16.20C), in the opposite part of the ENSO, strong trade winds push warm surface waters into the far western Pacific north of Australia. Evaporation is heightened in the western Pacific, further strengthening the trade winds. The movement of surface waters westward causes a strong upwelling of cool, deep waters in the eastern Pacific along the coast of South America. The Northern Pacific equatorial countercurrent, which flows counter to the prevailing winds, weakens. The effects on weather patterns are essentially the opposite of those in an El Niño year. The cold surface waters in the eastern Pacific force the jet stream northward, causing drought conditions across the southern U.S. and northern Mexico and heavy rains in the Pacific Northwest. At some point, the positive feedback that maintains this system reverses, and normal conditions or an El Niño return again.

The ENSO is a succession of alternating events. We do not completely understand what triggers the changes between El Niño and La Niña conditions, but the ENSO clearly depends on a complex interplay between oceanic currents and the atmosphere. Ocean waters are always exchanging energy (as heat) with the atmosphere, although they hold far more heat and move heat around the globe more slowly than do atmospheric winds. This exchange of heat energy with the ocean contributes to the movement of winds, and winds drive surface oceanic currents.

Globally, the El Niño of 1982–1983 caused droughts in Africa, Australia, and South America. Coral reefs off Costa Rica, Panama, Colombia, and the Galápagos Islands suffered drastic losses (50%–97% die-offs) because of the warmer water and loss of nutrients. In that same event, 85% of the seabirds on the Peruvian coast died or abandoned their nests; the same fate befell almost all of the 17 million seabirds living on tiny Christmas Island in the Pacific.

The El Niño of 1997–1998 was even more extreme, with the resulting drought exacerbating forest fires in Indonesia and Malaysia and harming the already threatened rainforests there. The rainforests of the Amazon basin also suffered from extreme drought. In many places, the unusual severity of this particular El Niño plus the other problems facing tropical forests (such as logging and fragmentation) combined to cause much greater damage than any of these factors would have by themselves (Wuethrich 2000). Other effects of that El Niño included drought in Australia and southern Africa, while severe flooding occurred in southern California and in the coastal deserts of northern Chile and Ecuador. The most recent extreme El Niño occurred from 2014 to 2016. During that event, the East Coast of the U.S. experienced some of its snowiest winters ever, and in 2014 Singapore experienced its driest February ever recorded, with temperatures as high as 35°C.

Short-term oscillations such as the ENSO are embedded in larger, longer-term phenomena. Over the past 5000 years, there may have been periods of more frequent El Niño events from 4800 to 3600 years before the present (BP), around 1000 BP, and after 500 BP. Long-term climate change, both natural and human induced, can exacerbate or dampen ENSO and other intermediate-scale climate and weather events.

## Temperature and rainfall predictability affect plant ecology and evolution

In this chapter, we have described patterns of variation in temperature and precipitation. These patterns are generally regular, with some predictability. Random effects act to introduce unpredictability into these patterns. The degree of unpredictability is one determinant of the types of plant communities found at different locations (Figure 16.22). Uncertainty is also an important determinant of some kinds of adaptations. For example, predictable variation is more likely to favor adaptation by phenotypic plasticity, while unpredictable variation is more likely to favor adaptation by a jack-of-all-trades strategy (see Chapter 9).

The distinction between random and predictable patterns of change depends on the life-span of the plant relative to the rate at which the climate changes. The ENSO, for example, varies enough in its details (e.g., intensity, occurrence, and duration) that for an annual plant it is effectively random, while for a *Sequoia sempervirens* (redwood, Cupressaceae) that lives for centuries, it forms a relatively predictable cyclic pattern.

At the longest time scale—tens of millions of years—continental drift and tectonic activity play a major role in changing climate. Mountain building, for example, creates rain shadows. The spread of grasslands across

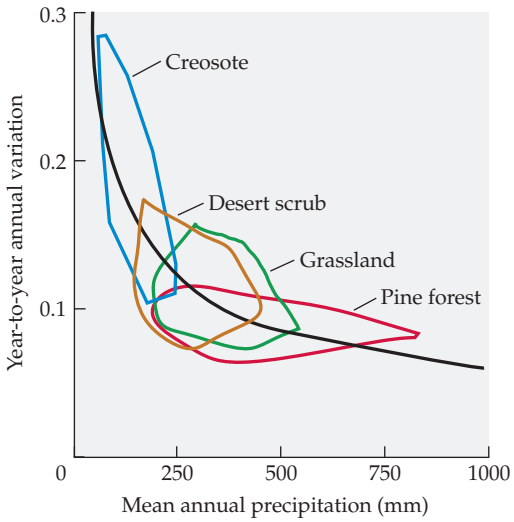



**Figure 16.22** Association between the variation and predictability of precipitation and the presence of different types of plant communities in Arizona. Desert areas dominated by creosote bush (lower Sonoran Desert) or desert scrub (upper Sonoran Desert) have low amounts of precipitation and high year-to-year variation. As precipitation amounts increase and variation decreases, plant communities become dominated by grasses and then by pine forest. (After G. Davidowitz. 2002. *Global Ecol Biogeogr* 11: 143–154.)

the middle of North America is associated with the rise of the Rocky Mountains during the Eocene epoch 34 million years ago (see Figure 17.1). As continents move with plate tectonics across the surface of the Earth, their climates change. For the past 100 million years, Australia has been moving northward; as a result, its climate has changed from mainly temperate and polar to mainly tropical and subtropical.

Changes in atmospheric $CO_2$ concentrations, both natural and anthropogenic, have exacerbated or caused large changes in global climate in the past as well as the present, but human additions of $CO_2$ to the atmosphere are now changing conditions rapidly. We now move on to examine the causes and ecological consequences of current human-generated global climate change.

## 16.4 Anthropogenic Global Climate Change Is Caused by Humans and Is Affecting Vegetation

These are the basic scientific facts about global climate change:

- Earth's climate is changing rapidly.
- Humans are directly responsible for much of the change in Earth's climate.
- These changes are already happening.
- Changes to Earth's climate will become increasingly severe in this century if nothing is done or if only modest steps are taken to mitigate climate change.
- Some, but not all, of the consequences of human impacts on Earth's climate are already irreversible.
- Scientists and others have known about, predicted, and warned about these climate effects for over 125 years.
- While there is considerable political controversy about human-caused climate change, it is very real and is not scientifically controversial.

In this section, we begin with the global carbon cycle and the role of plants in Earth's climate, explain the underlying science of human-caused climate change, and examine the evidence. We then briefly consider the consequences and predictions for weather and climate, plants and the natural world, and humans. We will also look at causes and some suggested actions that could be taken to reduce the enormous toll on humans and natural systems.

The global carbon cycle and our alterations of that cycle are central to some of the most dramatic and important of these changes to Earth's climate. One of the major consequences of alterations in that cycle is a continuing rise in average temperatures. Some of the other predicted changes in weather include shifts in rainfall patterns such that some areas will become wetter and others drier, changes in the seasonality of rainfall in many locations, and an increase in storm intensity, including heavier rainfall and stronger hurricanes. These changes are already affecting most organisms and ecological systems.

The global carbon cycle is intimately intertwined with plant ecology and with all life on Earth. Plants both drive and respond to the carbon cycle. To adequately address the role of plants and the consequences for plants in this great global drama, we must digress in places from our focus on plants to look more closely at the actions of people. As conditions for plant growth are changing over the face of the globe, dramatic changes—some predictable, but others unpredictable—are affecting agricultural systems and natural ecosystems in all biomes. Plant and animal distributions are shifting, natural communities now have different species compositions that will continue to change, some species have become extinct and many more will become extinct in the future, and some natural communities will decline or vanish altogether. Novel communities and ecosystems will result from the changes in species distributions and extinctions. Humans and human institutions (such as national economies) will be affected by these changes as well, both directly, by changing

weather patterns and rising sea levels, and indirectly, by alterations in agricultural productivity (for better or worse, depending on the location and other factors).

Anthropogenic (human-caused) global change encompasses far more than shifting weather patterns and the consequences of changes in climates. Other global changes caused by humans over the past 150 years include greatly altered patterns of land use and a large and continuing decline in the diversity of plants and animals living on the Earth. We discuss these other global changes, particularly focusing on the role of and impacts on plants and vegetation, in Chapter 19.

## The global carbon cycle is central to Earth's climates

The uptake of carbon by plants was introduced in Chapter 2, and we examined in detail how carbon moves through local ecosystems in Chapter 5. To review briefly, plants take up carbon in the form of $CO_2$ from the atmosphere and incorporate it into organic compounds using energy from sunlight. Carbon moves through ecosystem food webs via consumption of plants and plant parts by other organisms and by the actions of decomposers. Organic carbon accumulates in ecosystems for a period of time, both in living organisms and as soil organic matter. Carbon returns to the atmosphere primarily through the respiration of soil organisms. We now widen our focus to a global scale, examining the pools and fluxes of carbon in terrestrial environments, the atmosphere, and the oceans (Figure 16.23).

By far the largest active pool of carbon is the carbon dissolved in ocean waters. Active pools interact directly with living organisms over time scales of less than thousands of years and exist on or near the Earth's surface. The large pools of carbon contained in rock, oil, and natural gas are not considered active pools or part of the natural global carbon cycle, because without human intervention, most of that carbon is not involved in global carbon fluxes. Most of the marine carbon is found in deep ocean waters, and the exchange of this carbon with the atmosphere (its turnover) occurs slowly—a carbon molecule in the ocean stays there, on average, for 350 years. Surface waters exchange $CO_2$ with the atmosphere much more quickly. It takes about 11 years, on average, for a molecule of $CO_2$ in the atmosphere to be dissolved at the ocean surface and then released again. Slightly more $CO_2$ is dissolved every year than is released, so the oceans currently act as a net sink for carbon.

Another substantial pool of carbon is found in soil organic matter. A major flux of $CO_2$ into the atmosphere

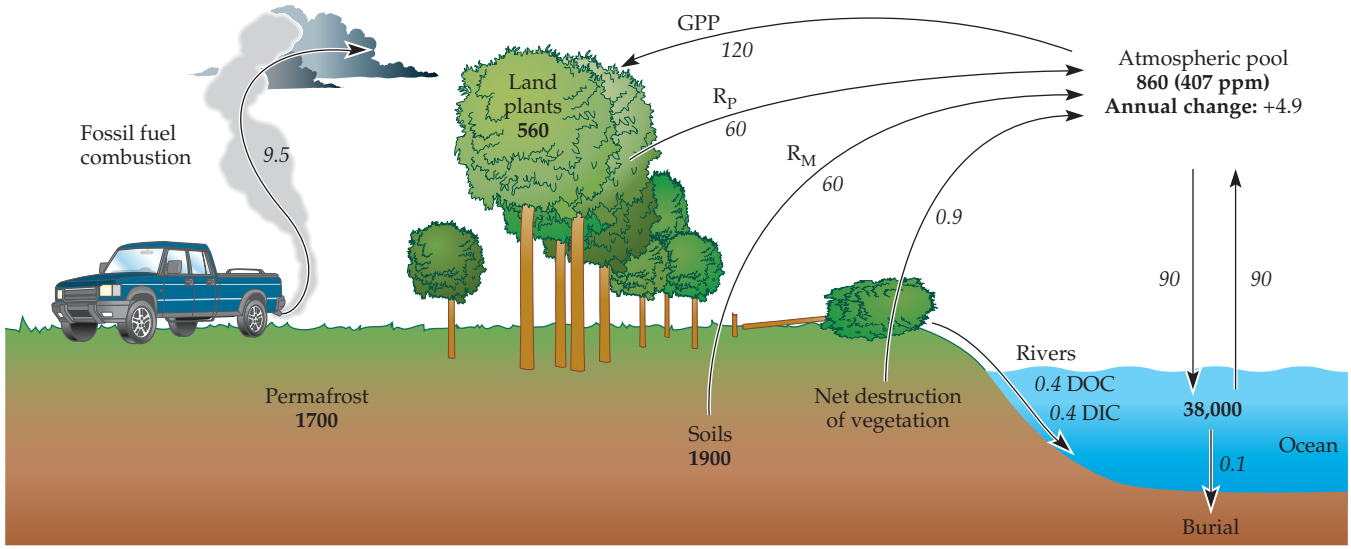



**Figure 16.23** The global carbon cycle. The numbers show the values for major global pools (in units of $10^{15}$ g C, shown in bold type) and fluxes (in units of $10^{15}$ g C/year, in italic type) between components of the land, oceans, and atmosphere. Notice that land plants take up about twice as much carbon annually as they release, and they store about 65% as much carbon as the atmosphere. The Earth contains about $10^{23}$ g carbon in total; most of this carbon is buried in sedimentary rocks. Only about $10^{18}$ g carbon is contained in active pools, as shown here; the largest of these is marine dissolved inorganic carbon. DOC, dissolved organic carbon; DIC, dissolved inorganic carbon; GPP, gross primary production; $R_P$, plant respiration; $R_M$, soil microbial respiration. (After W. H. Schlesinger. 1997. *Biogeochemistry: An Analysis of Global Change*. Academic Press: New York, based on D. S. Schimel et al. 1995. In *Climate Change 1994: Radiative Forcing of Climate Change and An Evaluation of the IPCC IS92 Emission Scenarios*. pp 35–71. J. T. Houghton et al. [Eds.]. Cambridge University Press; UK. Updates based on P. Friedlingstein et al. 2019. *Earth Syst Sci Data* 11: 1783–1838.)

comes from the respiration of soil organisms, primarily microorganisms (Schlesinger 1997). The atmosphere and living plants contain the two other major active pools of $CO_2$, with somewhat more carbon stored in the atmosphere than in plants. (The organic matter stored in multicellular animals is minuscule because the total biomass of animals is small relative to the global carbon budget.)

Gross primary production—the total flux of $CO_2$ into living terrestrial plants—removes the greatest amount of carbon from the atmosphere yearly. About half of this carbon is returned to the atmosphere through the respiration of terrestrial plants, so terrestrial global NPP (net primary production; see Chapter 5) is about $60 \times 10^{15}$g C/year. Marine phytoplankton—particularly cyanobacteria—exchange $CO_2$ directly with surface ocean waters and only indirectly with the atmosphere. NPP in the oceans is estimated to be about $35–50 \times 10^{15}$g C/year, or somewhat less than that of terrestrial plants. Most of the carbon fixed by marine phytoplankton is taken up by heterotrophic bacteria and zooplankton in surface waters. Much of the $CO_2$ taken up by the heterotrophic marine bacteria is quickly transformed into inorganic forms again and released into the water, while the zooplankton form the base of marine food webs. Only a tiny proportion of the NPP eventually sinks to the ocean bottom, where it is incorporated into sediments. The total pool of carbon in marine organisms at any one time is very small, although the turnover rate of carbon in this pool is very rapid.

These fluxes of $CO_2$ are natural, although they can be affected by human activities. Two additional important fluxes in the global carbon cycle, however, are entirely anthropogenic, at least at their current scales. These two fluxes, both of which add $CO_2$ to the atmosphere, result from the cutting and burning of large areas of forests and from the extraction and combustion of fossil fuels, including oil, natural gas, and coal. The immediate consequence of these additional fluxes is that atmospheric concentrations of $CO_2$ are rising. The actual concentration of $CO_2$ in the atmosphere is surprisingly small, although its total weight in the atmosphere is a hefty $3 \times 10^{12}$ metric tons. Before the Industrial Revolution of the nineteenth century, $CO_2$ was approximately 280 ppmv (parts per million by volume) of the atmosphere at the Earth's surface (Figure 16.24). The concentration of $CO_2$ was 407 ppmv on average in 2018; it was 370 ppmv on average when the previous edition of this textbook was written. If we continue our current practices, it could reach twice preindustrial levels—560 ppmv—by 2065. These higher levels of $CO_2$

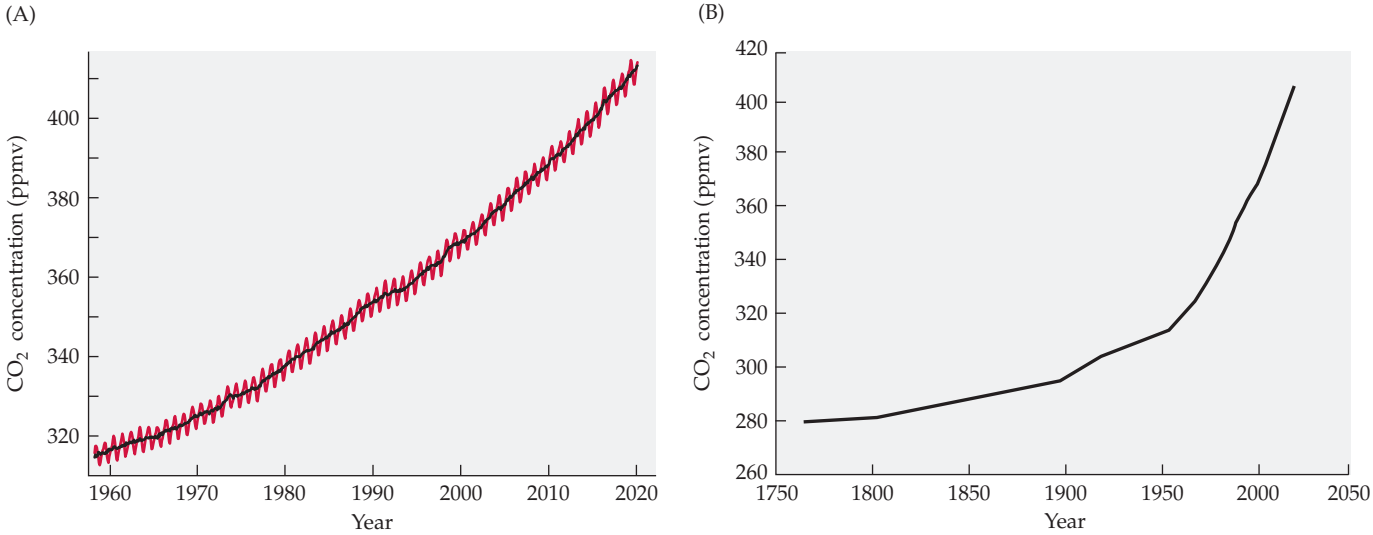



**Figure 16.24** Increases in atmospheric concentrations of $CO_2$. (A) Historic increases in atmospheric $CO_2$ (in parts per million by volume, ppmv) recorded at the Mauna Loa Observatory in Hawaii. This site was chosen because it is not subject to the variation in $CO_2$ levels found close to anthropogenic emissions, but results are similar to those at many other sites. Monthly values shown by data points; wiggling line is due to seasonal variation resulting from Northern Hemisphere terrestrial vegetation photosynthesizing most in summer and least in winter. The long-term trend is shown by the black line. Data for 1958 to1974 are from the Scripps Inst. Oceanography; data for 1974 to 2020 are from NOAA. (B) The trend in $CO_2$ concentrations over the past two centuries, determined from bubbles of air in ice cores from Antarctica; values since 1990 are from atmospheric values measured at Mauna Loa. (A after C. D. Keeling et al. 1976. *Tellus* 28: 538–551; K. W. Thoning et al. 1989. *J Geophys Res* 94: 8549–8565, with updates from NOAA/ESRL/Global Monitoring Division. 2020. [https://www.esrl.noaa.gov/gmd/ccgg/trends/]; B after P. M. Vitousek. 1992. *Annu Rev Ecol Syst* 23: 1–14, with updates from NOAA/ESRL/Global Monitoring Division. 2020. [https://www.esrl.noaa.gov/gmd/ccgg/trends/].)

have direct effects on plants as well as indirect effects mediated by changes in climate. We first discuss these effects, then examine the sources of the carbon humans are adding to the atmosphere.

## Increasing atmospheric $CO_2$ has direct effects on plants

For plants that use $C_3$ photosynthesis—that is, the vast majority of photosynthesizers—the concentration of $CO_2$ in the atmosphere is suboptimal (see Chapter 2). It is reasonable to assume, therefore, that rising atmospheric $CO_2$ levels might act to increase plant growth and ecosystem productivity. Elevated $CO_2$ levels can also potentially increase the efficiency of water and nutrient use (at least for $C_3$ plants), affect leaf chemistry (and therefore leaf litter decomposition rates as well herbivory), and alter patterns of carbon allocation within the plant. Because different plant species respond differently to enhanced $CO_2$ levels, global changes in atmospheric $CO_2$ concentrations might change intra- and interspecific competitive interactions among plants, leading to changes in community composition, possibly favoring fast-growing invasive plant species (although those results are currently inconclusive; Sorte et al. 2013). Adaptive evolutionary change in response to selection by elevated $CO_2$ could change the future genetic composition of plant populations. Elevated $CO_2$ could also alter local ecosystem carbon cycles, by modifying productivity and decomposition rates, and could change nitrogen cycling and other ecosystem-level processes. Some ecologists have predicted that although there will be large positive responses to higher $CO_2$ levels initially, they will be followed by adjustment and reduced responses over time, over a range of scales from individual plants to ecosystems. It is also very possible that responses to changing water availability, high temperatures, longer growing seasons, and other characteristics of climate change will have far greater effects on plants, plant communities, and ecosystems and that these will dwarf the direct responses to elevated $CO_2$. There is now an enormous amount of research on plant responses to rising $CO_2$ levels in both natural and agricultural systems.

$C_3$ plants are predicted for physiological reasons to respond more than $C_4$ plants to elevated $CO_2$, with the expectation that $C_4$ plants would be at a competitive disadvantage in an enhanced-$CO_2$ world (see Chapter 2). Stephanie Wand and her colleagues (1999) carried out a meta-analysis of studies on the responses of nondomesticated grasses to elevated $CO_2$. Across all studies, both $C_3$ and $C_4$ grasses had higher photosynthetic rates at higher $CO_2$ levels, with increases averaging 33% and 25%, respectively. Total biomass for $C_3$ and $C_4$ grasses increased by 44% and 33%, respectively, and leaf water use efficiency was also greater for both grass types. The $C_3$ and $C_4$ grasses did not differ significantly from one another in these responses, although $C_4$ species had consistently smaller average responses than $C_3$ species, and the two groups did differ in some of the other responses studied. It was surprising that $C_4$ grasses had strong positive responses to elevated $CO_2$ given their superior abilities to capture $CO_2$, and that photosynthesis and growth responses did not differ significantly between the two groups.

However, the studies in this meta-analysis were almost all short-term studies in controlled environments. In an outdoors free air $CO_2$ enrichment (FACE) experiment lasting 21 years, Melissa Pastore and coinvestigators (2019) found that photosynthesis acclimated to elevated $CO_2$, resulting in relatively modest responses in $C_3$ (11%) and $C_4$ (6%) grasses and $C_3$ legumes (15%) and dicots (12%), although water use efficiency (WUE) increased substantially in all groups. These increases were considerably less than short-term and instantaneous photosynthetic responses.

In an analysis of experiments on the effects of elevated $CO_2$ on trees (all with $C_3$ photosynthesis), Peter Curtis and Xianzhong Wang (1998) found that both photosynthesis and total growth increased substantially when $CO_2$ levels were elevated to twice ambient atmospheric levels (generally 350 ppmv, which was ambient at the time the experiments were carried out, although this varied among experiments). Stressful conditions such as low soil nutrient levels reduced these effects. Furthermore, they found no evidence that trees eventually would acclimate to higher $CO_2$ levels.

The relationship between forests, grasslands, and savannas in responses to elevated $CO_2$ is complex and depends not only on atmospheric $CO_2$, but on temperatures; water regimes; fire intensity, frequency, and seasonality; and grazing and other human land use changes. A literature review by William Bond and Guy Midgley (2012) concluded that the large increases in trees in recent times in mesic savannas and grasslands might be due at least in part to elevated $CO_2$, although in arid grasslands and savannas, woody species encroachment is probably due to other factors. These changes are profoundly altering not only the ecology of grasslands and their balance with forests, but also the global carbon cycle and the connections between terrestrial ecosystems and the atmosphere.

It is probably safe to say that many plants would have some positive direct responses to higher atmospheric $CO_2$ levels, including higher photosynthetic rates, greater growth, and higher water use efficiency, if nothing else changed. However, as we will see below, the indirect effects of climate change on plants, communities, and ecosystems that are likely to result from elevated $CO_2$ levels are probably much larger than these direct effects.

## The greenhouse effect warms the Earth due to greenhouse gases

What is it about $CO_2$ molecules that makes them act as greenhouse gases (GHGs), absorbing infrared (IR) wavelengths and heating the air surrounding them, when $O_2$ and other molecules such as nitrogen ($N_2$) that dominate the atmosphere on Earth do not absorb IR wavelengths? When a $CO_2$ molecule absorbs a photon of the right wavelength in the infrared, the $CO_2$ molecule vibrates faster and more energetically, with the bonds between the carbon and oxygens stretching and bending (Figure 16.25). These energetic vibrations transform electromagnetic radiant energy into molecular kinetic energy, which is then released as longwave radiant energy, or heat. The greater the number of $CO_2$ molecules in the atmosphere, the more IR radiation is trapped, and the more heat is retained and reemitted as heat at the Earth's surface.

$CO_2$ is the most important GHG contributing to the enhanced greenhouse effect because it is being added to the atmosphere in the greatest quantities, but it is not the only GHG. Water vapor is the most abundant GHG, followed by $CO_2$. Both occur in the atmosphere naturally. The other major naturally occurring GHGs are $O_3$ (ozone), $N_2O$ (nitrous oxide), and $CH_4$ (methane) (Figure 16.26), although their concentration has been increased by human actions. In addition, chlorofluorocarbons (CFCs), formerly widely used in refrigeration, aerosols, and making insulating foam (but now phased out, discussed below), are human-created greenhouse gases. Water vapor behaves very differently than these other GHGs, because it responds to changes in the temperature of the atmosphere rather than directly causing these

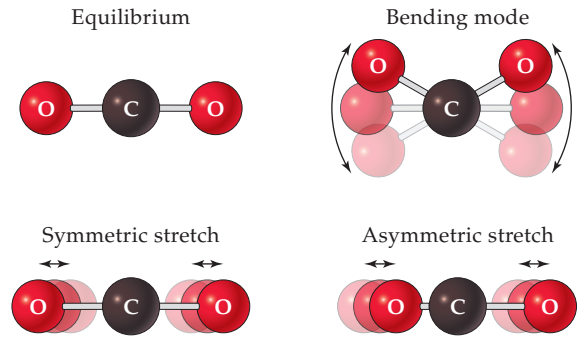



**Figure 16.25** The capture and absorption of an IR photon causes the $CO_2$ molecule to vibrate, stretching and bending the bonds between the carbon and oxygen atoms. Faster, more energetic vibrations change electromagnetic energy into kinetic energy of the $CO_2$ molecules, warming them; the $CO_2$ molecules then release this energy to their surroundings as longwave IR radiant energy, or heat. IR photons are not captured or absorbed by $O_2$ or $N_2$ molecules in the atmosphere. Courtesy of Jan Cami [http://www.astro.uwo.ca/~jcami/molspec_intro.shtml].

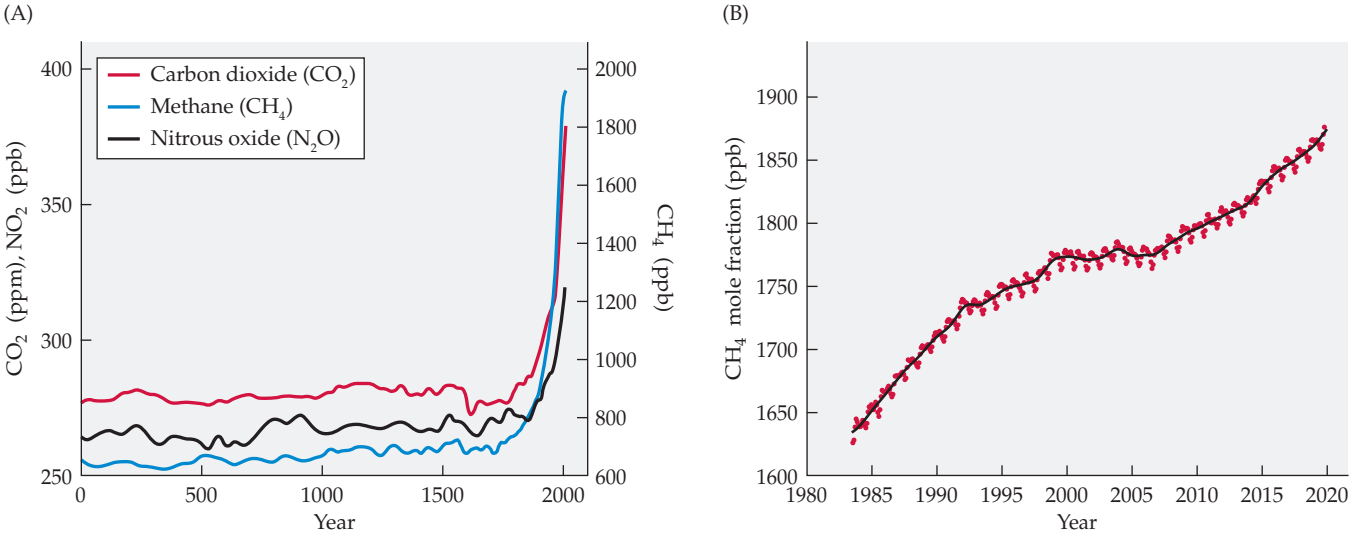



**Figure 16.26** Increases in atmospheric concentrations of methane and other major greenhouse gases. (A) Atmospheric methane ($CH_4$, in ppb), carbon dioxide ($CO_2$, in ppm), and nitrous oxide ($N_2O$, in ppb) during the past 2000 years. From the Intergovernmental Panel on Climate Change. (B) Global monthly averages of atmospheric $CH_4$ from 1983 to 2020; averages across multiple sites. (A after P. Forster et al. 2007. In *Climate Change 2007: The Physical Science Basis. Contribution of Working Group 1 to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.* S. Solomon et al. [Eds.], p. 135. Cambridge University Press: Cambridge, UK; B after E. J. Dlugokencky et al. 1994. *J Geophys Res* 99(D8): 17021–17043, with updates from NOAA/ESRL/Global Monitoring Division. 2020. [https://www.esrl.noaa.gov/gmd/ccgg/trends_ch4/]).

changes. Water vapor is a very dynamic part of the climate/atmosphere system; it evaporates, condenses, and precipitates rapidly depending on temperature, acting as a fast feedback on Earth's climate (Lacis et al. 2010). The other GHGs are far more stable in the atmosphere and do not readily condense or precipitate out. A molecule of $CO_2$ in the atmosphere can remain there for thousands of years. So, $CO_2$ and the other noncondensing GHGs maintain the entire terrestrial greenhouse effect, keeping Earth warm enough to be habitable for living organisms. If the noncondensing GHGs were gone, Earth's atmospheric, ocean, and land temperatures would rapidly drop by almost 40°C. Without the noncondensing GHGs, 90% of the water vapor in the atmosphere would rapidly condense out of the air, and global sea ice would become ten times as great, with only tropical seas being ice-free. It is the $CO_2$ molecules that are the main controlling factor in the greenhouse effect, with a boost from $CH_4$ and the other GHGs. Water vapor, clouds, and precipitation act quickly in response to temperature but in themselves they do not control temperature.

The concentrations of $CH_4$, $N_2O$, $O_3$, and CFCs have been greatly increased in the atmosphere by humans. Methane is released in large quantities by leakage during the drilling and transportation of natural gas, by digestion by anaerobic prokaryotes in cattle rumens (and then belched out by the cattle), and by the decomposition of trash in landfills and dumps. Methane released from garbage dumps can be captured and used as an energy source, but it rarely is; this could replace some combustion of fossil fuels as well as reducing emissions of this greenhouse gas to the atmosphere. Increases in $N_2O$ are largely a by-product of commercial agriculture, as it is released by microbial degradation of nitrogen-containing fertilizers; nylon production is another substantial source of $N_2O$. While humans produce far smaller amounts of $CH_4$, $N_2O$, $O_3$, and CFCs than of $CO_2$, these other greenhouse gases are much more potent absorbers of infrared radiation. $CH_4$ traps more than 20 times the amount of heat per molecule than $CO_2$, while $N_2O$ absorbs 270 times more. $CO_2$ emissions are responsible for about two-thirds of the anthropogenic intensification of the greenhouse effect, with most of the remainder due to $CH_4$ (about 19%), CFCs (10%), and $N_2O$ (6%) currently.

Ozone is produced in the troposphere (the lower atmosphere ) through a series of chemical reactions in areas with high levels of air pollution, particularly in the form of volatile hydrocarbons and CO (carbon monoxide) in the presence of NO (nitrogen oxide). Some ozone is also produced naturally from volatile hydrocarbon compounds emitted by vegetation and by forest fires. Ozone is harmful in the troposphere as a pollutant and also as a GHG, but it protects against damaging UV light when it is a stratospheric gas much higher in the atmosphere. Sometimes damage to ozone high up in the stratosphere—the creation of an "ozone hole"—is confused with climate change due to the greenhouse effect (Box 16B).

How long have we known about the greenhouse effect and increasing global warming due to higher atmospheric $CO_2$ levels? The first scientist to discover and experimentally demonstrate that $CO_2$ was essential for

## BOX 16B The Ozone Hole and the Greenhouse Effect

The human-caused intensification of the greenhouse effect is often confused with the problem of the ozone hole in the atmosphere. They are two different—although related—problems. The *ozone hole* refers to the loss of parts of the protective layer of ozone high in the stratosphere (the part of the atmosphere from 10 to 50 km above the Earth's surface), caused primarily by a class of artificial compounds called halogenated hydrocarbons (including CFCs and related chemicals). Historically, the stratospheric ozone layer absorbed much of the high-energy ultraviolet (UV) radiation coming from the sun, preventing it from reaching the Earth's surface. Halogenated hydrocarbons destroy ozone molecules in the stratosphere, allowing more UV radiation to reach the Earth's surface. This effect has been particularly severe at high latitudes in the Southern Hemisphere. The resulting increase in UV radiation poses a risk of damage to many organisms, including humans, including increases in cancer risk and damage to DNA. The ozone hole is a breakdown in stratospheric ozone, leading to an increase in damaging UV radiation. Because of these dangers, an international treaty was reached in 1987 to replace these harmful chemicals with safer alternatives that do not destroy stratospheric ozone. Over the more than three decades since then, this has saved millions of people from dying from skin cancer and reduced other harm to humans and organisms in nature. The ozone hole shrank to its smallest size ever recorded in 2019, which is great news. The ozone hole is not the same thing as the enhanced greenhouse effect, which is a buildup of heat-trapping gases in the atmosphere, leading to changes in global climates. The ozone hole is connected to the greenhouse effect and the problem of global warming specifically because CFCs act both as destroyers of stratospheric ozone and as greenhouse gases. More generally, both the ozone hole and changes to the greenhouse effect are anthropogenic changes in the Earth's atmosphere that have substantial consequences for organisms and ecosystems, including humans.

warming Earth's atmosphere—the greenhouse effect—and recognizing the possibility that high levels of $CO_2$ might cause Earth's climate to warm was an American scientist, Eunice Foote, in 1856, but her insights were essentially ignored, and her contributions were not rediscovered until 2011. Earlier, in 1824, the French mathematician Joseph Fourier had argued that the atmosphere prevents infrared radiant energy from escaping to space, therefore warming the Earth (the 'natural' greenhouse effect). In 1859, the Irish physicist John Tyndall experimentally verified the properties of the major greenhouse gases in the atmosphere, firmly demonstrating the existence of the greenhouse effect (along with his many other remarkable accomplishments). The Swedish chemist Svante Arrhenius, the "father of climate change science," quantified the greenhouse effect and was the first scientist to publish the prediction in 1896 that $CO_2$ from the combustion of fossil fuels would lead to global climate change. So, this information is not exactly new to science in the twenty-first century.

By the 1930s it was confirmed that $CO_2$ concentrations in the atmosphere were actually rising, and by the 1950s there were precise measurements of the magnitude of the increase. The prediction of global warming due to anthropogenic increases in greenhouse gases was common and accepted knowledge in the scientific community and among many educated people by the 1970s. Most scientists had accepted that human-caused global warming not only was a possibility, but was already occurring by the 1990s.

## 16.5 Humans Are Changing the Global Carbon Cycle

Where does the carbon responsible for these climate changes come from? Increases in atmospheric $CO_2$ are caused mainly by two factors: **deforestation** (the partial or complete removal of a forest) and the combustion of fossil fuels (Figure 16.27). Where are these things happening, and who is causing them? To fully investigate these questions, we would need to discuss not only the science involved, but also to touch on the social, political, and economic bases for anthropogenic $CO_2$ emissions. These topics can be treated to only a limited extent in a plant ecology textbook. Here we mention some of the factors that are involved.

### Fossil fuel combustion is the most important factor changing the greenhouse effect

The burning of fossil fuels provides the greatest anthropogenic contribution to atmospheric $CO_2$. The carbon contained in fossil fuels accumulated over eons as it was taken up from an ancient atmosphere by organisms now long extinct. Coal was first mined industrially as a fuel

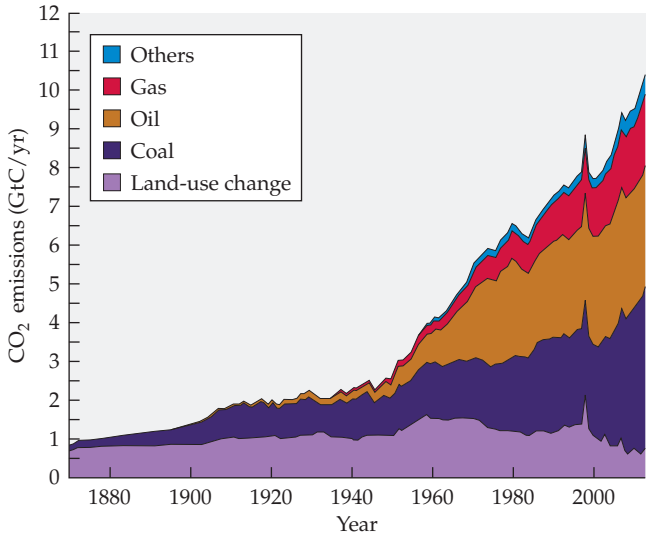



**Figure 16.27** Carbon emissions to the atmosphere (in GtC/year; a Gt [gigaton] is $10^{15}$g/year) from combustion of fossil fuels (gas, oil, and coal) and from carbon flux to the atmosphere from land use change (primarily deforestation) from 1870 to 2010. The filled portion of each segment indicates the amount of carbon emitted. (From GLOBAL CARBON PROJECT [2013] Global Carbon Budget https://www.globalcarbonproject.org/carbonbudget/archive/2013/GCP_budget_2013.pdf.)

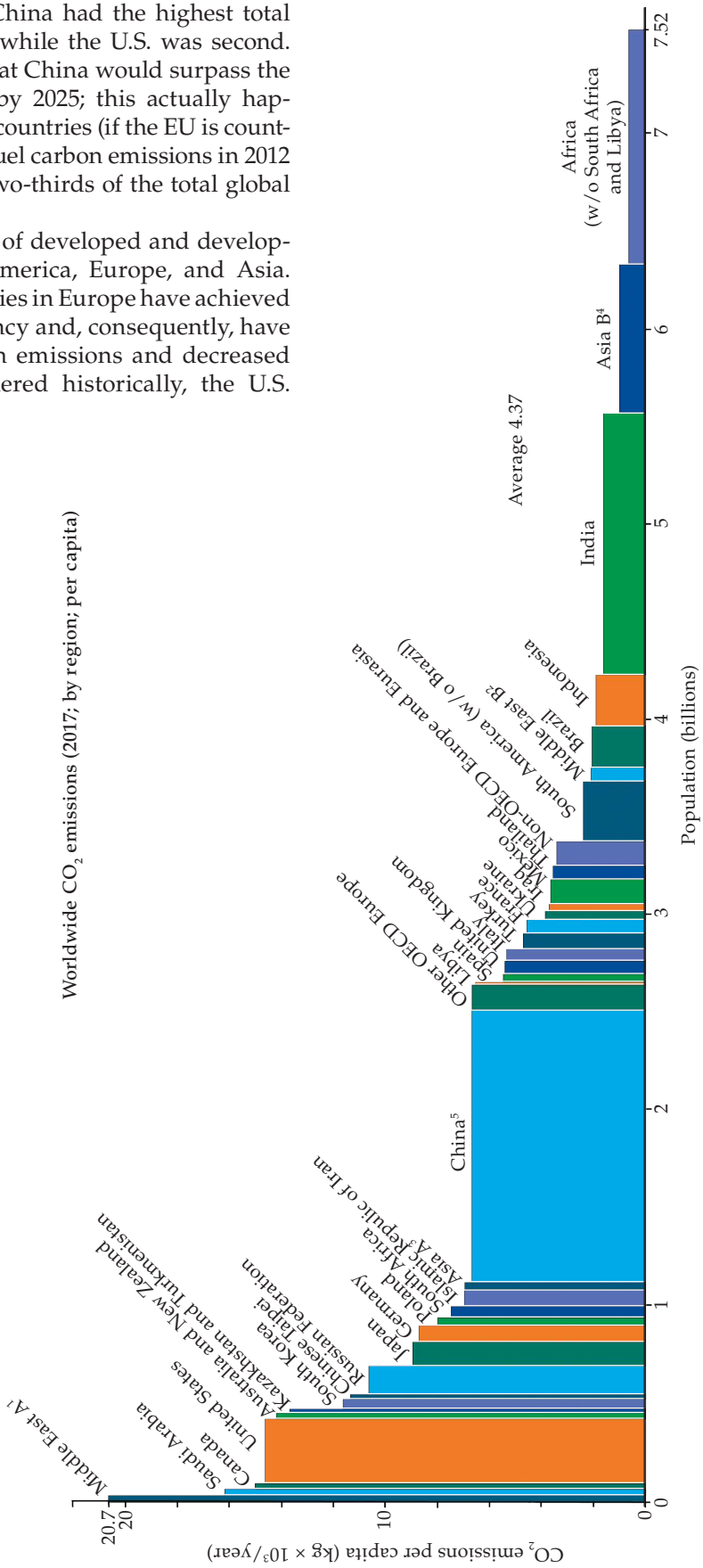
source in the mid-eighteenth century, and by the mid- to late-nineteenth century coal and oil were being extracted for energy. Vast amounts of carbon in the form of fossil fuels—coal, oil, and natural gas—have been extracted in less than 300 years from deep in the Earth, where they accumulated over many millions of years, combusted and released in a geological blink of an eye. The extractable pool of carbon—the portion currently economically feasible to remove—contained in the Earth's fossil fuels is about $4 \times 10^{18}$ g C, which is about 10% of the total active pool of carbon at or near the Earth's surface. The combustion of fossil fuels releases approximately $9.5 \times 10^{15}$ g C/year. A little over half of this carbon remains in the atmosphere; almost all the rest enters the oceans, and small amounts are currently unaccounted for. Not until the middle of the twentieth century did the amount of carbon released to the atmosphere by the combustion of fossil fuels become greater than that released by the cutting or burning of forests, but now it represents about 70% of all anthropogenic $CO_2$ released (see Figure 16.27).

$CO_2$ emissions differ enormously among the countries of the world, whether we look at the total amount released into the atmosphere by each country or consider how much, on average, each person in the country is responsible for producing (Figure 16.28A). There are different ways of looking at who is putting GHGs into the atmosphere; the details are nuanced and complex, yet the whole picture is in sharp focus. As of the

publication of this book, China had the highest total emissions of any country, while the U.S. was second. In 2005, it was predicted that China would surpass the U.S. in carbon emissions by 2025; this actually happened by 2007. The top ten countries (if the EU is counted together) in total fossil fuel carbon emissions in 2012 accounted for more than two-thirds of the total global emissions (Figure 16.28B).

This list contains a mix of developed and developing countries in North America, Europe, and Asia. Most industrialized countries in Europe have achieved increases in energy efficiency and, consequently, have both decreased per person emissions and decreased total emissions. If considered historically, the U.S.

(A)




**Attribution:**
Based on IEA data from IEA (2019) "World $CO_2$ E from Fuel Combustion", 2019 edition, www.iea.org/statistics, All rights reserved; as modified by AQAL Capital GmbH and Tom Schulz.
This map is without prejudice to the status of or sovereignty over any territor, to the delimitation of international frontiers and boundaries and to the name of any territory, city, or area.
This Work is licensed under a Creative Commons Attribution-Share Alike 3.0 Unported License

**Notes:**
Energy-related $CO_2$ emissions only; no other greenhouse gases
[1] Middle East A: Bahrain, Oman, Kuwait, Qatar, United Arab Emirates
[2] Middle East B: Israel, Jordan, Lebanon, Syrian Arab Republic, Yemen
[3] Asia A: Brunei Darussalam, Malaysia, Mongolia, Singapore
[4] Asia B: Asia w/o Asia A, China, India, Thailand, Chinese Taipei, Indonesia, Korea, Japan
[5] China: People's Republic of China, Hong Kong
Version: 17-Dec-2019 by AQAL Capital GmbH (https://aqalcapital.com) and Tom Schulz

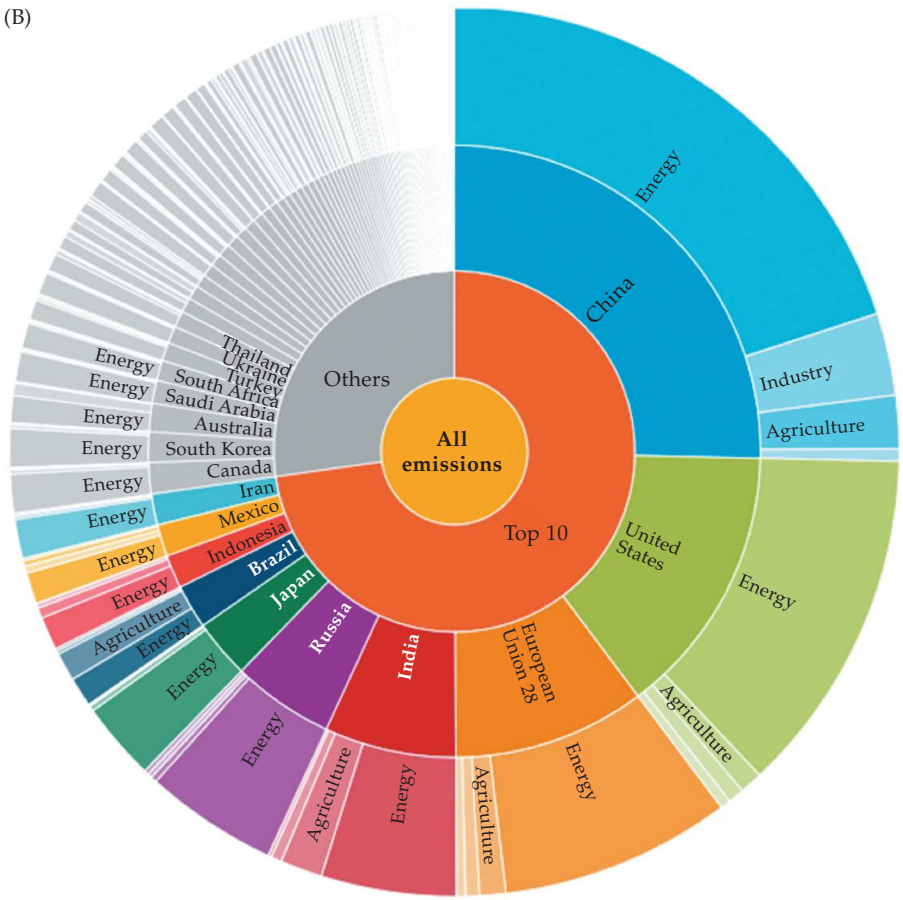



**Figure 16.28** (A) $CO_2$ emissions for various countries and groups of countries. The total emissions for each country or group of countries are indicated by the area of the bar. The height of each bar shows the emissions per capita (per person); the width is the population of the country or group. The bars are arranged in order of $CO_2$ emissions per capita. The European OECD as shown here is a group of industrialized Western countries . Non-OECD countries as shown here include a large number of less-developed countries in Eastern Europe and Euroasia. (B) Each 'slice' is proportional to a country's total annual emissions, and the economic sectors contributing most to emissions are indicated; energy is the dominant source of emissions. In 2013, the top 10 GHG emitters accounted for almost 75% of the global emissions total, and the top 3 (China, U.S.A., and the EU) emitted half of the GHG. The top carbon emitters in the EU were Germany, U.K., Italy, France, Poland, Spain, the Netherlands, Czech Republic, Belgium, and Greece. The bottom 100 together emitted less than 4% of the GHGs. (https://cdiac.ess-dive.lbl.gov/trends/emis/top2014.tot). Users can examine GHG emissions interactively at https://www.wri.org/blog/2017/04/interactive-chart-explains-worlds-top-10-emitters-and-how-theyve-changed. (B after J. Friedrich et al. 2017. *This Interactive Chart Explains World's Top 10 Emitters, and How They've Changed*. Online blog, April 11, 2017. CC BY 4.0. https://www.wri.org/blog/2017/04/interactive-chart-explains-worlds-top-10-emitters-and-how-theyve-changed.)

has added the most carbon to the atmosphere in total, followed by Germany and the UK. Considered differently, the 20 largest fossil-fuel-extracting, multinational public companies have produced the fuel responsible for nearly 30% of the total carbon added to the atmosphere; 5 of those companies count for about 12% of that total. Two-thirds of cumulative $CO_2$ emissions since the Industrial Revolution have been from the largest 90 oil, natural gas, coal, and cement producers (Heede 2014). Fossil fuel extraction and combustion have produced enormous profits for those companies, employed large numbers of people, and allowed many people to live much more comfortable, healthy and secure lives, particularly those in industrialized countries plus the wealthy people in poorer countries. However, the costs in money, lives, and damage to property and ecological systems has been and will be extremely large. On a very hopeful note, global atmospheric $CO_2$ emissions are steeply declining per unit of wealth produced, although the picture is complex (Figure 16.29).

The total carbon emissions of a country are the product of the average amount emitted per person (also called per capita or "per head") and the total number of people in that country. The U.S. per person emissions declined to about 80% of what they were from 1980 to 2013; the per person emissions in China were relatively very low in 1980 but increased by almost five times during that period of time, while in Germany in 2013 they were about 70% of what they were in 1980. Per person annual emissions range over more than three orders of magnitude, from over 70 tons of carbon in Qatar to 0.03 tons in Madagascar. Per person increases in rapidly developing countries such as India, with large and growing populations, have led to steep increases in those countries' total emissions. On the other hand, the tiny per person emissions of countries such as Madagascar (<0.1 metric tons per year in 2013) are largely due to most people living at bare subsistence levels. Of course, average emission rates per person are not equally distributed, and some people in any given country use far more energy in the form of fossil fuels and add much more carbon to the atmosphere than others.

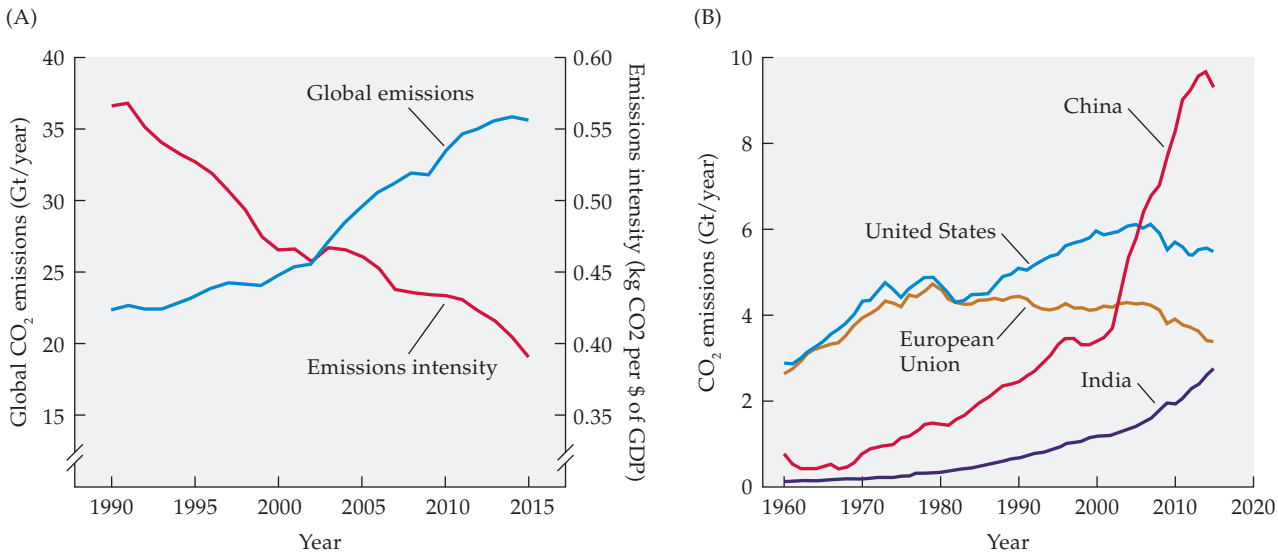



**Figure 16.29** (A) Global atmospheric $CO_2$ emissions are steeply declining per unit of wealth produced; although $CO_2$ emissions continue to increase, it takes far less in $CO_2$ emissions to produce a unit of wealth measured as gross domestic product (GDP) than it did in the past. (B) Annual $CO_2$ emissions from 1960 to 2015 for China, U.S., EU, and India in Gt/year. (A after R. B. Jackson et al. 2016. *Nat Climate Change* 6: 7–10; B after C. Le Quéré et al. 2016. *Earth Syst Sci Data* 8: 605–649. CC BY 3.0.)

What are people doing that causes GHG emissions from fossil fuels? Globally, electricity and heat production are the largest sources of emissions from fossil fuels (Figure 16.30). Different fossil fuels release very different amounts of carbon for each unit of energy generated—coal, for example, produces much more atmospheric carbon for the same amount of energy yield than natural gas. Other energy sources produce no carbon when used to generate electricity: nuclear power, hydroelectric power, solar power, wind, and geothermal energy all release no carbon directly. They may cause other environmental problems, however.

Emissions globally come from the major categories of agriculture, changing land use, transportation, residential use, and industrial use. There are many different ways to define and categorize fossil fuel use within these major categories and activities. Industrial uses include manufacturing almost everything, including chemical production and use of computers, including running data centers. For instance, it has been estimated that the creation of Bitcoin releases about 22 megatons of $CO_2$ annually (Stoll et al. 2019). Two large industrial sources of GHGs in industry are nitrogen fertilizer (about 3% of global emissions) and cement manufacture (about 5% of global emissions, from the calcium carbonate [$CaCO_3$] in limestone). Land use changes include the removal of forests, which results in the loss of $CO_2$ uptake by trees and soil, emissions from burning and decomposition of their wood, and the energy used to remove the forests and alter the land. Forest regeneration and **afforestation** (planting trees to create new forests) can reverse $CO_2$ loss to the atmosphere and decrease total emissions. Soils can act as a net sink or a large source for carbon, depending on how they are managed.

Transportation includes individual cars and trucks, civilian and military buses, trains, planes, and ships and all of the transportation uses for industrial and agricultural production, most of which involves combustion of oil (primarily in the form of gasoline and diesel). Residential uses include fossil fuels used for heat and cooking plus those used for electricity generation for these purposes plus refrigeration, cooling and electronics. There are also emissions from fuel extraction, refining, processing, and transporting the fuels.

These categories of emissions vary greatly by country and region. For example, in the U.S., emissions from electricity generation, transportation, and industrial uses dominate. Although cars and trucks became much more efficient due to standards set by the federal government, progress toward those standards is being reversed as of the writing of this book, and moreover, the most popular personal vehicles in the U.S. are large pickup trucks, which use the largest amount of fossil fuels. Hybrid gasoline-electric vehicles and electric vehicles remain a small proportion of total vehicles, and mass transportation is very limited in most parts of the country and faces stiff political opposition. Afforestation and forest and wetland restoration cause a net reduction in $CO_2$; they can result in zero net emissions if they equal the $CO_2$

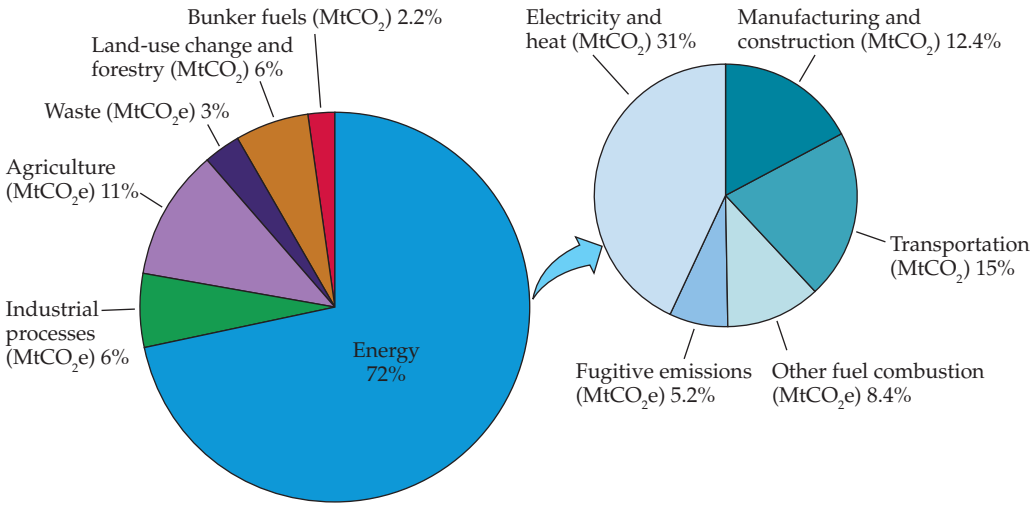



**Figure 16.30** Globally, energy production of all types accounts for 72% of all emissions; energy is further broken down at right. Major sources are electricity and heat (31%), agriculture (11%), transportation (15%), forestry and land-use change (6%) and manufacturing/construction/industrial processes (18% together). Waste and fugitive emissions contribute 8% of emissions. Bunker fuels are fuels used by ships. Units are metric tons of $CO_2$ equivalent, $MtCO_2e$. (CAIT Climate Data Explorer. 2017. Washington, DC: World Resources Institute. CC BY 4.0. http://cait.wri.org. CAIT data derived from several sources.)

emitted by deforestation and other land use changes. Agriculture is proportionately a smaller contributor to U.S. emissions compared with those of other countries, as are commercial and residential uses.

Globally, fossil fuel combustion contributed almost 80% of the GHG emissions from 1970 to 2011, and total emissions globally have increased by almost 90% since that time. Without question, there would be economic and other costs to changing our current patterns of fossil fuel consumption. But what are the alternatives? The consequences of our current actions—or lack of action—are already complex and costly in money, lives, quality of life and unintended outcomes, and they will only accelerate in the future.

## Deforestation and land use change also affect climate

Although trees naturally accumulate carbon in wood and release it when they eventually die, human activities have dramatically accelerated the rate at which this is happening. Cutting down a tree affects atmospheric $CO_2$ levels in two ways. First, if the tree is not replaced with another tree, the $CO_2$ it would have taken up will remain in the atmosphere. Second, if the tree is burned or decomposes, the carbon accumulated over its long life—which can be decades or centuries—is released to the atmosphere over a short time. Herbaceous plants, including most crops and garden plants, in contrast, do not accumulate carbon because they rapidly recycle it back to the atmosphere. The $CO_2$ that was taken up last summer by the wheat in your toast or the rice in your bowl is returned very shortly to the atmosphere as you metabolize those calories for energy.

If trees are logged to build things that remain intact for a long time (a good violin, perhaps, or a building that lasts a long time), then some of the carbon that was in their wood will be preserved for decades or even centuries, instead of being immediately released to the atmosphere, although the carbon in the roots, branches, and leaves that remain when the trunk is harvested might decompose quickly. If trees are logged to make newspaper, bulk mail advertisements, disposable chopsticks, or other short-lived items, the postdisposal fate of those items (whether they are burned, stored in a landfill, or recycled) will determine the fate of their carbon. The fate of the organic carbon in the soil of the logged forest will depend on a number of factors, including what happens to the land after the forest is logged. If new trees are planted where a forest is logged, the new trees will accumulate carbon, and net ecosystem production is likely to be positive (see Chapter 5 and Chapter 13). If the forest is replaced by buildings and pavement, that obviously will not happen. Removing forests and replacing them with agricultural lands, tree plantations, parking lots, expanding urban slums or residential developments has enormous ecological and other implications beyond the effects on the carbon cycle.

Two factors directly counter the emission of $CO_2$ to the atmosphere caused by forest destruction: forest regeneration and higher photosynthetic rates due to

higher $CO_2$ concentrations in the atmosphere. Forest regeneration has, in some cases, been dramatic. Large parts of the northeastern U.S., for example, were heavily logged and cleared for farming in the eighteenth and nineteenth centuries. The later abandonment of these farms due to a variety of social, economic, climate, and land management factors led to the regeneration of large areas of secondary forest, sufficient to act as a substantial carbon sink. Recently these lands have once again become sources rather than sinks for atmospheric $CO_2$, both because many of these forests have now reached maturity and their rate of carbon accumulation is declining, and because of renewed forest clearing for both lumber and suburban development.

Cutting trees for firewood and burning trees to clear land have been anthropogenic sources of carbon input to the atmosphere throughout human history. These activities were sufficiently limited before about 1800 that they had little or no measurable effect on atmospheric $CO_2$ concentrations (see Figure 17.7B). That is no longer the case, and the effects of these activities in tropical forests are particularly severe. Alexandra Tyukavina and her colleagues (2015) used satellite images to determine tropical forest losses across Africa, Asia, and the Americas by integrating data from remotely sensed images of tree height and forest fragmentation. They analyzed data from the Landsat satellites combined with data from the Geoscience Laser Altimeter system (GLAS) lidar instrument on the Ice, Cloud, and Land Elevation Satellite (ICESat) and high-resolution imagery from Google Earth to determine forest biomass (see Chapter 15). Their results are alarming. They reported that $0.59 \times 10^{15}$g C is lost from natural tropical forests each year, plus $0.43 \times 10^{15}$ g C/year is released to the atmosphere from managed land use (plantations, agroforestry systems, and subsistence agriculture). That is a total addition of 1.02 gigatons (Gt, or $10^{15}$ g) of carbon per year from tropical forests and managed lands. Forest losses from Latin America contributed about half of these land use change emissions. Total global tropical forest losses accounted for 11% of global $CO_2$ emissions in 2012. These changes have only accelerated by 2020, particularly in forests in the Brazilian Amazon. About one-quarter of the $CO_2$ emissions from tropical land use change are the result of forest degradation, including cutting trees for timber and fuel, and from burning forests for agriculture and from forest fires (caused in part by climate change), while the remaining 75% is from deforestation (Pearson et al. 2017). Tropical forest clear-cutting has been carried out both by people living in tropical countries and by corporations owned by people in developed countries. Chinese, Japanese, European, and North American corporations or governments are all active in tropical deforestation. So, when advanced industrialized countries report that they have no net emissions from forest loss, this can be misleading.

Developed countries have cleared large areas of their own forested land in the past, and continue to do so now. An example is the temperate rainforests of the northwestern United States and British Columbia in Canada, which have experienced large-scale deforestation and clear-cut logging for decades (see Figure 18.20). Where these old-growth forests are replanted as tree farms or undergo secondary succession, the young trees will once again serve as sinks for $CO_2$, although the removal of old-growth forests can have severe negative impacts on biodiversity as well as other ecological consequences. One important reason why forest destruction is concentrated in developing countries today is that forests in the wealthy industrial countries were altered or destroyed long ago. Both Europe and Japan, for example, have very few areas that are not heavily affected by human use. In Japan, virtually all lowland areas are either farmed or occupied by human structures; there are extensive forests on the mountainsides, but most of these are regularly logged. In the Swiss Alps, livestock are grazed even on some of the highest peaks. In Spain, the highest elevations contain some of the greatest plant species diversity, not because those areas are inherently more species rich, but because more productive lowland areas are either urban or heavily used for growing crops and for grazing domestic animals.

## 16.6 Agriculture Is a Major Source of Greenhouse Gases

Agriculture, deforestation, and land use change account for roughly a quarter to a third of the GHG emissions, both globally and in the U.S. Agricultural $CO_2$ emissions result from mechanized equipment and other sources. Carbon is released by burning crop residues—a huge source of air pollution in India and other countries. Soil cultivation can greatly speed up the decomposition of organic matter (see Chapter 5), increasing both $CO_2$ and $N_2O$ emissions, while no-till agriculture can reduce these emissions. $CH_4$ comes from several sources. When rice is grown in flooded fields (traditional paddy agriculture, see Figure 4.21), bacteria break down organic matter under water in anoxic conditions, releasing $CH_4$. Globally, rice production releases approximately $4.5 \times 10^{14}$ g annually, or about 20% of the global $CH_4$ production. The breakdown of manure produced by livestock can be a large source of $CH_4$ and $N_2O$ emissions, depending on how it is managed. Because they are ruminants and their digestion involves bacterial fermentation, cattle belch $CH_4$, accounting for about a third of U.S. agricultural emissions.

## 16.7 Global Climate Change Is Already Occurring

In September 2018, Hurricane Florence barreled into the Carolinas on the southeast coast of the U.S. Over a million people were evacuated, lives were lost, and extensive property damage resulted. The Atlantic and Gulf of Mexico coasts of the U.S. have always experienced hurricanes, but their intensity is increasing. Anthropogenic climate change contributed a 50% increase in the amount of rainfall and a great increase in the sheer physical extent of this hurricane. The storm surge, exacerbated by climate-change-driven sea level rise, was more than 3 m in some places. The greatest damage was from the approximately $6.8 \times 10^{13}$ liters of water dumped by the storm, causing severe flooding on the coast and inland. Human-caused climate change has also been shown to contribute to the severity of other hurricanes (Risser and Wehner 2017; Patricola and Wehner 2018). The previous year, Hurricane Harvey pounded southern Texas with 1.5 m of rainfall, causing catastrophic flooding, the deaths of dozens of people, and $125 billion in damage. The increase in storm intensity and in extreme rainfall is at least in part due to warmer ocean water; the energy of warm water is what drives hurricanes in general, and the more heat energy, the greater the effects. This kind of intense rainfall—a major consequence of human-caused climate change being repeated by storm after storm in many different parts of the Earth—has major effects on ecological systems and erosion and huge effects on people. Environmental damage from these and other storms includes not only the direct effects of flooding, but massive contamination from flooding of toxic waste disposal sites, waste from fossil fuel operations, vast quantities of waste pooled from pig farms, and other widespread and highly toxic pollution.

At the other extreme, drought and extended heat have contributed to unprecedented frequency, severity, and extent of fires in the western U.S. and other parts of the world. Fires are happening in forests at high elevations in Alaska and Canada that previously never experienced them. Fire seasons have extended long past their historical durations in regions where fire is a regular event. While the changes being experienced are complex and vary among ecosystems, increases in temperatures and more extended dry periods have resulted in increases in the probability of fires, their intensity, and the extent burned. In the summer of late 2019 to early 2020 in Australia, massive, catastrophic wildfires consumed millions of hectares of forest, woodlands, shrublands, and human habitation, killing people and enormous numbers of kangaroos, koalas and other native animals, and resulted in the evacuation of thousands of people. The air quality from such massive wildfires, even at great distances, reached dangerous levels for extended periods of time. Predictions are for even greater increases in fires in Australia, California, and other places. By mid-century, years without large fires in the western U.S. will become increasingly rare. The consequences of climate change are no longer restricted to the far future.

What is the evidence that the Earth is warming already and that climate changes have resulted from the rising concentrations of greenhouse gases? Weather and climate are complex phenomena, and natural variation can mask general trends. For example, in March 2018, four major snowstorms blanketed the northeast coast of the U.S. from New England to Philadelphia, with up to 90 cm of snow in places, loss of power to a million people, and several deaths. Does this mean the Earth is not warming? In a word, no. As already mentioned, warmer air holds more moisture, so a warmer winter can also be a snowier winter. Changes in Earth's climate occurring due to human alteration of Earth's energy balance (see Figure 16.2) result in greater heat retention in our atmosphere. Natural variation in weather contributes only a fraction of the changes in climate that we have already been seeing since the mid-twentieth century. The cumulative actions of people are pushing extreme weather events beyond that natural variability (Herring et al. 2018). Although considerable political debate on human-driven climate change exists (mainly in the U.S.), there is an overwhelming scientific consensus on the evidence (IPCC). The greatest uncertainty is not in the scientific models, but in what actions people will take. The sooner and more decisively we act, the lower the future damage, risks, and costs to people and the environment.

There are many different lines of evidence that have been amassed on human-induced global climate change. The average temperature of the Earth's surface has increased by about 1.1°C over the last 100 years, with about three-quarters of this change (0.6°C) occurring since 1975. Temperature increases are accelerating, with the amount of warming increasing each decade (Figure 16.31). July 2019 was the hottest month ever recorded globally, and Arctic sea ice reached record lows. The first two decades of the twenty-first century rank as the warmest in the 137-year period since global temperatures began to be reliably recorded, and the 5 warmest years ever recorded are (from the hottest) 2016, 2015 and 2019, 2017, and 2018. Summer temperatures in the past few decades in the Northern Hemisphere have been the warmest in the past 1000 years, based on analyses of tree rings and ice cores. Some regions have experienced much greater warming, while others have not warmed as much or have even cooled, but the Earth as a whole is warming steeply. High temperature extremes in the U.S. and other countries have increased dramatically.

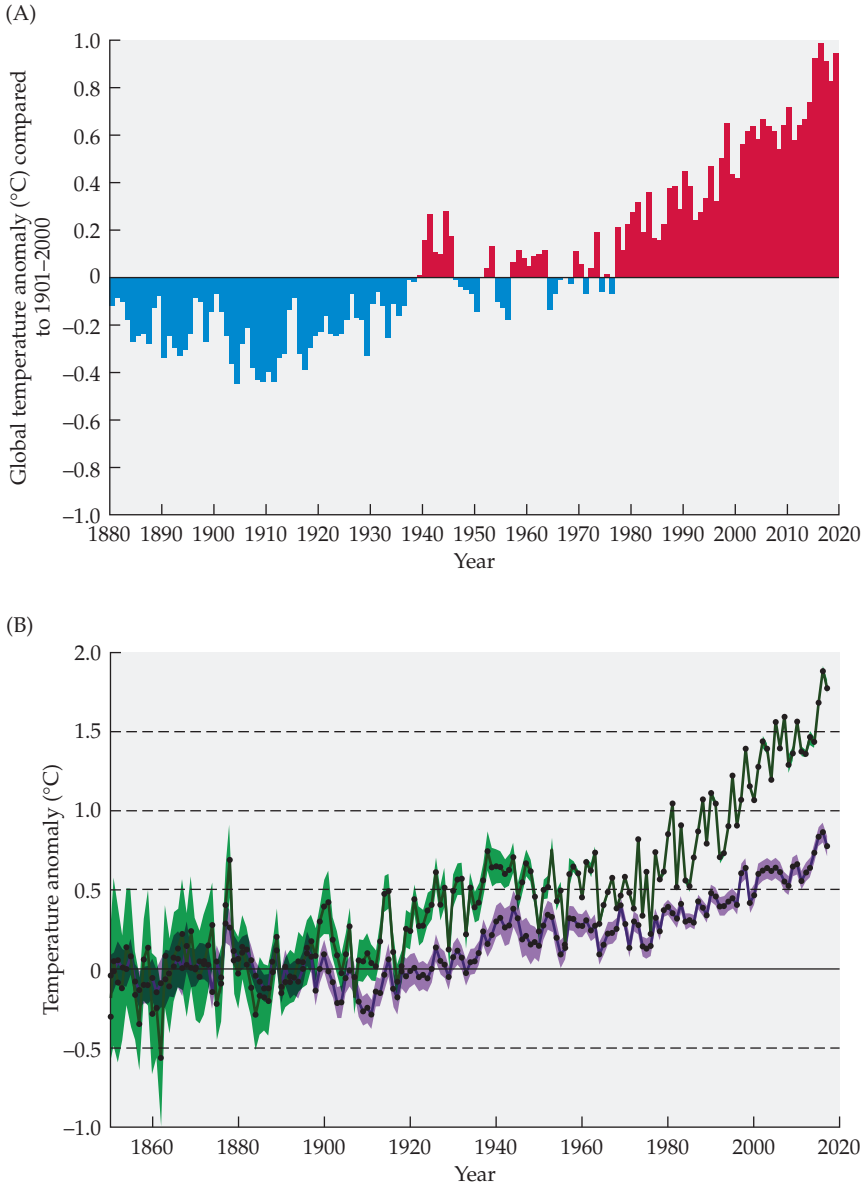


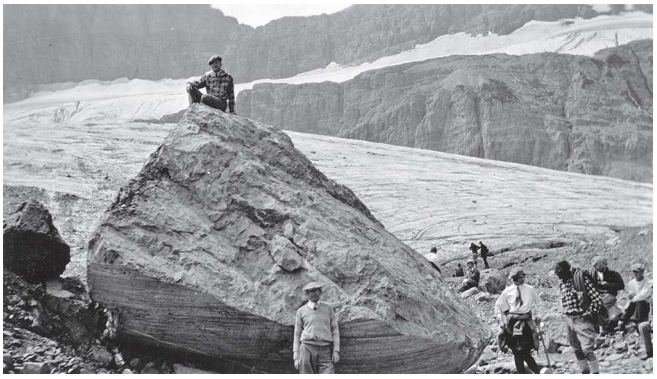

**Figure 16.31** (A) Global temperature anomalies for the entire Earth (land and ocean) from 1880 to 2020, as average annual °C difference from the 1901–2000 average (NOAA). (B) Global temperature anomalies for the land (green) and oceans (violet) as average annual °C difference from the 1850–1900 average. The land is warming much more rapidly than the oceans. Shading indicates confidence limits, and dots indicate means. (A after NOAA NCEI (National Centers for Environmental Information). 2020. *Climate at a Glance: Global Time Series*. https://www.ncdc.noaa.gov/cag/; B After Berkeley Earth. 2018. *Global Temperature Report for 2017*. http://berkeleyearth.org/global-temperatures-2017/.)

The greatest warming has occurred in northern latitudes (between 40°N and 70°N across Eurasia and North America), especially in Alaska, Canada, Siberia, Scandinavia, and the Antarctic Peninsula. The oceans have warmed more slowly than the land, because of water's higher heat capacity, and the Northern Hemisphere has warmed more rapidly than the Southern Hemisphere, in part because there is far more land area in the Northern Hemisphere and more ocean in the Southern Hemisphere. Nighttime temperatures are increasing more than daytime temperatures.

Sea level rise is a consequence of rising temperatures globally, and it also provides indirect evidence of global warming. Over the last century, average sea levels have risen between 13 and 20 cm on different coastlines, with much of that rise occurring from 1995 to 2019 (see below). The rate of sea level rise is accelerating, from about 1.5 mm/year in the twentieth century to about 3.4 mm/year currently. Historically, sea levels varied; since the last glacial maximum around 20,000 years ago, sea levels rose by about 120 m as the glaciers covering much of Europe, North America, and other parts of the world melted (see Figure 17.11 and Figure 16.34). However, for most of the past 2000 years sea levels had been more or less constant. That changed about a century ago. In the twentieth century, sea level rise occurred due to the thermal expansion of water (warmer water occupies more volume than colder water), but now the melting of continental ice is increasingly contributing to sea level rise. Melting sea ice such as the ice surrounding the North Pole does not contribute to rising sea levels (because the ice displaces the water it is floating on), and mountain glaciers do not hold enough water to change sea levels by much (although they have huge effects in terrestrial systems).

The retreat of glaciers worldwide is another consequence of warming temperatures globally, and also provides indirect evidence of the cumulative effects of rising temperatures. Montane glaciers are thinning, shrinking, and disappearing all over the world, from Glacier National Park in Montana, U.S.A. (down from 35 glaciers to 26, Figure 16.32), to the glaciers in the Andes (Bolivia's glaciers have declined by 40% in recent decades), the Alps in Europe, the Himalayas, the South Island of New Zealand,

(A)



Courtesy of Archives & Special Collections, Mansfield Library, University of Montana

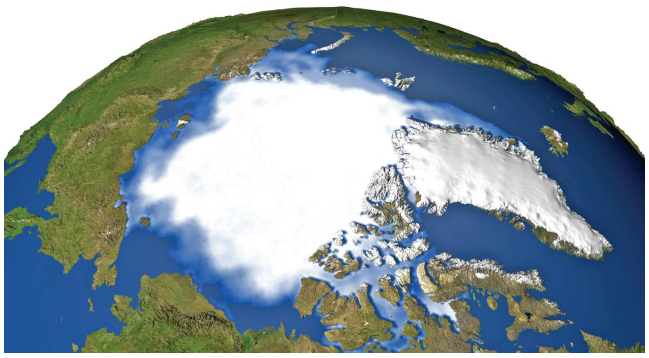
(B)

Courtesy of Lisa McKeon/USGS

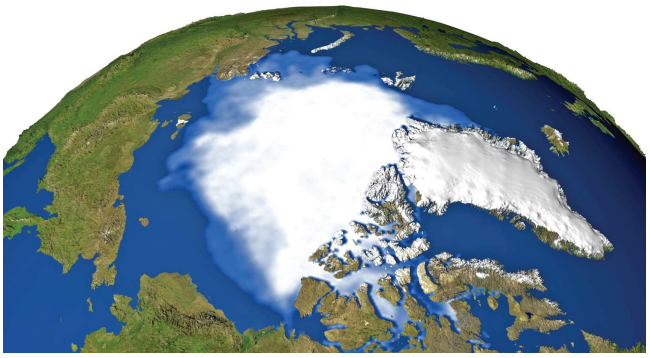
**Figure 16.32** A rapidly melting and retreating glacier on Chapman Peak in the northwest region of Glacier National Park, Montana, U.S.A. in photos taken at the same spot (A) circa 1910 and (B) 2007.

Mt. Kilimanjaro in Tanzania, and elsewhere. Arctic sea ice has shrunk by 40% on average since 1979, when measurements began (Figure 16.33). Permafrost in continental areas of the Arctic is also melting. This is particularly alarming for the huge glaciers covering much of Greenland. The changes in the Antarctic are more complex; sea ice is breaking up and becoming rearranged, but it is unclear whether its total extent is declining. However, there has been an accelerating loss of land ice in Antarctica, particularly in West Antarctica. This loss is not due to melting at the surface, as summer temperatures in Antarctica are mostly below freezing, but instead due to warmer ocean water melting the land ice around its edges, increasing the movement of melting ice to the ocean.

(A) Sea ice minimum 1979



(B) Sea ice minimum 2005



(C) Sea ice minimum 2019

**Figure 16.33** There has been a dramatic decline in the extent of Arctic sea ice between (A) 1979, (B) 2005 and (C) 2019. These satellite images show the ice in September, when it is at its minimum for the year. (From NASA Scientific Visualization Studio; https://climate.nasa.gov/interactives/climate-time-machine.)

## 16.8 Large Changes Are Predicted for Earth's Climates, but Some Impacts Can Still Be Mitigated

What kinds of changes are predicted for the future, and how will they impact natural systems and the humans that live in and depend on these natural systems? One common reaction to predictions of global warming by people in regions with cold winters is that warmer winter days would actually be pleasant and welcome. What could be wrong with getting to play ball or wash your car outside in the middle of winter? To answer that question, we need to consider more comprehensively what global climate change may entail. We first look at the predicted changes, and in the next section, we consider the results of these changes.

Climate scientists use a variety of direct and indirect measurements to document climate changes. Predictions about what future climates might be like are made by constructing and refining complex computer models (Box 16C). Various kinds of uncertainties are associated with these predictions, but the greatest uncertainty is in knowing what people will do—human actions in the present and near future will have enormous effects on how things turn out.

Warmer temperatures are driving large changes in the global water cycle, as water evaporates more readily from soils and the ocean, and more water vapor is contained in the atmosphere. The additional energy added to the ocean and atmosphere intensifies storms and rainfall. Uncertainties exist about consequences for changes in the jet stream, climate cycles, and oceanic circulation

### BOX 16C Understanding Past Climates and Predicting Future Climates

How do scientists know what the values for temperatures, precipitation, and other weather and climatic variables were in the past? How do they figure out what they were in the distant past? There are direct records of these variables—measurements that have been taken and recorded around the world—as well as indirect records of various kinds. Reasonably good instrumental records have been kept in many places for temperature, rainfall and snowfall, humidity, and air pressure at the Earth's surface for at least 100 years, but some locations are missing data entirely or for earlier years. For recent years, satellites and high-atmosphere weather balloons provide accurate measurements. There are also excellent recent data for global ocean temperatures, from satellites and various kinds of temperature probes placed in many areas.

Physical, biological, and chemical indicators offer insight into climates and weather further back in time. These indicators include measurements of tree rings from trees that are hundreds or thousands of years old, sediments at the bottoms of lakes and oceans, air trapped in ancient ice cores, and annual density bands in long-lived corals.

How do scientists make predictions about future climates? Atmospheric scientists use complex computer simulation models, called general circulation models or global climate models (GCMs), that incorporate existing weather and climatic data, as well as what is known about the major processes that determine climate, to predict climate change. Once a model is constructed, changes in some factor—such as $CO_2$ levels in the atmosphere—can be introduced to see how the climate might be affected.

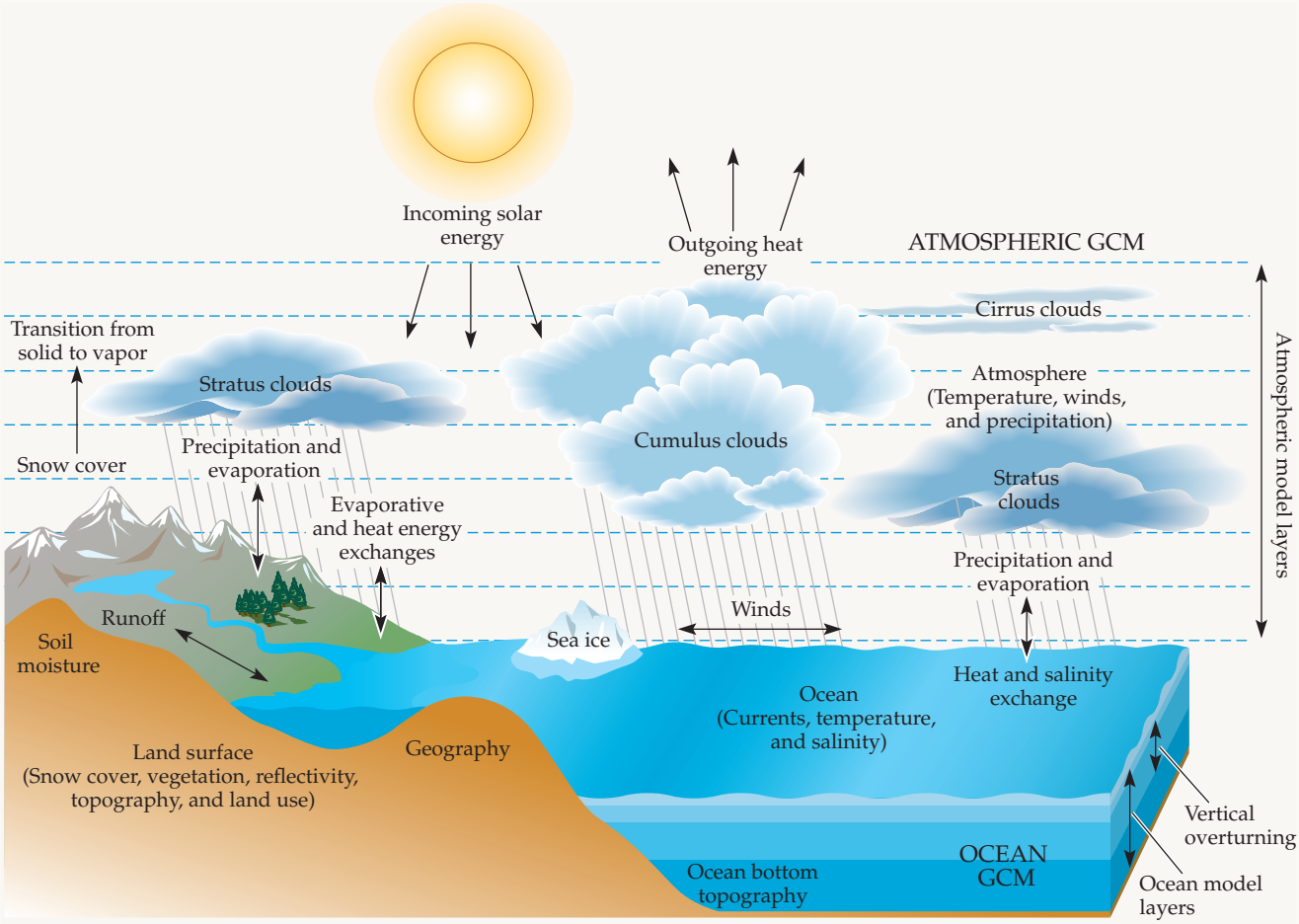
GCMs take into account detailed data on glaciers, land, ocean and atmosphere, including vegetation, incorporating lags and feedbacks between model elements.

Many scientists must work together to construct a GCM, and only a small number of major GCMs exist. These models incorporate large amounts of known data and make predictions about future temperatures, rainfall, snow cover, sea level, and other climatic factors. Recent research in this area has included efforts to incorporate more accurate satellite data on the Earth's radiation balance into GCMs (Wielicki et al. 2002). These efforts have revealed greater variation in radiation fluxes between the Earth's surface and the atmosphere than was previously thought to exist, particularly in the tropics. Differences in cloudiness appear to account for this variation; clouds both increase reflection of incoming solar radiation (a cooling effect) and reduce infrared radiation to space (a warming effect). A complex set of feedbacks between the Earth's surface and the atmosphere can affect the degree of cloudiness.

Predictions based on GCMs are uncertain due to limitations in what is known as well as uncertainty about what human actions will be taken (e.g., whether $CO_2$ emissions will be reduced). They are tested in a number of ways, including 'backcasting' by starting from historic conditions to see how well models 'predict' present conditions. These models are being continually refined and improved, and they are becoming increasingly accurate and detailed. They offer a reasonable set of predictions that will help us to anticipate the effects of global climate change at global and regional scales.

patterns. Two predicted changes of particular interest are an increase in storm severity and extreme weather, and an increase in ENSO events. Both changes are linked to the same cause: an increase in ocean temperatures. Warmer ocean water provides more energy to feed storm strength, intensifying rainfall events. Averaged across the entire globe, by 2050 temperatures are predicted to rise by approximately 2°C to 2.5°C. Longer-term predictions are more uncertain, but by the end of the twenty-first century, temperature increases may range from 1.5°C to 4.5°C. In comparison, global temperatures have risen by less than 6°C since the most recent glacial maximum (ice age) 20,000 years ago (see Figure 17.10). The average temperature over the year is only 3°C higher in Milan, Italy, than in London, England, and 4°C higher in Charlotte, North Carolina, than in New York City. What might seem to be a small average temperature increase can mean a very different climate. By 2050, cities in Europe will see a rise of summer temperatures of 3.5°C. Most cities globally will experience dramatic climate change—in both temperature and associated weather events like flooding or severe air pollution from fires—while over a fifth of them will suffer from conditions that currently do not exist for any city (Bastin et al. 2019).

Temperature increases will not be evenly spread across the globe. Land areas will continue to warm much more than oceans, and areas in the far north will continue to warm much more than those to the south. Summer warming is expected to be particularly great in central and northern Asia. More hot days and periods of extended intense heat are predicted over all land areas, and the average humidity is expected to increase. Cold




Some of the elements in a global climate model (GCM). The model shown incorporates atmospheric, terrestrial and oceanic components. (After U.S. Global Change Research Program and Climate Change Research Initiative. *Our Changing Planet: The Fiscal Year 2003. A Supplement to the President's Fiscal Year 2003 Budget*. Global Change Research Information Office [GCRIO]: Palisades, NY.)

waves will become fewer and less intense. Glaciers and ice caps will continue their dramatic retreats (see Figure 16.32). Jonathan Overpeck and associates (2005) predict the total loss of summertime sea ice in the Arctic in less than 100 years. This would be the first time in 800,000 years that the northern polar ice cap has completely disappeared in summer. Sea ice reflects more solar radiation back to space than dark ocean water, which can absorb more of the sun's energy. But reduced sea ice coverage could also lead to more heat loss in winter from the open ocean than would be lost from ice, so the feedbacks resulting from loss of Arctic sea ice are complex and uncertain.

Reductions in the density of Arctic water resulting from warming and from changes in salinity due to melting glaciers and sea ice could potentially induce changes in thermohaline circulation (THC) patterns (see Figure 16.13), although there is considerable uncertainty about this prediction. If that happens, it could lead to reductions in the THC transfer of heat to northern and western Europe (Jacob et al. 2005; O'Hare et al. 2005), leading to much colder temperatures there.

Global climate models predict large changes in rainfall patterns. Substantial increases in yearly rainfall are expected in the Arctic and subarctic regions. Winter precipitation is predicted to increase in tropical Africa and at mid-latitudes in the Northern Hemisphere, and summer rainfall is predicted to increase in southern and eastern Asia. Drought will become more prominent in Australia, Central America, and southern Africa, particularly in winter. Midcontinental areas will generally become drier, which does not augur well for the world's breadbasket regions (see Chapter 18). Temporal patterns are expected to shift as well. Everywhere year-to-year rainfall will become more variable, with increased flooding in some years and drought in others.

Inland areas are vulnerable to flooding from heavy rainfall events overflowing rivers and streams, but the greatest implications for natural systems and human societies are effects on coastlines from an increase in sea levels. Sea levels are estimated to rise between 30 and 120 cm by 2050, with a possible increase in mean sea levels of between 30 and 250 cm by 2100 (Figure 16.34). There is some possibility of truly catastrophic changes in sea level; the complete melting of the ice sheets of Greenland and Antarctica would raise ocean levels by as much as 100 m. Many current coastal areas may be permanently under water. Higher storm surges make coastlines more vulnerable, because hurricanes push large amounts of water toward coasts. When that water hits shallow nearshore

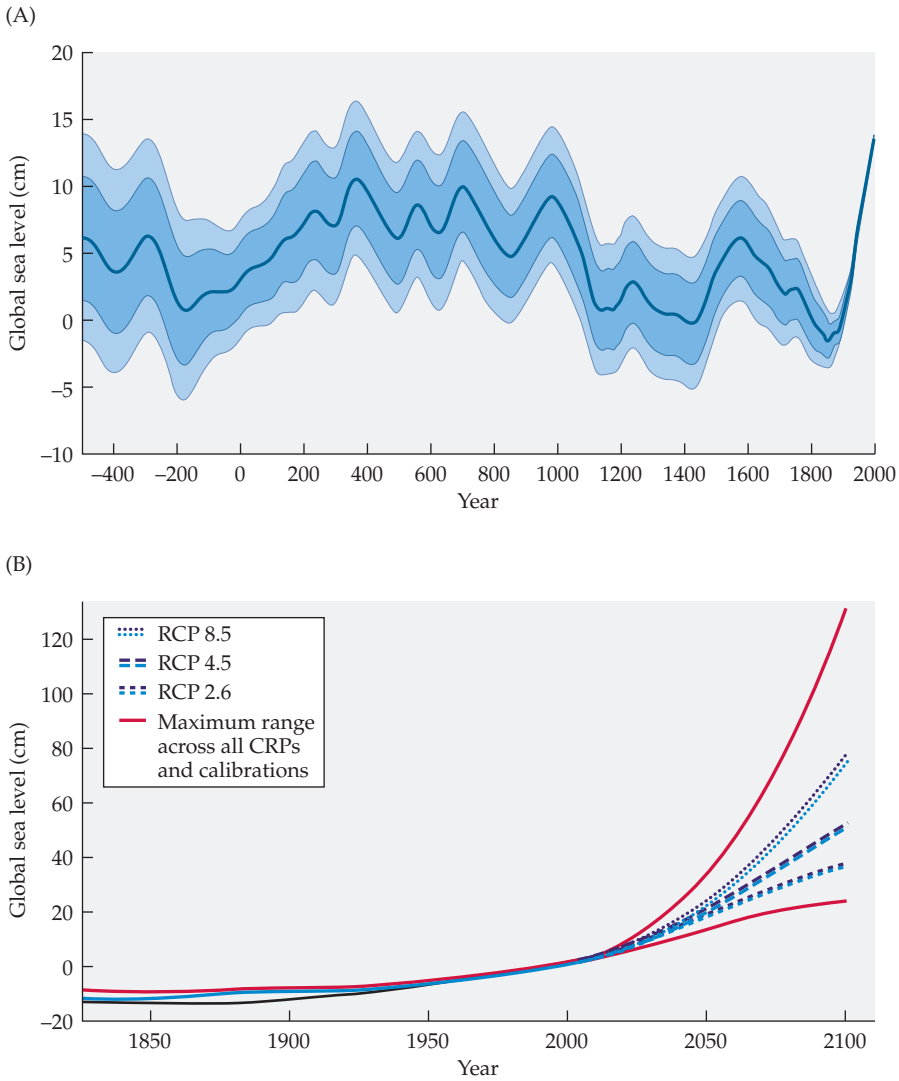



**Figure 16.34** (A) Observed and reconstructed mean sea level for the last 2500 years. (B) Projected mean sea level for six future scenarios based on three climate models; RCP 8.5 represents continuation of current $CO_2$ emissions, RCP 4.5 represents climate forcing due to a moderate reduction in emissions, and RCP 2.6 represents severe reductions in emissions. The red lines indicate the fifth percentile of RCP 2.6 and 95th percentile of RCP 8.5, corresponding to the lower bounds and upper bounds of the lowest and highest model estimates. (After R. E. Kopp et al. 2016. *Proc Natl Acad Sci USA* 113: E1434–E1441.)

water, as found along the Atlantic and Gulf coasts of the United States, there is a storm surge—a large temporary increase in the local sea level. Because in the eastern U.S. (as in many other parts of the world) the land rises from the sea very gradually in a wide coastal plain, flooding due to storm surges can penetrate far inland. A rise in sea level can compound the damage due to storm surges, because adding even 50 cm to a storm surge will move floodwaters substantially farther inland. Because human populations are increasingly concentrated along the world's coastlines, they are more vulnerable to these effects (Emanuel 2005). The consequences of sea level rise are already being felt, from low-lying island nations such as the Maldives, to highly populated coastal regions in Bangladesh, China, and the U.S. Some of the most vulnerable areas include the Pearl River Delta in China (120 million people), Alexandria and the Nile Delta of Egypt (8 million), and southern Florida, U.S. (up to 7 million). By 2050, 27 million people in Bangladesh may be at risk or displaced due to sea level rise. Coastal ecosystems, including temperate salt marshes, tropical mangrove ecosystems, and ecologically rich estuaries around the world, risk permanent inundation. The risk to natural systems is shared by humans.

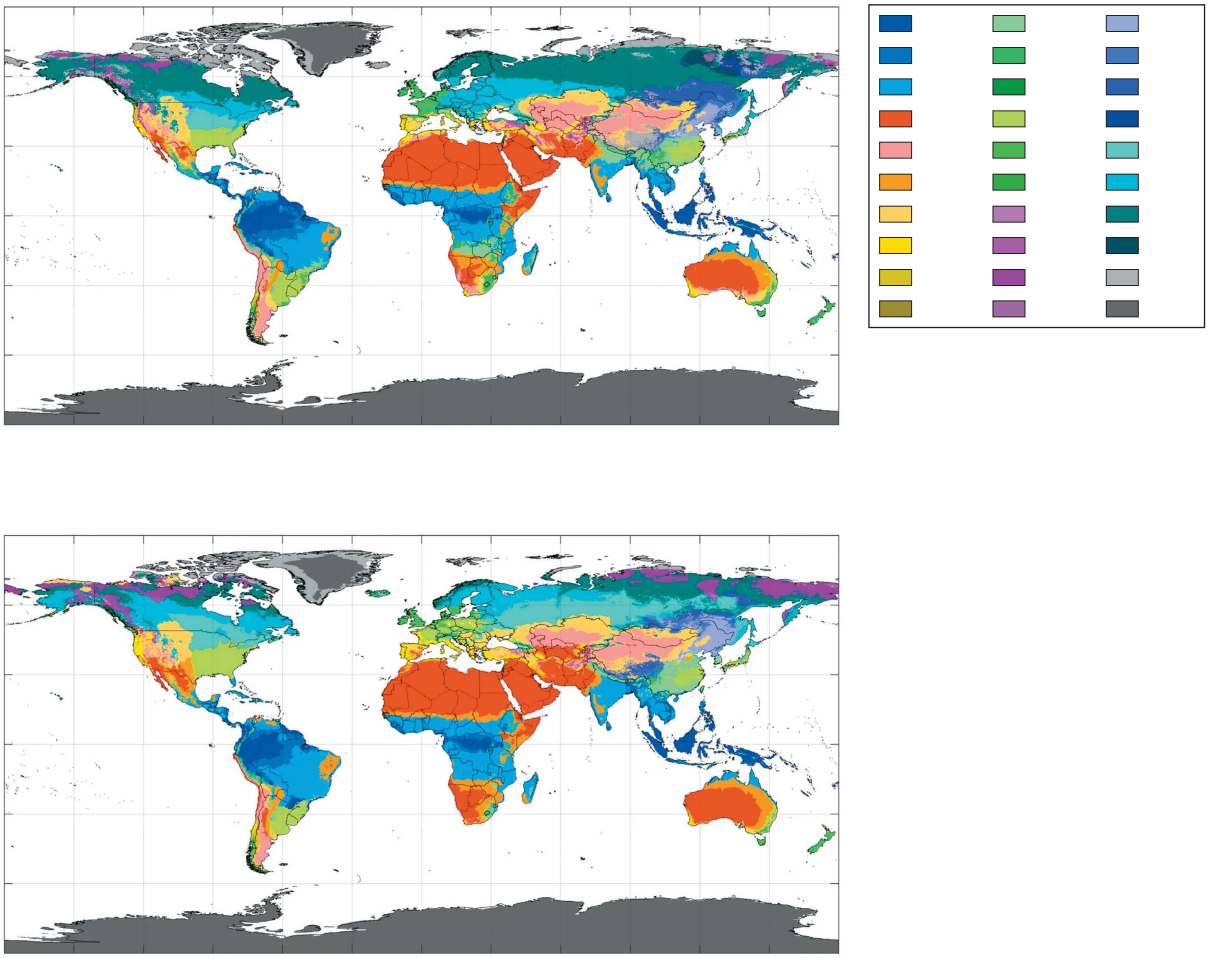
## 16.9 Changing Climates Are Affecting Species and Ecological Systems

Regional and global changes in climate are likely to have large and only partially predictable effects on plants. Some species, ecosystems, and vegetation types will benefit from these changes and increase, while others will lose ground, declining or disappearing (Figure 16.35). So what are these places that we stand to lose? The Sundarbans is the largest tropical mangrove forest in the world (over 4200 $km^2$), is a UNESCO World Heritage site, and is part of the World Network of Biosphere Reserves. It is located on the Bay of Bengal in eastern India and southern Bangladesh at the delta of three major rivers. The Sundarbans is rapidly getting eaten away due to rising sea levels, and pollution and land use changes have reduced freshwater flow such that sediment replenishment no longer keeps up with losses due to sea level rise. Changes in temperatures and rainfall, particularly during the monsoon, may be stressing the organisms. This area is named for the sundari mangrove trees, which include many species such as *Excoecaria agallocha* (genwa, Euphorbiaceae), *Aegiceras corniculatum* (khalsi, Primulaceae), *Bruguiera gymnorrhiza* (kankra, Rhizophoraceae), and *Xylocarpus granatum* (dhundal, Meliaceae). The Sundarbans is home to many species of birds, Bengal tigers, chitral deer, the fishing cat, and a wealth of reptiles, including several species of crocodiles, giant water monitors, king cobras, rock pythons, Russell's vipers, dog-faced water snakes, chameleons, salvator lizards, and mouse geckos, and many species of amphibians. The waters are home to olive ridley sea turtles, green turtles, Ganges dolphins, Indo-Pacific hump-backed dolphins, Irrawaddy dolphins, and finless porpoises as well as a diversity of fishes and molluscs. This ecological system, rich in many different kinds of species and protected as a biosphere reserve, is being diminished and threatened by climate change and other human actions beyond the borders of the reserve.

Globally, plant growth, reproduction, and nutrient allocation will all be affected by changing climates. Interactions of plants with other species—competitors, pollinators, herbivores, and pathogens—will also change. For example, plants that are dependent on mutualistic interactions with pollinators or dispersers will suffer if environmental changes reduce or eliminate those animals. Climate change is predicted to alter ecosystem cycling for many major nutrients. Soil respiration and nitrogen mineralization rates, for example, have been shown to increase substantially when ecosystems are warmed experimentally (Rustad et al. 2001). The distributions of many plant species are already changing and will continue to change, resulting both in shifts in the locations of plant communities and in different combinations of species growing together. In a study of 1700 species of plants and animals, Camille Parmesan and Gary Yohe (2003) reported significant range shifts poleward across all species, and a speeding up of spring phenology. In a meta-analysis of 143 studies on the responses of plants and animals to global climate change, Terry Root and colleagues (2003) concluded that there were already significant, measurable effects of climate change. A study of flowering dates for 385 British plant species showed that 16% began flowering significantly earlier in the 1990s than in the previous four decades (Fitter and Fitter 2002).

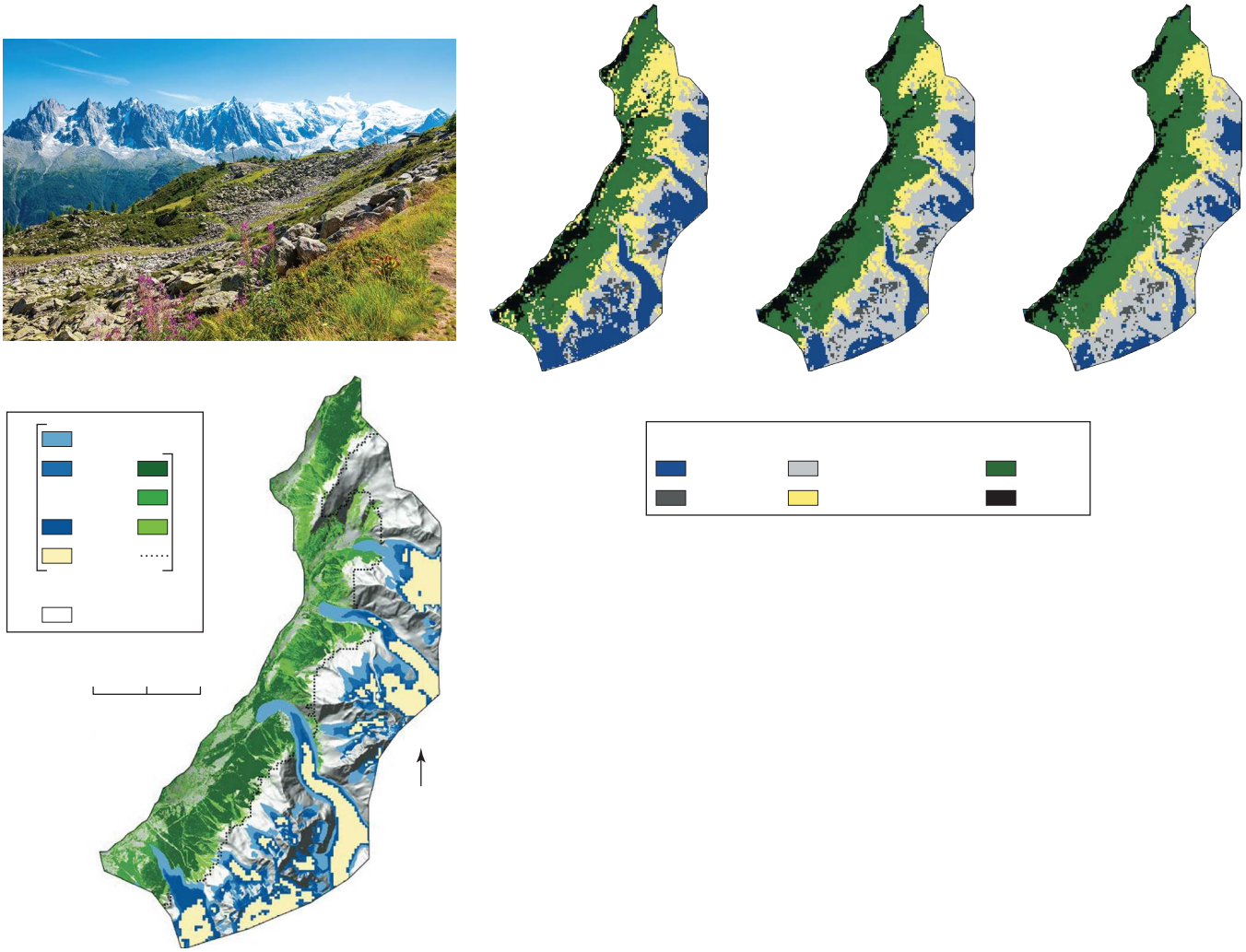
The increases in productivity resulting from higher $CO_2$ levels may be amplified in some ecosystems by warmer climates and longer growing seasons. Other ecosystems will be affected by increased drought, canceling the positive effects of higher $CO_2$ levels on photosynthesis and growth. Changes in temperatures and rainfall regimes are expected to have major consequences for agricultural regions. Some highly productive areas are predicted to become unproductive, while some regions that are currently marginal agriculturally will become highly productive. The crops that are grown in particular areas may change substantially, with grain belts shifting and frost-sensitive crops being grown at higher latitudes.

As we have seen, atmospheric $CO_2$ levels have fluctuated widely over the course of the Earth's history (see Figure 17.7), and global and regional climates have fluctuated widely as well (see Figure 17.10). Biomes and the species inhabiting them have changed dramatically



over the eons. Many plant species have moved as climate has changed. After the most recent glaciation in North America, for example, species moved northward to recolonize areas that emerged from beneath the glaciers as temperatures warmed (see Figure 17.14). The climate changes of the past and those occurring now have two big differences. First, the current changes are happening at much faster rates, far greater than the rates at which many species can migrate or adapt. Second, the changes are far greater in magnitude than have ever been experienced by humans and many other organisms, as some aspects of climate are being altered beyond what has been experienced on the Earth in thousands or hundreds of thousands of years.

Habitat reduction and fragmentation are likely to make it more difficult for species to migrate in response to climate change. Parks and preserves that currently protect populations, species, or communities may become unsuitable refuges. If species are unable to reach new favorable habitat, or if habitat in newly suitable



environments is not protected, those species will disappear. Species will not move together, but migrate (or fail to migrate) individually (see Figure 17.14). As a result, many vegetation and community types may disappear and be replaced with new species combinations that have new properties.

Arctic tundra and boreal forest species, facing the greatest degree of climate change, may find themselves with no place to go as warming shrinks the extent of suitable habitat. Alpine species forced to move to higher altitudes may face similar problems. Many plant species are reported to be moving to higher elevations in mountainous areas, including trees that are becoming established above current tree lines. In Europe, current plant species richness on 30 peaks in the Alps has increased by 70% over the oldest accurate historical records, apparently as a result of colonization by lower-elevation species. Likewise, butterflies, birds, and other species that interact closely with plants have been moving upward in elevation and poleward (reviewed by Hughes 2000). Einar Heegaard and Vigdis Vandvik (2004) inferred that an important factor determining the community composition of alpine vegetation in Norway was the effect of time of snowmelt on the outcome of competitive interactions among plant species. They predict that changes in the timing of snowmelt will alter dominance patterns and that competitive exclusion, rather than simple intolerance of warmer conditions, will eliminate many of the alpine species adapted to conditions at the highest altitudes.

As elevation increases on mountains, available land area declines, and this reduction in habitat combined with warming forcing species upward can potentially result in declining populations of montane species, and the risk of extinctions. Bradley Carlson and his colleagues (2014) predicted changes in distribution for 36 high-elevation forest, meadow and alpine heath species in the Chamonix Valley (Figure 16.36A) and surrounding alpine areas in France. Using aerial photographs from 1952, 1979 and 2008, they incorporated information on rising tree line, glacial retreat and primary succession in species distribution models and predicted future

distributions. Glaciers covered 30% of the study area in 1850, but had declined to 19% of the area in 2008; forest cover increased from 17% in 1952 to 26% of the study area in 2008. Forest area was predicted to increase to over 30% of the area by 2051–2080, while the glaciers were projected to decline steeply to approximately 10% of the area by that time (Figure 16.36B). Habitat for the meadow and heath alpine species was predicted to decline by 2012–2050, and to decline further from 2015–2080 (Figure 16.36C, D, E). However, this decline was not as steep as might be expected, because primary succession in the area left by the retreating glacier was predicted to allow colonization by these species. The forests were predicted to continue to expand in future years in their model. Of course, once the glaciers have disappeared, further climate change is likely to result in reductions in habitat for the high elevation tree species as well as for the alpine meadow and heath plants.

Range shifts, range contractions and expansions, and further changes in phenological patterns are likely to occur as changes continue to take place not only in average temperatures, but also in the length of growing seasons, the timing of snowmelt or even the existence of snowpack, the duration of drought and extreme heat events, fire regimes, and soil moisture availability at different times of the year. Furthermore, the ecology of organisms at all levels, from populations to ecosystems, will be affected by changes in many of these factors at the same time. Due to the combined effects of multiple stresses, species currently threatened or endangered are more likely to go extinct, and additional species are likely to become endangered. Small islands close to sea level may disappear entirely as oceans rise, taking their endemic and rare plant and animal species with them (as well as the homes of the people living there). Even far from coastlines, increases in storm intensity, sometimes coupled with the effects of deforestation, are expected to result in major disturbances to plant communities by causing flooding in river valleys and landslides, mudslides, and avalanches on steep slopes.

## 16.10 Responses to Ongoing and Predicted Climate Change

Is there anything that can be done about the dramatic changes predicted for climate and life on Earth? Recognizing the enormous and catastrophic scope of already occurring and more severe predicted changes, the most effective and straightforward action would be a serious global effort to reduce and eventually eliminate the things we are doing to change our climate, most obviously the emission of GHGs. While efforts are being made in this direction, they are incomplete, tentative, and slow, with frequent reversals. Several other approaches have been proposed to complement reduction of GHG emissions by transforming dependence on fossil fuels to use of alternative energy sources.

One set of actions is ecological: protection of existing forests and afforestation, protection of soils and wetlands with high organic matter content from erosion to reduce methane emissions, and protection and restoration of shoreline marshes such as saltwater marshes and mangrove forests. Environmental protection and ecological restoration have a myriad of benefits in addition to reducing the acceleration of climate change (see Chapter 19).

A second set of actions focuses on protecting people and property in floodplains and along coastlines through land use laws and civil and environmental engineering: protecting infrastructure by zoning laws to limit construction, raising existing buildings structurally, and constructing large-scale devices designed to block storm surges from inundating coastlines, tunnels, subways, and other human infrastructure. These measures do nothing to halt or reverse climate change but seek to reduce the damage to people and their property. A major limitation is that as climate change accelerates, this becomes a never-ending battle and an increasing drain on public and private financial resources.

A third, highly controversial set of proposals is collectively called **geoengineering**, the deliberate large-scale manipulation of the planetary environment to counteract anthropogenic climate change. As Earth inexorably heats up and damage and costs become unbearable, these controversial approaches may claim wider support. There are great risks and costs associated with implementing them, but desperate times may increase the call for desperate measures. One of the most prominent approaches is called solar radiation modification (SRM). It involves the injection of small particles into the upper atmosphere to reflect sunlight back to space, increasing Earth's albedo and reducing the amount of energy from sunlight reaching Earth's surface. Detailed climate models show that this would indeed cool the planet. It would also change precipitation regimes in a complex manner, and have the potential for unintended negative effects to climate, people, and ecological systems (Trisos et al. 2018). Advocates emphasize that geoengineering should only be undertaken, if at all, in conjunction with strenuous efforts to reduce $CO_2$ emissions. It might be the only way to reduce catastrophic ecological and human losses, but entails enormous risks and uncertainties. Other geoengineering proposals include methods to slow the melting of Arctic ice and

continental glaciers, reduce the ferocity of hurricanes, and manipulate marine systems in various ways, but none of these are as well supported and well developed as solar radiation modification. A major effort has also been made to directly remove $CO_2$ from the atmosphere and sequester it in the ground or under the oceans. While this can work on a small scale, no one has been able to figure out how to do it on a large enough scale to matter, especially without insurmountable financial costs. It is totally unclear whether any of these will ever be deployed or whether solar radiation modification is worth considering even with its risks, given our current headlong rush to heat our planet and irreversibly change climates on Earth.

## Summary

- Climates are determined by the mean, variability, and seasonality of temperature and precipitation over long periods of time (decades to centuries) and can be defined from local to global extents; weather refers to local daily conditions.
- Mean temperature in an area is determined by the amount of incoming solar radiation, which varies daily, seasonally, yearly, and over the course of centuries.
- At intermediate time scales of years and decades, cycles such as the ENSO cycle have large effects on weather patterns. At time scales of tens to hundreds of thousands of years, changes in the Earth's orbit around the sun and in the tilt of the Earth's axis create long-term cycles in solar radiation and in climates on Earth.
- Patterns of precipitation are caused by differential heating and cooling of the atmosphere at different latitudes.
- The Hadley cells, in which warm air rises over the tropics to high in the atmosphere and falls to create deserts at around 30° north and south of the equator, affect wind and water currents and drive regional climates in the tropics and subtropics and affect climates globally.
- The rain shadow effect is caused by the flow of moist air across islands or continents and its interaction with mountain ranges. Areas on the windward sides of mountain ranges tend to be wet, while the leeward sides of mountain ranges (the rain shadows) tend to be dry.
- The global carbon cycle is intertwined with life on Earth and is central to Earth's climate.
- The greenhouse effect is the retention and radiation of heat energy in the atmosphere and is due to $CO_2$ and other greenhouse gases. It is natural and essential to life on Earth.
- Humans are increasing the greenhouse effect by adding enormous amounts of greenhouse gases into the atmosphere primarily from the combustion of fossil fuels, heating the planet, and drastically changing Earth's climate.
- The consequences of human-caused climate change for humans and for the natural world are far-reaching and profound.

Courtesy of S. Scheiner

17

# Paleoecology

A fundamental principle of ecology is that across space and time organisms are distributed in a heterogeneous manner (see Table 1.2, principle 1). In other chapters, we have focused on spatial patterns and the ecological and evolutionary processes that create them, including patterns of vegetation and diversity. We have also investigated processes that occur over relatively short and medium time spans, such as individual plant growth (see Chapter 6), population change (see Chapter 8), and community successional change (see Chapter 13). The world is constantly changing (see Table 1.2, principles 5 and 6), and viewed in the big picture, these chapters are "snapshots" of the world of today—a moment in time. In this chapter we focus on very long-term patterns over time and over large spatial extents. The world was a very different place at various times in the past. Today, if you stand at a spot in rural northern Illinois, you will see cornfields all around you. Just 150 years ago, the scene would have been one of rolling prairie. Earlier—20,000 years ago—you would have been standing in tundra, with sunlight glinting off the massive glaciers to your north. Go even further back, and you might have been standing in the middle of a dense forest, or in a swamp dominated by tree ferns.

In this chapter, we consider some of these long-term changes in plants and plant communities, the scientific methods used for understanding them, and some of their implications. Knowledge of the past both informs our understanding of the present by telling us how we got where we are now, and provides critical data for the management of today's ecosystems (Swetnam et al. 1999). We cannot, of course, cover the entire history of the Earth, or even of terrestrial plants, in one chapter. The ecology of many earlier periods is still poorly understood. The field of **paleoecology**—the study of historical ecology—depends heavily on inferences from and analogies with

Above: Teide National Park is contained within the caldera of Mount Teide, one of the volcanos that make up the island of Tenerife in the Canary Islands of Spain. This photo was taken at 2000 m elevation looking towards the peak of the still active volcano.

current species and communities. Our understanding of the very distant past is especially tenuous because of a dearth of fossils from those times and because many ancient species have no modern analogues. Our knowledge of the past, however, is increasing rapidly. The development of new methods, such as the use of carbon isotope ratios to identify $C_3$ and $C_4$ plants (see Chapter 2) and the creation of extensive computerized databases of the fossil record, contributed to making paleoecology an active and lively field of research.

In this chapter, we first survey several geological eras and periods (Figure 17.1), especially those that represent significant times of change for plant communities. Then we look in detail at the past 20,000 years. Because of the amount of detail involved, we focus on North America. The history of this recent period is now well known and provides special insight into current plant communities and biomes. We take up the future of Earth in Chapter 16 and Chapter 19.

## 17.1 Plants Invaded the Land in the Paleozoic Era

Plant life began in marine environments, and fossil evidence tells us that plants first invaded land in the

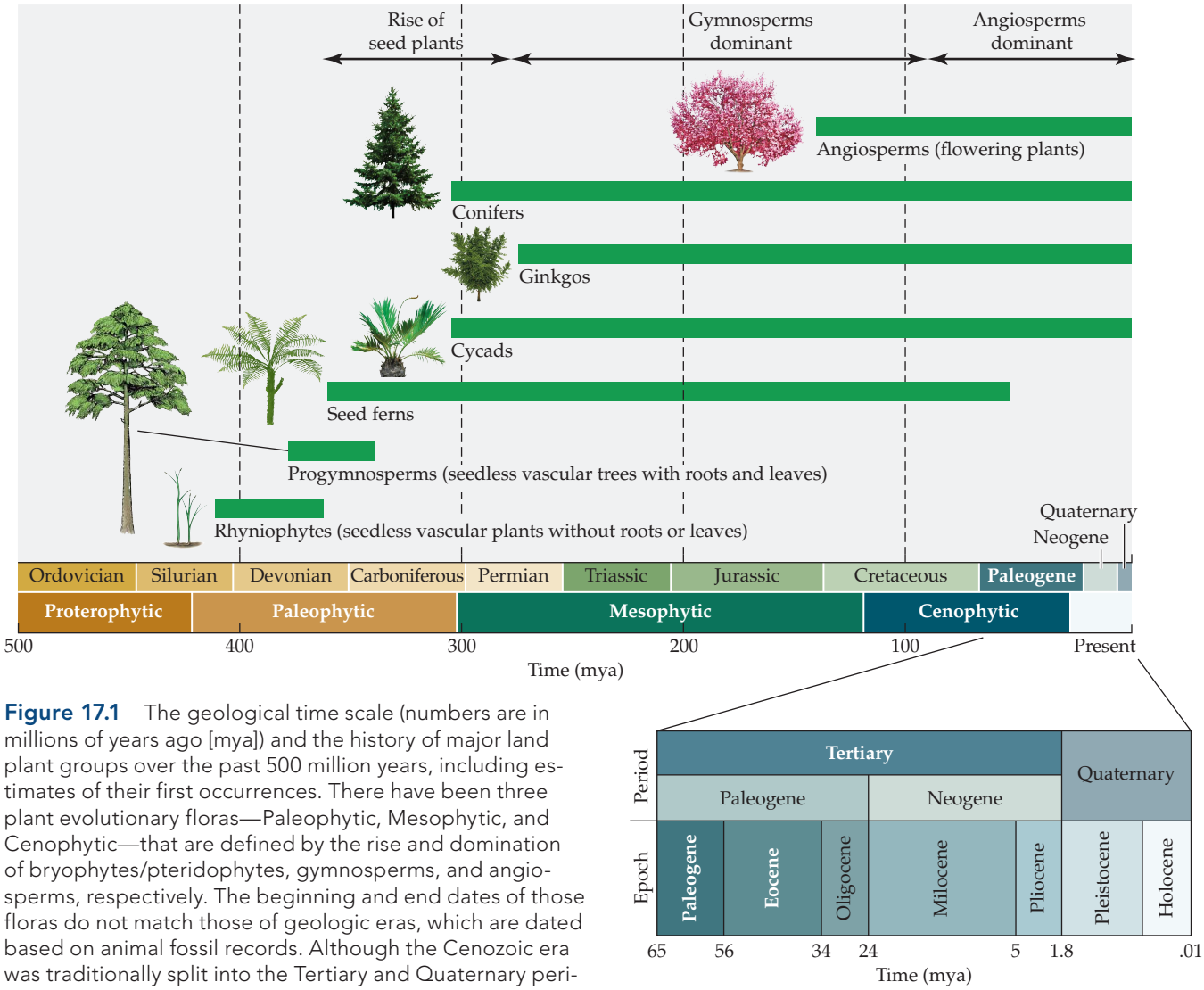
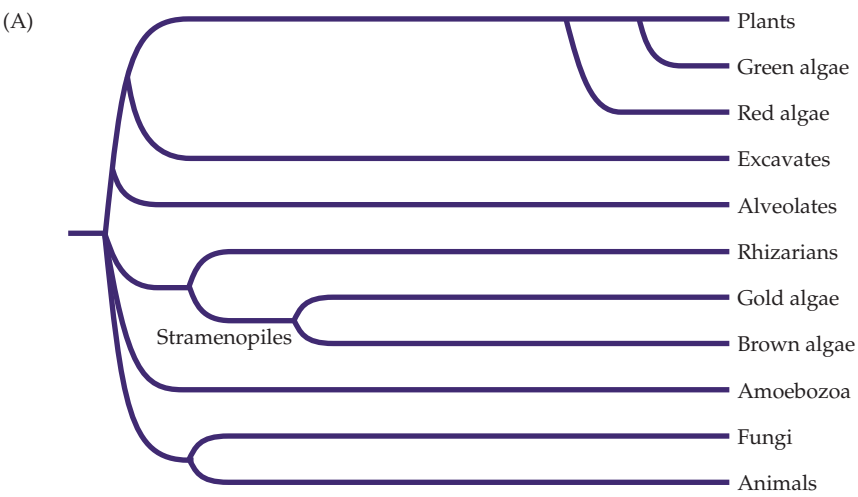



**Figure 17.1** The geological time scale (numbers are in millions of years ago [mya]) and the history of major land plant groups over the past 500 million years, including estimates of their first occurrences. There have been three plant evolutionary floras—Paleophytic, Mesophytic, and Cenophytic—that are defined by the rise and domination of bryophytes/pteridophytes, gymnosperms, and angiosperms, respectively. The beginning and end dates of those floras do not match those of geologic eras, which are dated based on animal fossil records. Although the Cenozoic era was traditionally split into the Tertiary and Quaternary periods, a newer system now divides it into two periods of more equal duration, the Paleogene and the Neogene. (After D. M. Hillis et al. 2020. *Life: The Science of Biology*, 12th ed. Oxford University Press/Sinauer: Sunderland, MA.)

Paleozoic era, during the late Ordovician through Silurian periods. In Chapter 3, we considered some of the evolutionary challenges they faced in accomplishing this major transition. Land plants are descended from the Charophyta, a group of branched, filamentous green algae (Figure 17.2). Today, many green algae grow along seashores, where they are subjected to daily drying as the tides rise and fall. It is likely that land plants evolved under similar conditions. The invasion of plants drastically changed terrestrial environments physically and biotically, setting the stage for two major (and separate) animal invasions (by arthropods and by vertebrates) as well as for all subsequent plant evolution.

During the Silurian and Devonian periods, the terrestrial landscape was completely transformed. This period of time saw an increase in the morphological



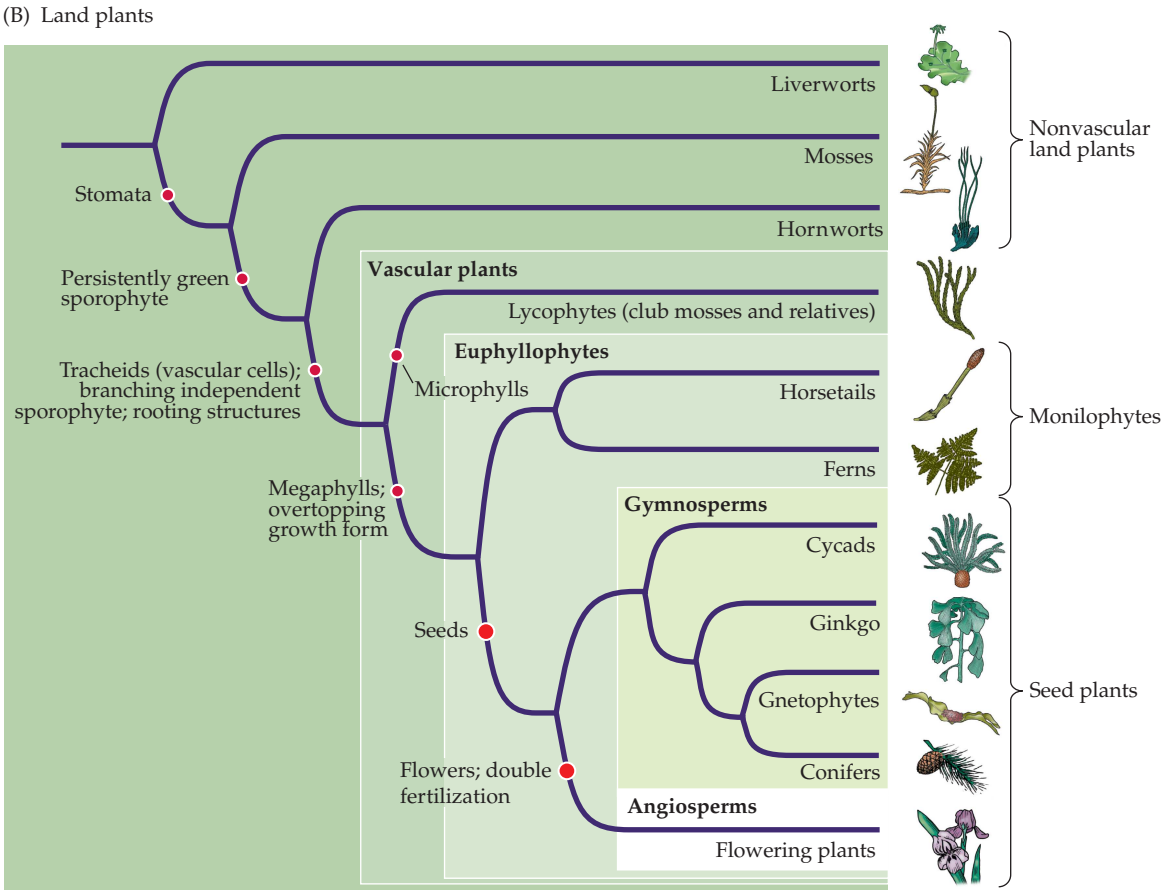




**Figure 17.2** (A) Land plants are most closely related to green algae, and together they are most closely related to red algae. The relationship of this group to other eukaryotes is uncertain. The other two major groups of terrestrial, multicellular organisms—animals and fungi—are most closely related to each other. Most other groups of eukaryotes consist exclusively or primarily of single-celled organisms that live in a variety of marine, freshwater, and terrestrial environments. (B) The major groups of land plants showing their relationships. (B after W. K. Purves et al. 1998. *Life: The Science of Biology*, 5th ed. Oxford University Press/Sinauer: Sunderland, MA.)

Courtesy of S. Scheiner

**Figure 17.3** *Marchantia* is a liverwort, a nonvascular plant. The 3 cm tall structures in the center are the reproductive organs of the female (left) and male (right) gametophytes. Early terrestrial plants may have had a similar form.

Painting by Jan Vriesen, © Denver Museum of Nature and Science

**Figure 17.4** A forest in what is now Colorado as it might have looked 300 million years ago. A shallow sea lies to the east, and rivers choked with gravel flow from the mountains to the sea. The banks of the rivers are colonized by forests of scale trees (*Sigillaria*, Lycopodophyta) and giant horsetails. The trees form a very minimal canopy with little shade. They have green trunks, and even their surface roots are green; the whole tree can photosynthesize. Forests of primitive conifers grow in the higher terrain of the foothills and mountains.

and ecological complexity of both plants and animals and the consequent dramatic radiation of major phyla. At the beginning of the Silurian, the only vegetation consisted of ground-hugging mats in very damp areas, dominated by plants related to liverworts (the Marchantiophyta, Figure 17.3). Based on fossil evidence, these early plants were generally mechanically supported by turgor pressure; their gametophytes produced both male and female reproductive organs, like those of modern bryophytes and ferns; and they had, at most, only rudimentary roots (Bateman et al. 1998). The forms of their sporophytes were simple, but at least one species, *Rhynia gwynne-vaughanii*, had already evolved the ability to spread by clonal growth.

By the end of the Devonian, interactions with vegetation had created soils in many habitats, vascular tissue and then trees had evolved, and tall forests rose above ancient wetlands, dominated by giant lycopods up to 40 m tall (Figure 17.4). With the evolution of root systems and mycorrhizal symbioses (see Chapter 11), land plants diversified and spread to mesic and drier uplands. By the end of the period, many major vascular plant phyla—Psilotophyta (whisk ferns), Lycopodiophyta (clubmosses, Figure 17.5), Sphenophyta (horsetails), and Pteridophyta (ferns)—had evolved, as had the seed plants (Bateman et al. 1998). Today, the first three of these groups grow only as small understory plants. In contrast, ferns grow both as small plants throughout the world and as trees in tropical and subtropical areas, as well as in temperate rainforests in Australia and New Zealand. Like modern plants, early plant life had alternation of generations (see Chapter 6), but unlike most modern plants, the gametophyte (haploid) generation was generally the largest and longest lived. Diploid plants have numerous advantages, however, and during this period lineages arose in which the sporophyte (diploid) generation became increasingly larger and longer-lived and the gametophyte became smaller and shorter-lived.

We can speculate about the selective processes that created this diversification. The evolution of vascular tissue allowed plants to increase in height. This ability probably triggered an "arms race." There is generally little advantage to a plant in becoming very tall unless it is likely to be shaded by its neighbors (see Figure 9.7). If one individual in a population is taller than others, it is likely to outcompete its neighbors for light (see Chapter 2). Similarly, if an entire species in a community is taller than the others, it may outcompete other species in that community for light. Natural selection will then favor individuals of those other species that can grow even taller. This process continues until other constraints balance selection for height. Selection for increased height, as well as other selective forces, resulted in the evolution of wood, which in turn enabled the great diversity of forms of trees and shrubs to evolve. The evolution of roots, in contrast, was probably driven by a combination of abiotic factors as well as competition for nutrients and water. As a result of these adaptations, terrestrial plants were able to spread from wet areas, eventually increasing over nearly the entire surface of the Earth that is not covered by water or ice.

Courtesy of S. Scheiner

**Figure 17.5** *Lycopodium obscurum* (ground pine, Lycopodiaceae) is a clubmoss, a vascular plant approximately 10 cm tall. It has a stem, but no true roots or leaves.

As plants diversified, biotic interactions probably came to play an increasingly important role in their evolution. Many of the ecological interactions we see today evolved very early, including mycorrhizal symbioses and herbivory (see Chapter 11) during the Devonian, and animal pollination and seed dispersal (see Chapter 6) during the Carboniferous and Permian periods.

The Carboniferous period is so named because of the enormous amounts of fossil carbon—oil and coal—that were deposited during that time. Coal deposits were generated mainly from the remains of large vascular plants living in extensive wetlands, including ferns and other spore-producing plants, as well as early gymnosperms (DiMichele et al. 2001). The fossilization is believed to have been a result of the burial of large amounts of organic matter in the anoxic soils of these widespread swamps, and possibly of the evolution of lignin, the tough material that hardens cell walls and is resistant to biodegradation (see Chapter 3 and Chapter 11). The excess of photosynthesis over microbial decomposition of the organic matter (see Chapter 5) resulted in a 50% increase in atmospheric $O_2$ concentrations and a subsequent increase in wildfires (Berner et al. 2007). The organic sources of oil are less well understood; they appear to be primarily marine autotrophic and heterotrophic plankton but may also include wetland plants.

As climates changed in the late Carboniferous—mainly by becoming drier—the swamps disappeared from most areas, and gymnosperms and other, now extinct, seed plants began to dominate the landscape. At the same time, oxygen levels in the atmosphere dropped substantially. These evolving lineages obtained better control over their stomata, allowing them to minimize water loss and to colonize a wider range of habitats (McAdam and Brodribb 2012). This change in the vegetation was rapid, resulting in a nearly complete change in the species present in most locations, with plants from dry uplands rapidly replacing those of the former wetlands.

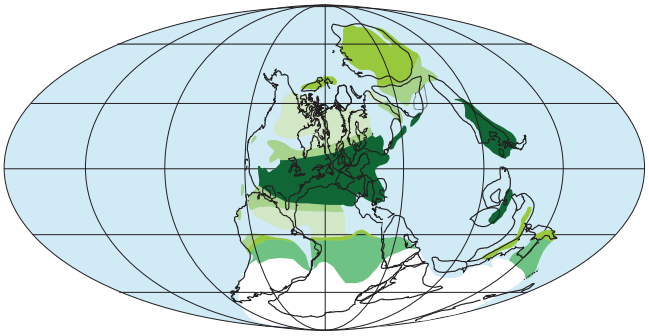
## 17.2 The Mesozoic Era Was Dominated by Gymnosperms and Saw the Origin of the Angiosperms

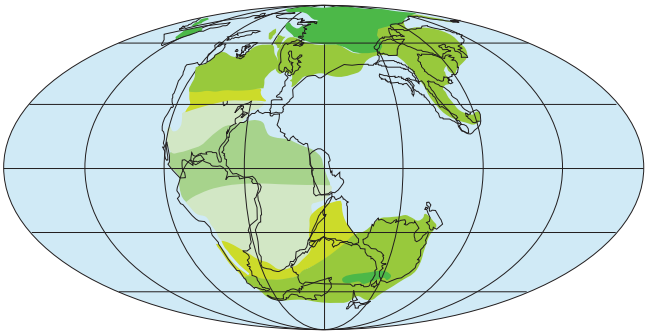
### Gymnosperms were the first group of dominant seed plants

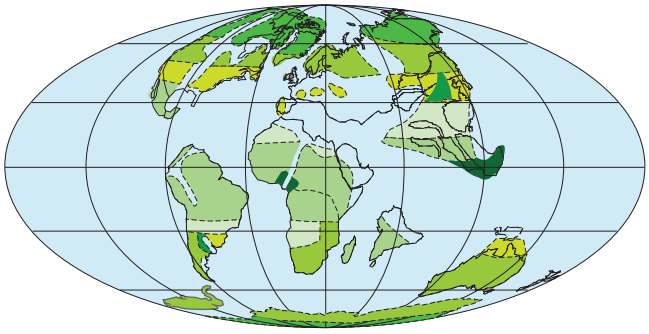
At the beginning of the Mesozoic era, all of the continents were joined together into one supercontinent, **Pangaea**. They have been carried apart and reshuffled by plate tectonics since that time (Figure 17.6). The coalescence of Pangaea, along with increases in atmospheric $CO_2$ and temperatures, led to widespread continental climates, greater seasonality, and widespread aridity. Any photosynthetic benefits of the high $CO_2$ levels may have been more than counterbalanced by the combination of high temperatures (from the greenhouse effect, see Chapter 16) and low or seasonal rainfall.

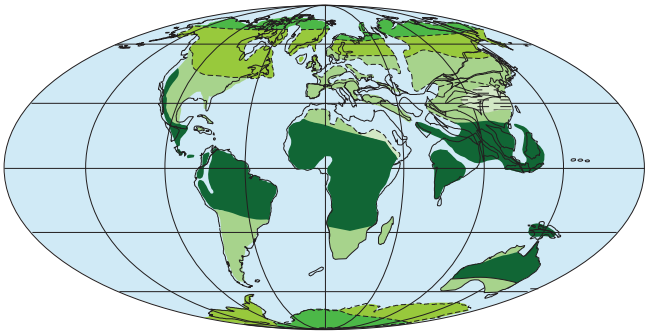
By the late Jurassic period, 150 million years ago, Pangaea had separated into two large continents, northern Laurasia and southern Gondwana. These continents, in turn, would break up into our current continents by the end of the Cretaceous period. The Atlantic Ocean was forming by the early Cretaceous, first with the separation of North America and Eurasia about 130 million years ago, and then with the separation of South America and Africa about 90 million years ago. As a result of these continental movements, areas that were once inland became coastal. Coastal areas are both wetter and less seasonal than inland regions (see Chapter 16). Thus, overall conditions for plant growth tended to improve throughout the Jurassic and Cretaceous periods.

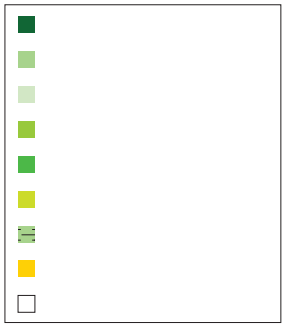
What was happening to the evolution of the major groups of plants during the formation and later breakup of Pangaea? We can consider the history of land plants as the successive dominance of three major groups (see Figure 17.1). Plants in the Pteridophyta (ferns) and Pteridospermatophyta (seed ferns, now extinct) dominated the flora during the late Paleozoic, when climates were warm



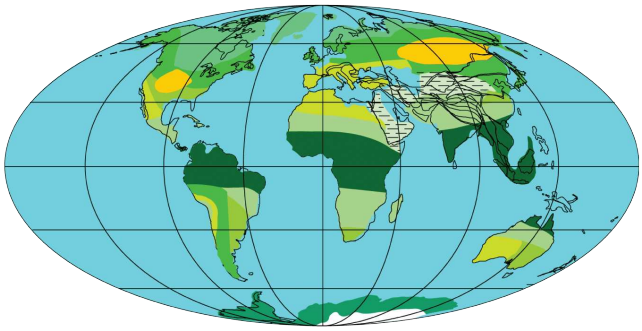








(B) Early Jurassic



(D) Middle Eocene

(E) Modern world

verwet
ummerwet
al desert
nperate
perate

and wet. Despite the similarity in their names, ferns and seed ferns are not closely related; **seed ferns** are an extinct division (phylum) that superficially resembled ferns but reproduced through seeds rather than spores and, unlike ferns, did not have free-living gametophytes.

With increasing aridity and **continentality**—influence of land masses on weather, in this case due to the formation of Pangaea—during the first half of the Mesozoic, the flora became dominated by conifers, cycadophytes, and other gymnosperms, and seed ferns gradually went extinct. Nonflowering families, genera, and species from the Mesozoic that persist today include ferns such as *Osmunda* (an understory species in North American forests), *Ginkgo biloba* (ginkgo, Ginkgoaceae), *Sequoia* (a conifer genus that today includes the giant sequoia of California, see Figure 18.18), and *Araucaria* (a conifer genus of the Southern Hemisphere that today includes several important forest trees, such as *A. excelsa*, the Norfolk Island pine of the South Pacific, and *A. araucana*, the monkey puzzle tree of the southern Andes of South America).

The angiosperms came to dominate the terrestrial fauna during the Cretaceous. The early Cretaceous was a time of warm temperatures and very little difference among the seasons globally. Atmospheric $CO_2$ levels were somewhere between three and four times higher than they are today (Figure 17.7), contributing to warmer temperatures through the greenhouse effect (see Chapter 16). The warm climate and ready availability of $CO_2$ made the early Cretaceous a period generally favorable to plant growth. By the late Cretaceous, however, these favorable conditions had gradually declined, with climates becoming much more seasonal and larger differences in temperature existing between the equator and the poles.

The plant communities of the Mesozoic were similar in physiognomy to today's in some ways. There were distinct biomes dominated by trees, shrubs, or herbs. However, the trees were predominately conifers; in contrast, today's tropical forests and many temperate forests are dominated by angiosperms, with coniferous forests mostly found in temperate or boreal regions. Savannas are a particularly arresting example of how ancient biomes differed from those of today. Today's savannas are dominated by grasses along with scattered angiosperm trees (see Figure 18.27). In the Mesozoic, there were no grasses; instead, ferns provided most of the ground cover, while the trees were largely gymnosperms.

If we consider the modern distribution of ferns, we might think of them as a largely tropical group, as they are most abundant and diverse in the tropics. Interestingly, most of today's fern families originated in moist temperate regions (above 30°N and S latitude) during the Mesozoic (Skog 2001). During this time, ferns were noticeably absent from the tropics. Long-distance dispersal of spores allowed fern families to spread widely in temperate regions, so during much of the Mesozoic, many families were found in both the Northern and Southern Hemispheres. Only later did they spread into the tropics.

## The breakup of Pangaea happened as the angiosperms rose to dominance

The rise of the angiosperms was a critical event in the history of plants. Angiosperms most likely originated in the middle Jurassic, with the major clades arising in the late Jurassic and early Cretaceous (Bell et al. 2010). They initially diversified in the tropics and by the middle to late Cretaceous were spreading throughout the world (Crane and Lidgard 1989). Angiosperms now account for approximately 75% of all land plant species and a level of plant diversity not seen previously on Earth. That movement across the globe happened in two ways, by migration and also passively as the continents themselves moved (see Figure 17.6). We can see evidence of the latter by looking at where related groups of plants are living today: the family Proteaceae is found in Africa, Australia, South America, and India, which were all previously part of the continent of Gondwana (Figure 17.8).

Despite these changes in which groups dominated the Earth's flora, some broad ecological patterns have

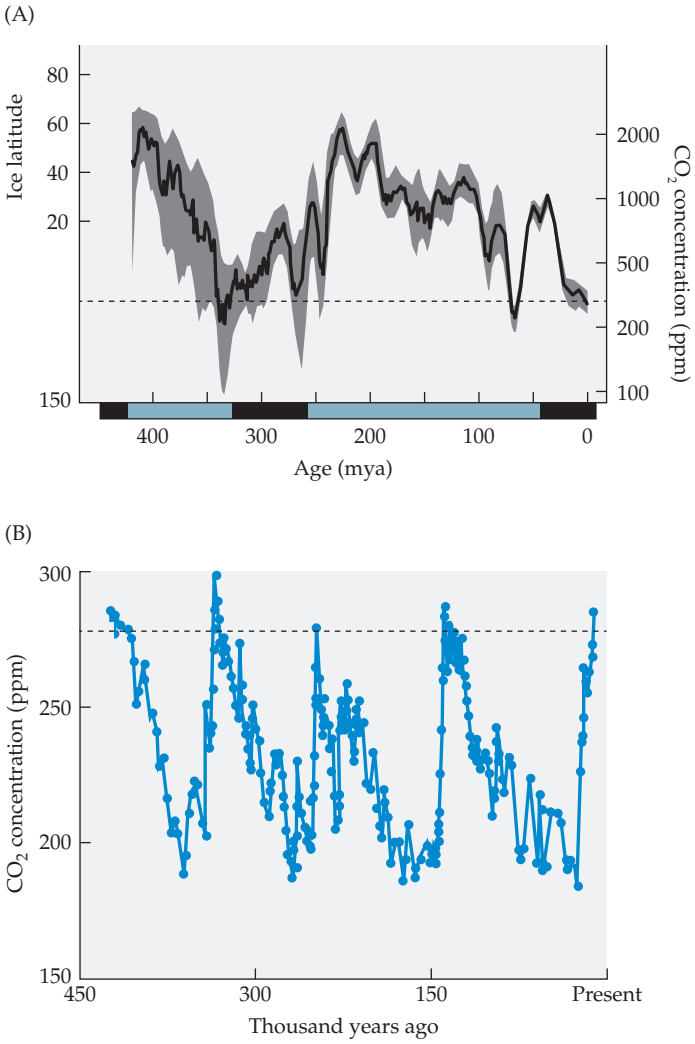



**Figure 17.7** (A) Estimated concentrations of atmospheric $CO_2$ over the past 400 million years. The shaded areas above and below the solid line are upper and lower limits of the estimate. The black bars indicate periods when the Earth's climate was relatively cool; the shaded bars indicate periods of relative warmth. Atmospheric $CO_2$ levels have generally decreased over the past 175 million years. The dashed line indicates the preindustrial atmospheric $CO_2$ level. (B) Estimated concentrations of atmospheric $CO_2$ over the past 450,000 years, measured from trapped air in ice cores from the Antarctic. Over this period, $CO_2$ levels varied, but there was not a trend either up or down. As of 2019, atmospheric $CO_2$ levels are approximately 410 ppm, higher than any concentration in the past 450,000 years. (A after G. L. Foster et al. 2017. *Nat Comm* 8: 14845/CC BY 4.0; B after J. R. Petit et al. 1999. *Nature* 399: 429–436.)

(A)

Courtesy of G. Fox

(B)

Courtesy of S. Scheiner

**Figure 17.8**
(A) *Protea cynaroides* (king protea, Proteaceae) is the national flower of South Africa. (B) *Banksia ericifolia* (heath-leaved banksia, Proteaceae) is native to New South Wales, Australia.

not changed. One similarity between ancient and modern floras is the pattern of species richness across the globe. During the Cretaceous, the greatest species richness was in the tropics, with richness decreasing toward the poles—the same pattern as today (see Figure 19.1).

The separation of the continents and extensive mountain building during the Cretaceous both contributed to regional climate change and helped to accelerate speciation rates in both plants and animals through geographic isolation (see Chapter 9). These environmental changes helped the angiosperms to spread, proliferate, and diversify. Their rise, in turn, created profound changes in the structure of communities. First, the nature of animal communities went through a dramatic transformation. Insects began to diversify at this time; today their species numbers vastly overwhelm those of any other group of animals. In some cases, such as the diversification of lepidopterans (butterflies and moths), the timing of insect radiations can be linked directly to the evolutionary radiation of angiosperms.

The evolution of the flower—a key feature of angiosperms—opened new possibilities for more specialized relationships between plants and pollinators. The pollination of earlier vascular plant taxa depended mainly on wind (as in conifers) or on beetles (as in cycads), which mainly consumed pollen and only occasionally transferred some of it to other plants. With the evolutionary diversification of flowers came nectar rewards, which, with diversification in flower shapes, colors, and scents, increased the reliability of insects in moving pollen to receptive stigmas of other flowers of the same species. Today, some of the best-known examples of mutualisms are those of flowers and their pollinators (see Chapter 6). The evolution of fruits similarly opened new possibilities for interactions between plants and their animal seed dispersers (and seed predators).

Another important difference between angiosperms and the groups that preceded them is that angiosperms include many herbaceous and small woody species that often have faster growth and life histories. In conifers, seed maturation can take from 6 months to 3 years, while almost all angiosperms can mature seeds in just weeks to months. More generally, angiosperms encompass an extremely wide diversity of plant architectures, life histories, and reproductive modes.

At the end of the Cretaceous, tropical forests were found in what is now southeastern North America. These forests had open canopies with relatively tall conifers similar to *Sequoia* and *Metasequoia* (Taxodiaceae), with an understory consisting of a variety of angiosperms, including rosette palms. Forests with a similar mix of overstory conifers and understory palms and other angiosperms can still be found in parts of Florida. At this time, the middle of North America was covered by a vast shallow sea. The areas west of the sea were covered by a broad-leaved evergreen community with a variety of species like those in the southeast. Temperate forests, dominated by broad-leaved evergreen conifers, cycads, and ginkgos, covered what is today the northwestern United States and southwestern Canada. In those forests, the angiosperms were generally deciduous species growing along streams and in other disturbed habitats. The far northwestern areas of North America—above 60°N—were covered by broad-leaved deciduous forests with a mixture of conifers (such as Taxodiaceae and *Ginkgo*) and angiosperms, including species related to birches (Betulaceae) and elms (Ulmaceae). The milder climates of the Cretaceous allowed these broad-leaved communities to exist where now there is taiga and tundra. Thus, by the late Cretaceous some plant communities were

starting to more closely resemble ones that we might find today, although the locations of such communities might have been very different.

### The boundary between the Cretaceous and Tertiary periods resulted in big changes in the flora and fauna

The Mesozoic era came to an end 65 million years ago. There is strong evidence that a large asteroid (10 ± 4 km in diameter) slammed into what is now the Caribbean Sea near the current Yucatán Peninsula. The asteroid struck from the southeast, throwing vast amounts of debris into the atmosphere to the north. This debris eventually settled to the ground, forming a distinct geological layer marking the transition between the Cretaceous and Tertiary periods, known as the K-T boundary.

How this asteroid impact led to massive global change—including the extinction of the dinosaurs—is still an area of active research and controversy. According to one hypothesis, the debris thrown up by the impact may have been sufficiently dense to darken the skies for long enough that photosynthesis was largely shut off. Other evidence suggests that the impact radiated enough heat to start massive and widespread fires that lasted for several years. The debris remaining in the atmosphere from the impact and the fires is likely to have led to a drastic cooling in the world's temperatures for a period of years to decades. It has also been hypothesized that the disruption in the global carbon cycle caused by the asteroid impact led to major oscillations in global climates for the next million years. It is now generally accepted that this event is tied to the extinction of the dinosaurs, along with 70% of the marine species that existed at the time.

In North America, the flora of the Paleocene epoch contained many fewer species than the flora of the late Cretaceous. In the southern part of the continent, almost 80% of existing plant species became extinct, while polar forests had a 25% extinction rate. Even as far away from the impact as New Zealand, a study of fossil pollen and spores found that a diverse flora consisting of gymnosperms, angiosperms, and ferns was abruptly replaced with one consisting of a small number of fern species (Vajda 2001). Recovery occurred only gradually over more than a million years, with the eventual evolution of many new species, particularly angiosperms.

## 17.3 The Cenozoic Era Was Dominated by Angiosperms

In the early Eocene epoch, global temperatures began to fall, following falling global $CO_2$ levels. In North America, the Rocky Mountains and the Sierra Nevadas continued rising, creating a midcontinent rain shadow (see Chapter 16). Along with continued mountain building in the Himalayas, these geological changes led to a cooling in the global climate. The cooling was strongest in the Southern Hemisphere: The South Pole was covered by ice during the Oligocene epoch, but the North Pole was not. This cooling may have been reinforced by another effect: Massive mountain building during this period resulted in the weathering of large amounts of rock. Chemical reactions during weathering remove $CO_2$ from the atmosphere (Ruddiman and Kutzbach 1989). Temperatures remained high enough, however, that throughout the Miocene epoch, the vegetation was similar to that which today is found in tropical regions, even at high latitudes.

These changes in climate had important ecological consequences. By the Oligocene epoch, broad-leaved evergreen forests in the middle of North America had thinned out to become tropical woodlands and savannas. Grasses, which had originated about 65 million years ago, expanded globally so that grass-dominated biomes (see Chapter 18) were widespread by about 35 million years ago. Grass eaters, in turn, evolved from ancestors that ate leaves from trees or shrubs. Most notably, horses evolved in North America during this time. (Eventually horses migrated to Eurasia over the Bering land bridge between Alaska and Siberia, and they went extinct in North America 10,000 years ago. Today's wild herds in North America are descendants of horses brought by the Spanish just 500 years ago.)

The drop in $CO_2$ levels may have had a second important effect on plant ecology beyond the change in temperature. Studies of the carbon isotope ratios of fossil plants show that $C_4$ angiosperms first evolved from $C_3$ ancestors during the Oligocene. $C_4$ plants are more efficient than $C_3$ plants at taking up $CO_2$ (see Chapter 2), especially at low $CO_2$ concentrations. During the late Miocene, tropical and subtropical grasslands dominated by $C_4$ grasses spread widely, especially between 8 and 7 million years ago. Their spread has often been explained as a response to the drop in atmospheric $CO_2$ concentrations. However, the drop in $CO_2$ had already become substantial by the early Miocene, some 16 million years earlier. This delay in the spread of $C_4$ grasslands is unexplained. Some researchers have suggested that changes in atmospheric $CO_2$ concentration were a precondition for the rise of $C_4$ grasslands (Sage et al. 2012), with climate change being the most important factor (Pagani et al. 1999; Edwards et al. 2010).

We can look at changes in the distribution of $C_3$ and $C_4$ grasses in response to much more recent climate changes in order to understand what might have happened during the Miocene. Studies of fossils from 10,000 years ago in two lake beds in Mexico and Guatemala support the hypothesis that climatic drying was responsible for the

local spread of $C_4$ grasses (Huang et al. 2001). Atmospheric $CO_2$ concentrations were the same at both sites; thus they can be ruled out as a cause of differences in the vegetation. At the Mexican site, $C_3$ grasses spread as the local climate became rainier at the end of the most recent glaciation; at the Guatemalan site, $C_4$ grasses spread as the local climate became drier. Studies of grasslands in southern Africa also suggest that the relative proportion of $C_3$ to $C_4$ plants varies with climate (Scott 2002).

How will Earth's vegetation change in response to the current rise in atmospheric $CO_2$ levels? Will it favor the spread of $C_3$ plants over $C_4$ plants (see Chapter 16)? It seems likely that the answer will depend on how changes in the concentration of $CO_2$ affect regional climates.

## 17.4 Many Different Methods Are Used to Uncover the Past

How do we know so much about the history of plant communities? Studies of the past use a variety of techniques. Of primary importance is the study of fossils. From these remains of leaves, stems, flowers, pollen, and other plant parts, we can reconstruct past communities and climates (Prentice et al. 1991).

The sources of fossils depend on the habitat where they were formed. Because most organic material decomposes quickly, **anoxic**—oxygen-free—conditions are ideal for its preservation. In nonmarine habitats, these conditions are most often found at the bottoms of lakes and bogs. Sediments formed in lakes and bogs are the primary source of **macrofossils**, such as leaves, flowers, stems, and seeds. For information from hundreds of thousands or millions of years ago, we rely primarily on hard-rock fossils, in which the organic material has been replaced by nonorganic material, as well as fossils in organic deposits of coal. Material from more recent periods may also still be organic. Bogs, which have highly acidic waters that slow decay, are especially useful in this regard.

Fossils can be retrieved from lakes and bogs by removing a core from the sediments at the bottom. A long tube or pipe, open on the bottom, is pushed down into the sediments, and the core is then carefully drawn up. Thin slices of the core are examined; fossils, if present, are identified, and the slices are aged using isotopic dating.

In addition to delivering macrofossils, these cores yield pollen grains (Figure 17.9) and other **microfossils**. Wind-borne pollen often lands in a bog or lake and settles to the bottom. Because pollen can travel long distances, cores from lake or bog sediments provide a profile of the community over a wide area around the body of water. In contrast, macrofossils usually provide a picture only of the plants growing immediately adjacent to the area studied.

**Palynology**—the study of spores and pollen—has strengths and weaknesses as a method of studying the past. One strength is that pollen is readily preserved, and large amounts of it can usually be found. A weakness is that not all species are present or equally represented, or even identifiable. Only the pollen of wind-pollinated species makes it into lakes and bogs in quantity; animal-borne pollen is largely missing. Thus, in temperate regions sediment core samples are biased toward trees—because most temperate-zone trees are wind-pollinated—and wind-pollinated herbaceous groups, especially grasses and sedges. In wet tropical regions many tree species are animal pollinated, similarly creating a biased pollen record (meaning that some groups that were present appear to be absent, while other groups are overrepresented in the pollen record and appear to have been more abundant than they actually were).

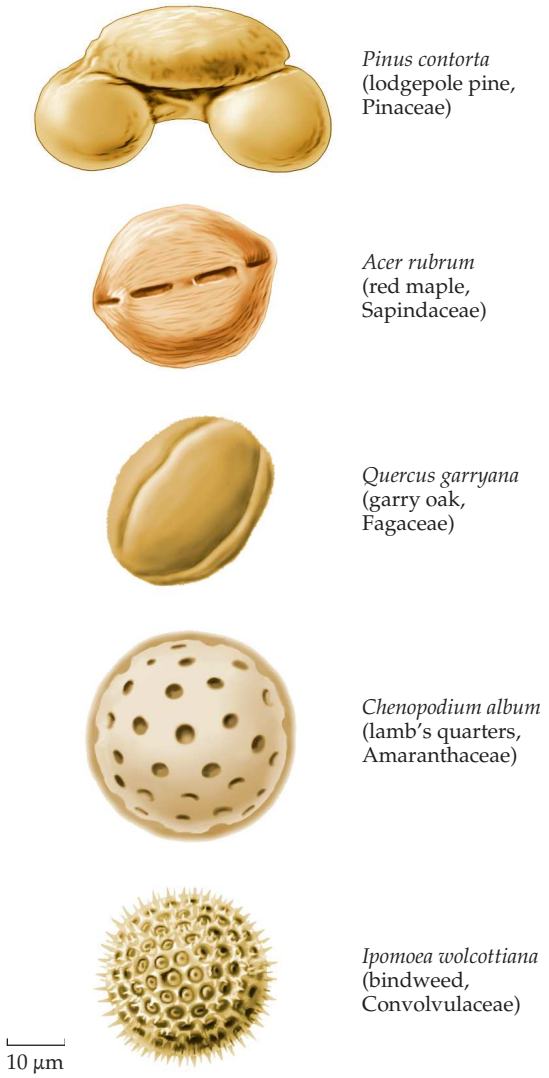



**Figure 17.9** Pollen grains preserved in sediments can be used to identify plant species present at some time in the past. (After E. C. Pielou. 1991. *After the Ice Age.* University of Chicago Press: Chicago.)

The level to which pollen can be identified varies among taxa. Pollen from some trees can be identified to species. In many genera, however, only subgenera can be distinguished. For example, among pines (Pinaceae), the pollen of *Pinus resinosa* (red pine) and *P. banksiana* (jack pine) cannot be told apart, although they can be differentiated from that of *P. strobus* (white pine). In some groups, such as the grasses, pollen can be identified only to family. Despite these limitations, palynology has been key to reconstructing communities and migration patterns in temperate regions.

Another way in which plant material can be preserved is by drying. This means of preservation has provided macrofossils in arid regions of North America where packrats (*Neotoma*) live. As their name implies, these rodents collect large amounts of plant material, which they store in their dens. The dens have a latrine area, where they urinate on the plant material, caking it with mineral salts. Urination provides substantial evaporative cooling for the den and, incidentally, provides excellent conditions for preservation of the plant material. The accumulations of packrats are referred to as **middens**. A given den site may be used for centuries by successive generations of packrats, with its middens growing to several meters across. A fossilized midden may be preserved for thousands of years in a dry climate. Samples from packrat middens can be radiocarbon-dated and the plant material identified, allowing researchers to reconstruct changes in the plant community surrounding the den site (Díaz et al. 2012).

## 17.5 Vegetation Change in the Recent Past Has Been Dominated by the Waxing and Waning of Glaciers

We now leave the distant past and jump ahead to a point just 20,000 years before the present, when much of the Northern Hemisphere was covered by glaciers. Over the previous 2 million years, during the Pleistocene epoch, the Earth had experienced a series of major glaciations. Starting about 850,000 years ago, there were gradual periods of glacial buildup, followed by abrupt switches to warmer weather and short periods of glacial retreat (called interglacials) (Figure 17.10A). Each of these periods of alternating cold and warm weather lasted about 100,000 years. These cycles are caused mainly by changes in the tilt of Earth's axis, combined with differences in the wobble

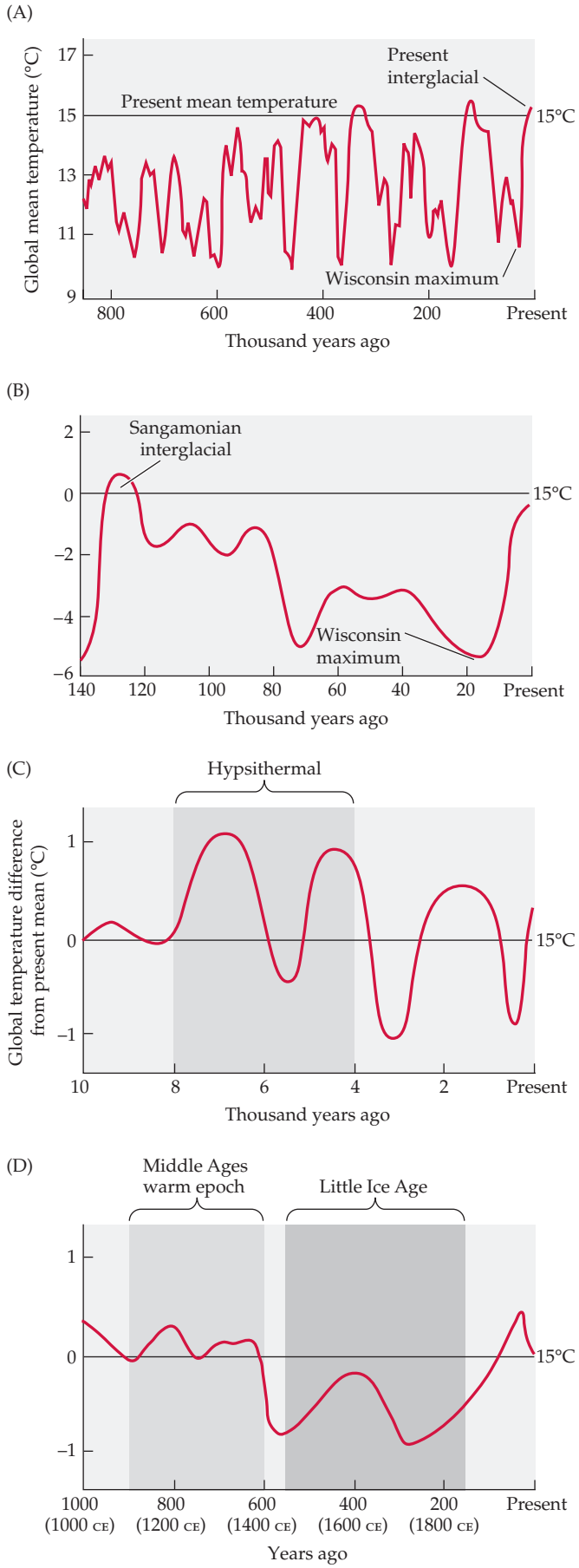



**Figure 17.10** Mean global temperatures over geological time for various periods: (A) the past 850,000 years; (B) the past 140,000 years; (C) the past 10,000 years; (D) the past 1000 years. (After D. M. Gates. 1985. *Energy and Ecology*. Oxford University Press/Sinauer: Sunderland, MA.; D. M. Gates. 1993. *Climate Change and Its Biological Consequences*. Oxford University Press/Sinauer: Sunderland, MA.)

of the axis and ellipticity of Earth's path around the sun (see Chapter 16). The various glacial advances and retreats are known by various names. In North America, the most recent glacial advance is called the **Wisconsin glaciation**, named for one of the locations of its southernmost extent (Figure 17.11). Because each glacial advance destroyed most of the traces of previous ones, we focus here on the Wisconsin glaciation, for which the best data exist.

What has past climate change meant for plants? As glaciers advanced and retreated, plants in northern latitudes migrated across continents, following moving climate zones. In North America, diversity increased in middle and high latitudes (Figure 17.12). In the tropics, climate change resulted in the contraction of forests and the expansion of grasslands. Our study of these changes in plant distributions has led to new understanding of the causes of diversity patterns across the globe, as well as new insights into the nature of plant communities.

## At the glacial maximum, climates and habitats were very different from today

The ice had begun to advance about 2.6 million years ago, and starting about 115,000 years ago, it spread rapidly in the Northern Hemisphere. At their maximum about 26,500 years ago, the glaciers were 3–4 km (about 2 to 2.5 miles) thick at their greatest depth—about as high as seven of the tallest Manhattan skyscrapers stacked one on top of another (or twice as high as the

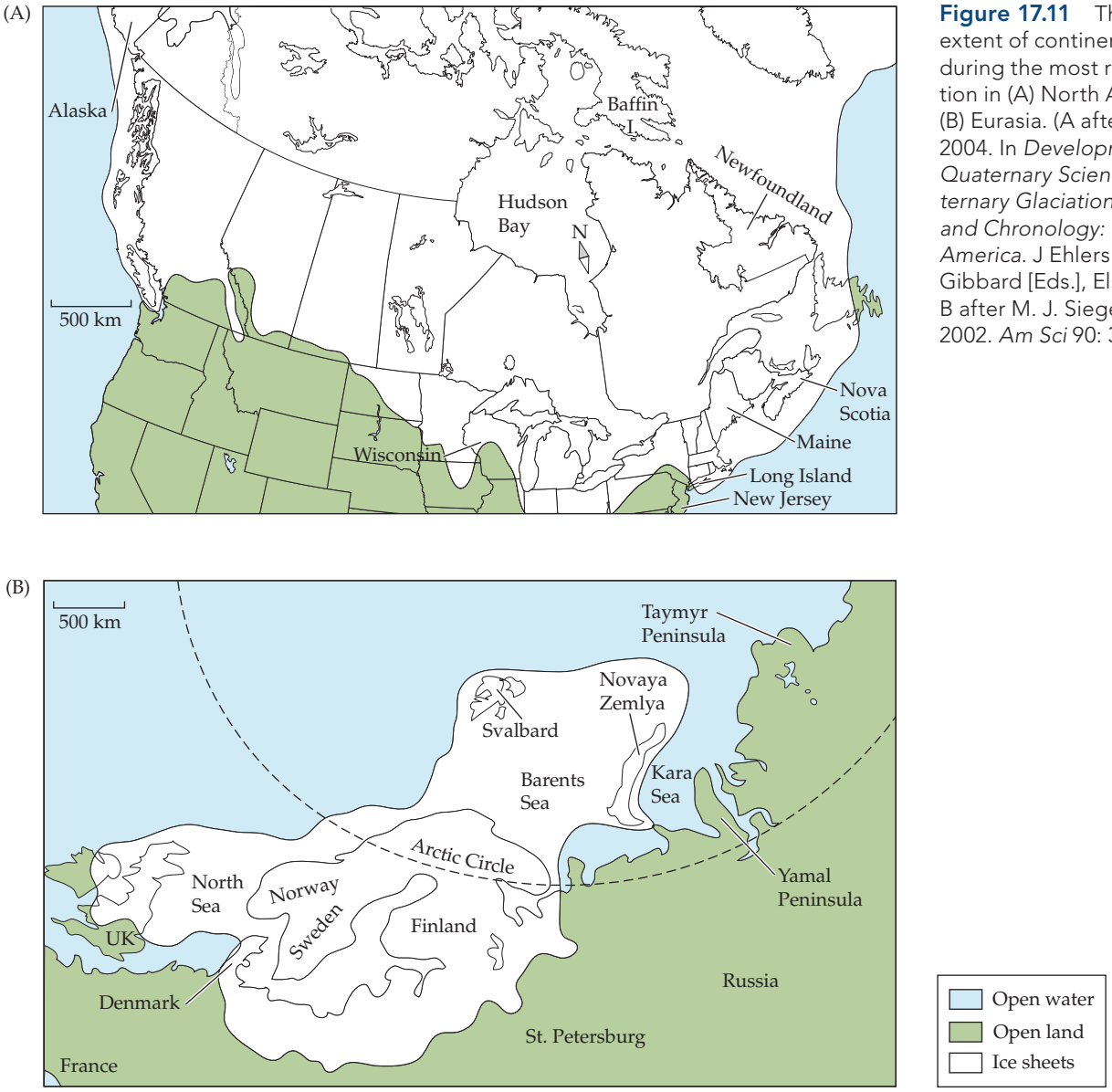



**Figure 17.11** The greatest extent of continental glaciers during the most recent glaciation in (A) North America and (B) Eurasia. (A after A. S. Dyke. 2004. In *Developments in Quaternary Sciences, 2. Quaternary Glaciations - Extent and Chronology: Part II: North America*. J Ehlers and P. L. Gibbard [Eds.], Elsevier B. V.; B after M. J. Siegert et al. 2002. *Am Sci* 90: 32–39.)

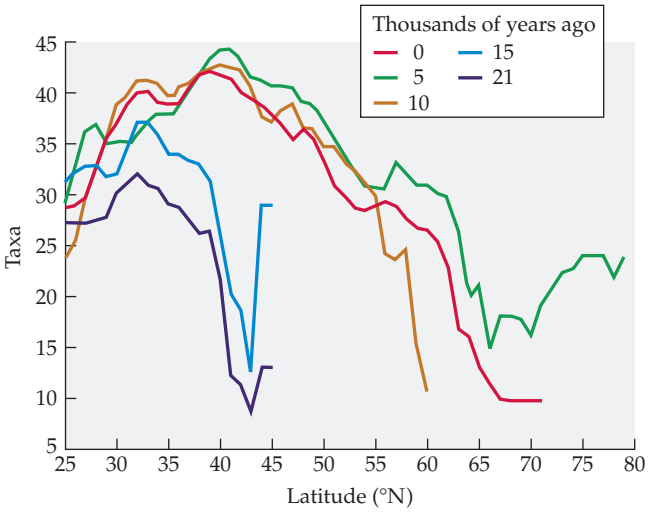



**Figure 17.12** Average plant species richness in North America over the past 21,000 years. For each curve, richness was estimated for each band of 1° latitude. At the height of the glaciation 20,000 years ago, no species were found farther north than about 45° latitude. In the far north, species richness peaked about 7000 years ago during the temperature maximum. (After D. Griffin. 2011. *Plant diversity gradients in eastern North America 21,000 BP to present*. International Biogeography Society Biennial Meeting: Heraklion, Greece.)

Grand Canyon is deep) and heavy enough to compress the bedrock underneath. Global temperatures were 5.5°C to 6°C colder on average than they are today. About 25% of Earth's land area was under ice even in summer. In North America glaciers covered most of Canada, with a southern border that extended from Puget Sound to New York City. Northwestern Europe was largely covered by ice, with glaciers covering most of England and Ireland and all of Iceland and extending as far south as northern France and central Germany and Poland and eastward into western Russia. In the Southern Hemisphere, there were few glaciers except in the southern parts of Chile and Argentina, in Antarctica, and in high elevations in the Andes Mountains, because not much continental land area exists at high latitudes. Much of Asia was ice-free except for high elevations, because although it was as cold as North America, it was too dry for snow to accumulate into glaciers. Curiously, and important for human evolution and history, marshy, ice-free land now known as Beringia existed far to the north extending from Siberia in Asia to current-day Alaska, and humans and many mammals and birds made this their home for thousands of years. Some of these people migrated from Asia to North America eventually, peopling North, Central, and South America. Beringia also was important for plant and animal migrations. The southeastern coast of North America was a refuge for species pushed southward by the glaciers. Similar refuges existed in southern Europe. With so much water tied up in glaciers, sea levels were 85 to 130 m lower than they are today. This greatly extended coastlines, created land bridges to areas that are now islands, and created many other changes in climates and the features of Earth's surface. For example, Great Britain was a peninsula of Europe until about 6200 BCE.

Patterns of vegetation across the globe were very different from today. South of the glaciers, and in other ice-free areas such as eastern Siberia, were vast stretches of tundra and **muskeg**—a grassy bog habitat with scattered, stunted conifers. The tree species of the taiga and temperate forests that today cover these areas were then found far to the south. In southeastern North America, along the coastal plain, there were forests of deciduous trees such as *Quercus* (oak), *Carya* (hickory), *Fagus* (beech), and *Acer* (maple). Along the southern reaches of the Mississippi River valley were coniferous forests dominated by species such as *Picea* (spruce), *Thuja* (cedar), *Abies* (fir), and *Larix* (larch). In the Amazon and southeast Asia, tropical rainforests shrank greatly and were replaced by tropical grasslands and savannas.

The discovery of these distributions addresses an important question in plant community ecology: whether communities are highly coordinated units or simply collections of coexisting species (see Chapter 12). The assemblages of species that existed 20,000 years ago, at the most recent glacial maximum, have no modern equivalent. For example, *Tsuga canadensis* (eastern hemlock, Pinaceae) and *Acer saccharum* (sugar maple, Sapindaceae) are codominants in forests around the Great Lakes today. Yet during the glacial maximum, *T. canadensis* was found mainly along the Mississippi River, while *A. saccharum* was found several hundred kilometers to the west along the Gulf Coast and parts of inland Texas. In the late 1970s and early 1980s, palynological studies of past species distributions and the mapping of northern migrations in response to climatic warming demonstrated that the species compositions of plant communities have been reshuffled many times. These data demonstrate that communities are not highly coordinated units, and they also show that communities rarely reach an equilibrium of either species composition or population structure (Davis 1981).

Studies of the refuges where plant species survived during the glacial maximum also point to some of the important causes of the contemporary differences between North American and European plant communities. As mentioned in Chapter 19, temperate and boreal communities in North America contain approximately

50% more vascular plant species than equivalent European ones. This difference is due to the many European species that went extinct during the Pleistocene (see Table 19.1) but survived in North America because they were able to reach refuges.

Human migrations have had important direct and indirect effects on plant communities. Indirectly, the arrival of humans alters the animal community, including the herbivores. The human invasion of Australia some 50,000 years ago coincided with the extinction of all land mammals, reptiles, and birds that weighed more than 100 kg, as well as many of those weighing 45 to 100 kg. Humans arrived in North America possibly as early as 30,000 years ago, but certainly by 12,000 to 15,000 years ago. This human migration was followed by a major extinction about 10,000 years ago of most of the large mammal species in the New World.

A similar mass extinction of large animals occurred in Madagascar following human arrival about 2000 years ago. David Burney and associates (2003) used abundance patterns of the spores of dung fungus to infer the extinction of large herbivores within 200 years of human arrival. They speculated that the lack of large herbivores led to an accumulation of plant biomass in grasslands, savannas, and along forest edges, making these communities increasingly flammable. The change in fire regimes transformed savannas to thorny deserts and short grasslands. However, the exact timing of human colonization is highly contentious, with some researchers dating human arrival to as early as 10,000 years ago (Hansford et al. 2018) and others dating it to as late as 1350 years ago (Anderson et al. 2018).

Animal extinctions can create pitfalls for present-day plant ecologists trying to understand plant communities. Today, if you travel the western deserts of the United States, you find many shrubs with sharp thorns and spines. Yet, except for cattle, horses, donkeys, and other introduced species, there are only a few large browsers, such as bighorn sheep, pronghorn, and mule deer. Only after studying the fossil record would you realize that these plants evolved during a time when North America had many large native browsers, including horses, camels, elephants, and giant ground sloths, all of which are now extinct.

Humans also alter plant communities directly. The widespread, frequent fires characteristic of much of modern Australia first began around the time of the human invasion, possibly due to the use of fire by the Aboriginal peoples for a variety of purposes, including hunting along the fire front, promoting the growth of preferred plants, and promoting animal populations that prefer open areas. Using the distinctive carbon isotope ratios of $C_4$ and $C_3$ plants (see Box 2B), Gifford Miller and colleagues (Miller, Fogel, et al. 2005) examined the diets of the herbivorous bird *Dromaius novaehollandiae* (emu) and its much larger extinct relative *Genyornis newtoni*, as reflected in fossil eggshells. Both species switched their diets at around the time of the human invasion, from mainly $C_3$ to mainly $C_4$ plants, as they would have had to if grasslands replaced many woodlands following the introduction of fire. Reasoning that such a shift should also be found in other herbivores, the researchers studied fossil teeth of wombats (*Vombatus ursinus*) and found a similar switch. A simulation study by members of the same research team (Miller, Mangan, et al. 2005) suggested that the introduction of fire to Australia may have increased albedo and surface heating enough to cause regional climate change, preventing monsoon rainfall from penetrating most of the continent and leading to desertification.

## Modern plant communities began to appear as the glaciers retreated

Beginning about 20,000 years ago, the climate slowly began to warm and the glaciers retreated, more rapidly starting about 14,000 years ago. In North America, for several thousand years, as the ice melted, large lakes (much bigger than the current Great Lakes) dominated the interior of the continent. Areas that had been under ice became tundra and then forest or grassland (Figure 17.13). The transition from tundra to forest may have happened within a few hundred years in some locations (Post 2003). Since that time, the average global temperature rose, peaked about 7000 years ago, and then declined (see Figure 17.10B, C). If not for human activities bringing on a new period of rapid global warming, we would likely be seeing the growth of glaciers in the next few thousand years.

During this warming period, trees and other plant species migrated northward. If we just look at aggregate patterns, it seems as if this migration was part of a steady progression (see Figure 17.12). However, if we look more closely, we see that each tree species was unique in its migration pattern (Figure 17.14). Many of these patterns were deciphered by Margaret Davis, who used pollen samples gathered from lakes and bogs throughout eastern North America. Among the fastest-moving species was *Pinus strobus* (white pine, Pinaceae), which migrated at an average rate of 325 m per year. In contrast, *Ulmus* (elm, Ulmaceae) moved at a rate of 250 m per year, while slowpoke *Castanea dentata* (American chestnut, Fagaceae) moved a mere 100 m per year (Table 17.1).

From 1993 *Bull Ecol Soc Am* 74:278

Margaret Davis

One might imagine that wind-dispersed and animal-dispersed species would have migrated at about equal rates and that among animals, birds

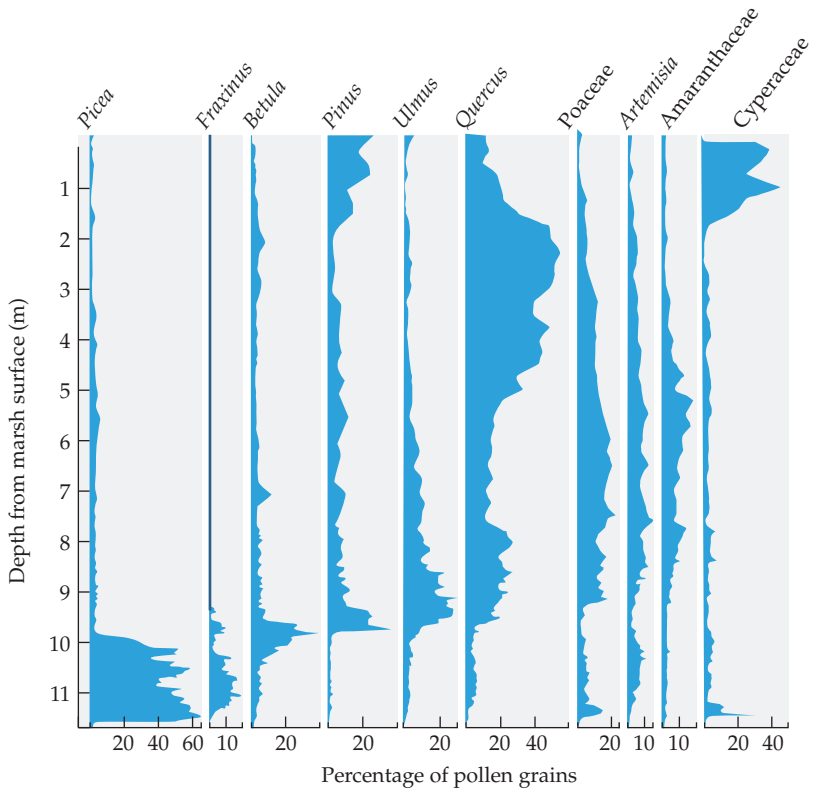



**Figure 17.13** Example of a pollen diagram for Kirchner Marsh, Minnesota. The deepest samples are the oldest. About 20,000 years ago, this area was dominated by *Picea* (spruce) and *Fraxinus* (ash). These were followed by a succession of species: *Betula* (birch), *Pinus* (pine), *Ulmus* (elm), and *Quercus* (oak). The subsequent decrease in oaks and increase in grasses (Poaceae) indicates the existence of a woodland vegetation. Currently the area is again dominated by pines mixed with oaks, but now with a sedge (Cyperaceae) understory. (After T. C. Winter. 1962. *Science* 138: 526–528.)

would have moved seeds faster than mammals. Yet Table 17.1 clearly shows that wind-dispersed seeds generally traveled much faster than animal-dispersed seeds. The exception, *Quercus* (oaks), is dispersed mainly by small mammals in eastern North America. The bird-dispersed species (e.g., *Fagus* and *Castanea*) were actually the slowest to migrate northward.

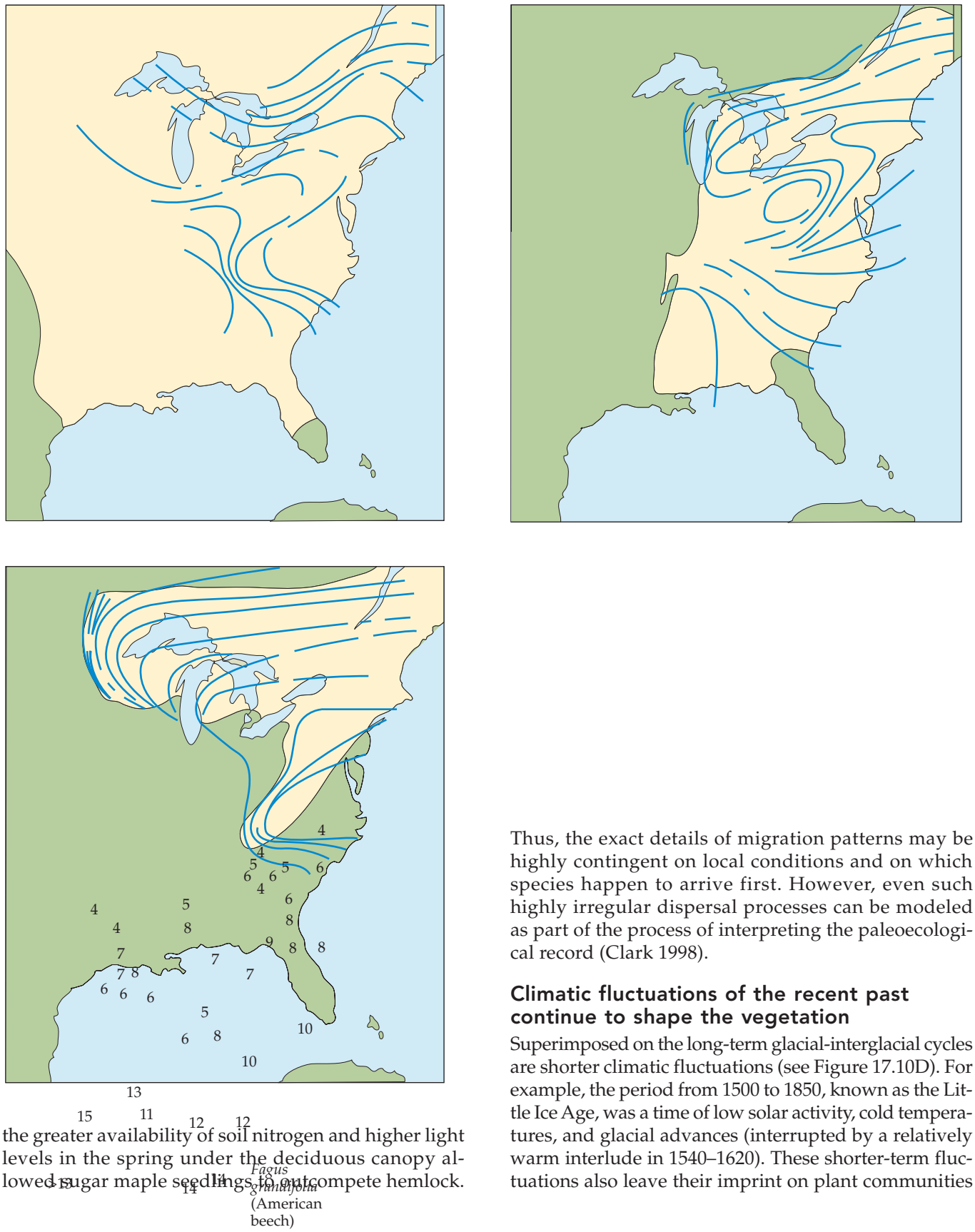
The migration of a species can be very irregular. For example, the migration of *Fagus grandifolia* (American beech, Fagaceae) over the last 7000 years has included both periods of rapid expansion and periods of relative stability (Woods and Davis 1989). During this period, beech migrated northward and westward from the Lower Peninsula of Michigan into the Upper Peninsula and Wisconsin. Some of the delays are attributed to the barrier presented by Lake Michigan; long-distance dispersal by blue jays or passenger pigeons may have been responsible for breaching this barrier. Other starts and stops may have been responses to climatic fluctuations. Beech arrived in Wisconsin on the western shore of Lake Michigan approximately 6000 years ago but did not spread farther west and north for another 1500–2000 years, probably due to cold winter temperatures.

Migration depends on the successful establishment of seedlings. When we look closely at the frontier of a migration, we find a patchwork of populations, in which establishment and spread are highly dependent on local conditions (Lloyd 2005). Margaret Davis and her colleagues (1998) studied the spread of eastern hemlock into an area in the western Upper Peninsula of Michigan. Four communities dominated by white pine were invaded by eastern hemlock about 3000 years ago over a period of 800 years. Over the next 2000 years, climate changes caused the water table to rise and the soil to become wetter. Eastern hemlock became the dominant species, and white pine disappeared from all but one site. In contrast, four other nearby sites that were initially dominated by sugar maple and oak were never invaded by eastern hemlock, possibly because

**TABLE 17.1** Average rates of northward range expansion for trees in eastern North America following the most recent glacial retreat

| Dispersal agent | Species | Rate (m/yr) |
|---|---|---|
| Wind | *Pinus banksiana/resinosa* | 400 |
| Wind | *Pinus strobus* | 300–350 |
| Wind | *Picea* spp. | 250 |
| Wind | *Larix laricina* | 250 |
| Wind | *Ulmus* spp. | 250 |
| Wind | *Tsuga canadensis* | 200–250 |
| Wind | *Abies balsamifera* | 200 |
| Wind | *Acer* spp. | 200 |
| Animals | *Quercus* spp. | 350 |
| Animals | *Carya* spp. | 200–250 |
| Animals | *Fagus grandifolia* | 200 |
| Animals | *Castanea dentata* | 100 |

*Source*: M. B. Davis. 1981. In *Forest Succession Concepts and Applications*, D. C. West et al. (Eds.), pp. 132–153. Springer: New York.




the greater availability of soil nitrogen and higher light levels in the spring under the deciduous canopy allowed sugar maple seedlings to outcompete hemlock.

Thus, the exact details of migration patterns may be highly contingent on local conditions and on which species happen to arrive first. However, even such highly irregular dispersal processes can be modeled as part of the process of interpreting the paleoecological record (Clark 1998).

## Climatic fluctuations of the recent past continue to shape the vegetation

Superimposed on the long-term glacial-interglacial cycles are shorter climatic fluctuations (see Figure 17.10D). For example, the period from 1500 to 1850, known as the Little Ice Age, was a time of low solar activity, cold temperatures, and glacial advances (interrupted by a relatively warm interlude in 1540–1620). These shorter-term fluctuations also leave their imprint on plant communities

and often help explain puzzling aspects of current species distributions.

Vertical migration in response to climate fluctuation has been very important in the American Southwest. This region has many high mountains, plateaus, and other vertical features, and during the Pleistocene much of the plant migration consisted of elevational changes. During periods of maximum glaciation, the climate was cooler and wetter, and pine forests and pinyon-juniper woodlands spread out over the landscape. Large lakes dotted the landscape in what is now the Great Basin desert. Their remnants can be seen today as the Great Salt Lake and the Bonneville Salt Flats. During interglacial periods (like the present), the low-elevation areas became hotter and drier, and desert shrubland became dominant in these areas.

Studies of these patterns in the Grand Canyon in Arizona have disclosed an important phenomenon, called **vegetation lag**. Such lags occur when changes in the vegetation happen more slowly than the rate at which the climate changes. Using evidence from packrat middens, Kenneth Cole (1985) found that as the climate warmed, species disappeared locally as the lower edges of their ranges moved upward. However, these species were not immediately replaced; it took centuries for new species to become established (Figure 17.15). The studies of migration in the Great Lakes region described previously also found evidence of vegetation lag, with species taking 1000 to 2000 years to reach and then dominate areas after the climate had become suitable for them. The studies of the Grand Canyon vegetation also reinforce the notion that each plant species migrates at its own rate, with plant communities changing through time as species combinations change (Cole 1990).

(A)

(B)

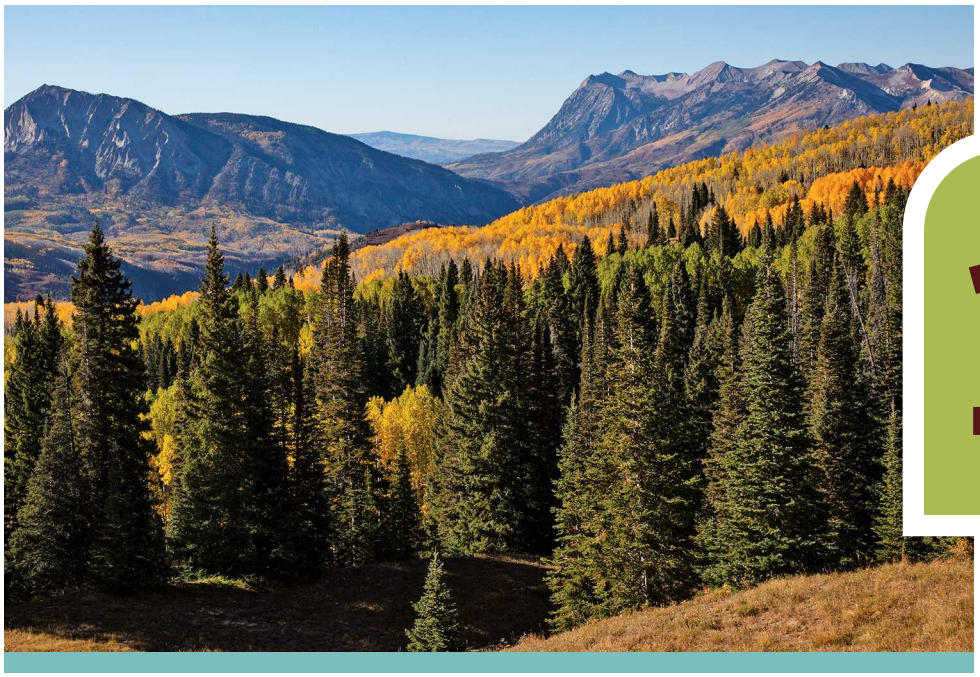
Photograph and reconstruction courtesy of K. Cole

**Figure 17.15** An area near the southern rim of the Grand Canyon in Arizona at 1500 m elevation. (A) Today, this area is a desert shrubland of *Ephedra viridis* (Mormon tea, Ephedraceae), *Coleogyne ramosissima* (blackbrush, Rosaceae), and *Opuntia engelmannii* (prickly pear cactus, Cactaceae), with a few junipers (*Juniperus* sp., Cupressaceae) along the cliff tops. (B) A reconstructed image of the same site as it looked 20,000 years ago, based on plant fossils from packrat middens. The bottom slopes support a woodland of juniper, *Artemisia tridentata* (sagebrush, Asteraceae), and *Atriplex canescens* (shadscale, Amaranthaceae). In the shady alcoves of the cliffs grow *Pseudotsuga menziesii* (Douglas fir, Pinaceae) and *Abies concolor* (white fir, Pinaceae), with a few *Pinus flexilis* (limber pine, Pinaceae) along the cliff tops.

Plant species may also undergo evolutionary adaptation in response to climate change. Rapid climate change—of the kind now occurring in response to rising concentrations of atmospheric $CO_2$—imposes very strong natural selection on populations. Studies of tree species point to the conclusion that adaptation and migration occur simultaneously (Davis and Shaw 2001); they are not alternative responses to climate change. Will plant populations be able to meet the challenge of today's changing climates through some combination of adaptation and migration? Evolutionary responses are limited by the amount of heritable variation within a population (see Chapter 9). Migration is limited both by dispersal biology (e.g., wind can move pine seeds only so far) and by anthropogenic changes in land use that present barriers to species movement. Together, these factors suggest that rates of extinction are likely to increase as a consequence of current global climate change.

## Summary

- The plant communities that we see today are not static entities; they are dynamic systems that change in a complex fashion on both ecological and evolutionary time scales.
- The Paleozoic era witnessed the establishment of plants on land and the origins of many of the basic patterns of plant and community diversity.
- The early Mesozoic era saw the dominance of the first seed plants, the seed ferns and the gymnosperms.
- An explosion of new forms and patterns occurred in the Cretaceous with the diversification of the angiosperms.
- During the Cenozoic, angiosperms continued to evolve, and more recognizably modern species and communities appeared.
- These changes in the Cenozoic were the result of ecological interactions and evolutionary change, driven by environmental and climate change caused primarily by continental drift, mountain building, and variations in the Earth's orbit, and evolving biotic interactions such as competition, herbivory, and pollination.
- Today's temperate plant communities were shaped mainly by patterns of glaciation and other climate changes during the past 2 million years.
- The last glacial retreat began about 12,000 years ago, with temperatures first rising, peaking at about 7000 years ago, and then dropping.
- Along with these climate changes have come migrations of species, resulting in shifting mosaics of species and communities across the landscape.



Courtesy of Jeff Mitton

18

# Biomes and Physiognomy

Our planet has a complex pattern of climates, which in turn has a major role in creating the complex pattern of vegetation and community types we find on it. Across a continent, this variation falls into regular patterns. Here we introduce several of these broad patterns. The first of these is a north-south gradient in vegetation form due to temperature. A second major pattern is variation from west to east in response to changes in average precipitation. Such patterns exist on all continents because of the latitudinal differences in solar radiation and the movements of weather systems described in Chapter 16. Ecologists divide these large-scale vegetation patterns into units called **biomes**: major biogeographic regions that differ from one another in the structure of their vegetation and in the form of their dominant plant species (Clements 1916). Biomes represent the greatest spatial extent at which ecologists classify vegetation. A biome can be very large, like the taiga (northern coniferous forest), or much more restricted in size, like the temperate rainforest. The same biome can be found on more than one continent or island, like tropical rainforest in South America, Australia, Southeast Asia, and central Africa. In this chapter, we examine these broadscale patterns and look briefly at the major characteristics of and threats to the world's major biomes.

Above: The Ragged Mountain Wilderness from the Ruby Mountains, Colorado, U.S.A.

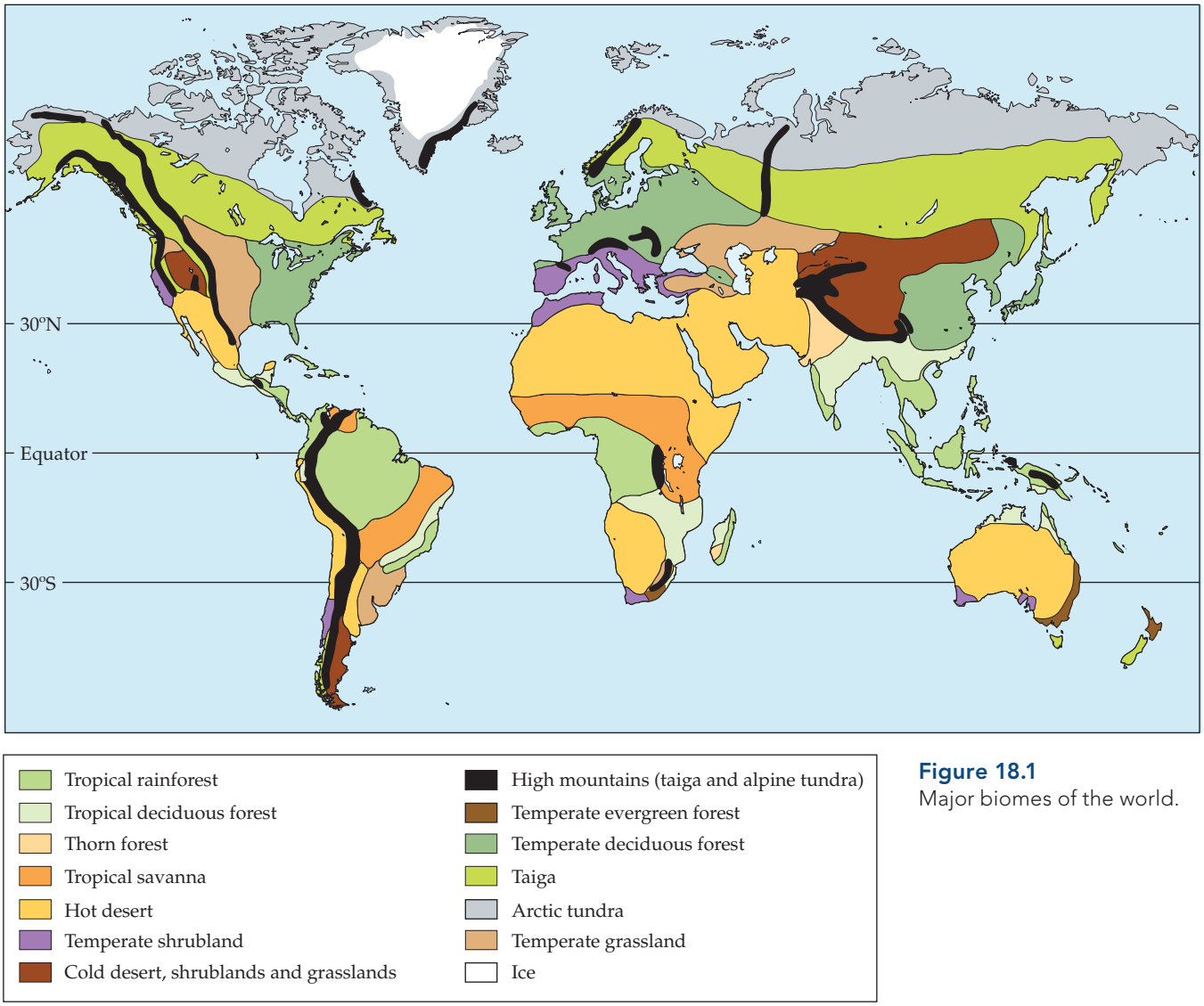



**Figure 18.1**
Major biomes of the world.

## 18.1 Vegetation Can Be Categorized by Its Structure and Function

### Plant physiognomy varies across the globe

Classification of the world's vegetation into biomes (Figure 18.1; Table 18.1) provides us with useful categories that describe major aspects of vegetation, such as its structure, function, and characteristic adaptations. The type of biome not only tells us a great deal about what kinds of plants might be found growing in a given area, but also indicates something about the types of animals and other organisms likely to be living there and the major environmental constraints on living things. Biomes are strongly determined by climate, particularly by temperature, rainfall, and seasonality (Figure 18.2). Similar climates in different parts of the world contain similar biomes, although they may differ in important ways, such as their evolutionary history. For example, temperate evergreen forests are similar in being dominated by trees that stay green all year, but they differ in being dominated by needle-leaved conifers in North America, by broad-leaved beeches in South America, and by broad-leaved eucalypts in Australia.

Biomes are defined by the physiognomy of the dominant or most obvious plants. Thus, we recognize different types of forests (e.g., evergreen or deciduous, broad-leaved or needle-leaved), shrublands, and grasslands. However, we must not lose sight of the trees for the forest: There may be a tremendous amount of variation within a given biome. In her study of communities

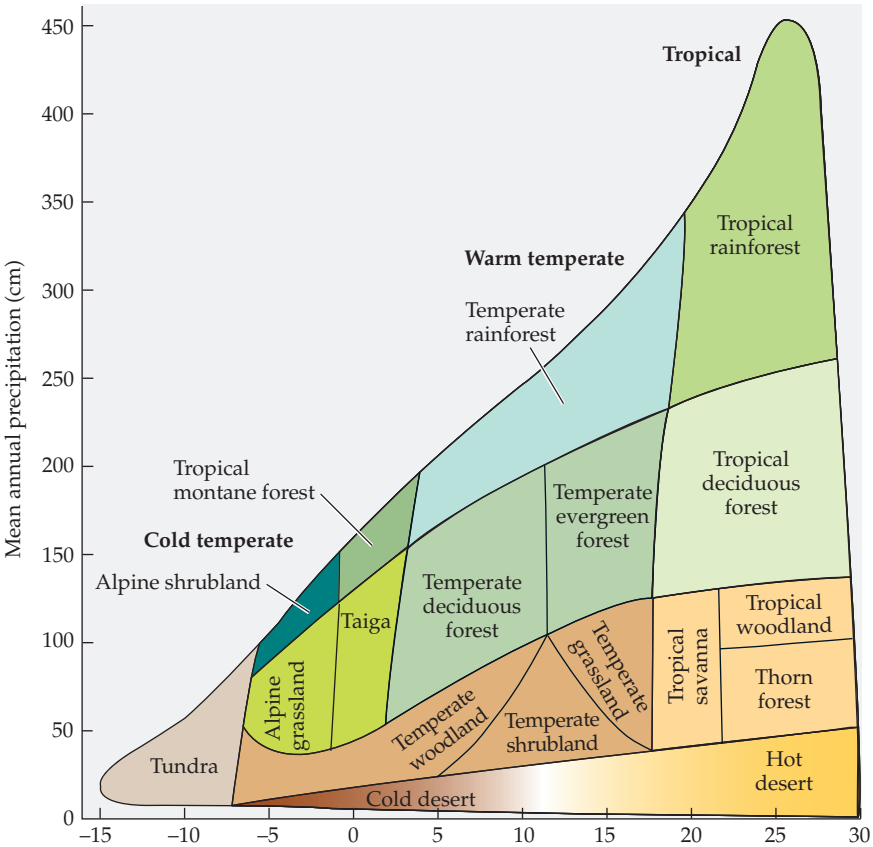



**Figure 18.2** The distribution of biomes is determined by climate, especially mean annual temperature and precipitation. In regions within the dashed lines, other factors—such as fire, grazing, and seasonality of precipitation—strongly affect which biome is present. Climate can also interact with factors such as soil type to determine biome distributions. (After R. H. Whittaker. 1975. *Communities and Ecosystems*, 2nd ed. Macmillan: New York.)

within a single biome, the temperate deciduous forest of eastern North America, E. Lucy Braun recognized 12 different major sub-biomes and a number of minor variants (Braun 1950). Within a biome, there may be patches of vegetation that appear "not to belong," such as riparian forests found along streams in grasslands or deserts. The occurrence of these patches reminds us that climate is only one factor in determining vegetation. Local variation in soils and in topography, especially as it affects microclimate, can influence the type of vegetation found in an area. In turn, the animals living in a given biome are determined by the vegetation there, as well as by climate and other factors.

Biomes are useful, if somewhat arbitrary, descriptive classifications, rather than quantitative or objective categories. Boundaries between biomes may represent the range limits of the dominant species, whereas other species may have ranges that span those boundaries. While we draw sharp boundaries on a map, in reality the boundaries of biomes are often fuzzy. Similarly, the categories that we define ignore the fact that a particular place might not fit easily into any given biome. Furthermore, different scientists might include a place in different biomes. Some use fewer, broader groupings, whereas others use a more detailed breakdown of vegetation types. The definitions of biomes given here may not exactly match those given in other texts. There is a broad consensus, however, on the rough categories and approximate definitions of biomes.

Our use of *temperate rainforest*, for example, may not be identical to someone else's use of that term, but it will be close. If you were to search the scientific literature, you would find many different descriptions of places called rainforests. There is, however, no single, rigorous definition of the term—it is, in fact, used to refer to forests occurring in places where rain is abundant and there is no extended dry season. Tropical rainforests occur where there is substantial year-round rain; temperate rainforests tend to occur where there is heavy winter rain and lighter summer rain but considerable summer fog. The term *rainforest* is thus useful as a broad description, but not as a rigorously defined category or unique entity.

Biomes are defined on their physiognomies, ignoring the evolutionary relationships of the species. The result is that the same biome in different parts of the world may contain unrelated species, while different biomes may contain relatively closely related species. Alternatively, we can take those evolutionary relationships into account in our classification. For example, we separate tropical forests based on the amount of precipitation (see below). But, the tree species in the moist tropical forests of South America are more closely related to those in South American dry tropical forests than they are to moist tropical forests in Africa and Asia (Slik et al. 2018). These evolutionary relationships can reflect the history of the Earth. The trees in African and South American moist tropical forests are more closely related to each other than either is to species in moist tropical forests of Asia because Africa and South America were part of a single continent 100 million years ago (see Figure 17.6).

**TABLE 18.1** Major biomes of the world: dominant growth forms and general climatic conditions

| Biome | Dominant growth form | Angiosperms or gymnosperms (dominant or common) | Temperature | Moisture |
|---|---|---|---|---|
| Tropical rainforest | Broad-leaved evergreen trees | Angiosperms | Hot | Wet |
| Tropical montane forest | Broad-leaved evergreen trees | Angiosperms | Mild | Wet |
| Tropical deciduous forest | Broad-leaved deciduous and semievergreen trees | Angiosperms | Hot | Seasonally dry |
| Thorn forest | Broad-leaved deciduous trees | Angiosperms | Hot | Dry |
| Tropical woodland | Broad-leaved semievergreen trees and grasses | Angiosperms | Hot | Moderate |
| Temperate deciduous forest | Broad-leaved deciduous trees | Angiosperms | Seasonally cold | Moderate |
| Temperate rainforest | Needle-leaved evergreen trees | Gymnosperms | Seasonally cool | Wet |
| Temperate evergreen forest | Needle-leaved or broad-leaved evergreen trees | Gymnosperms or angiosperms | Various | Various |
| Temperate woodland | Needle-leaved evergreen or broad-leaved deciduous trees and grasses | Both | Mild | Moderate |
| Taiga | Needle-leaved evergreen trees | Gymnosperms | Cold | Moderate |
| Temperate shrubland | Evergreen shrubs, annual forbs | Angiosperms | Mild | Moderate |
| Temperate grassland | Perennial grasses | Angiosperms | Seasonal | Moderate |
| Tropical savanna | Perennial grasses | Angiosperms | Hot | Moderate |
| Hot desert | Shrubs, succulents, annual and perennial grasses, annual forbs | Angiosperms | Hot | Dry |
| Cold desert | Shrubs | Angiosperms | Mild | Dry |
| Alpine shrubland | Deciduous shrubs | Angiosperms | Cold | Moderate |
| Alpine grassland | Perennial grasses | Angiosperms | Cold | Moderate |
| Tundra | Perennial grasses, sedges, shrubs, forbs | Angiosperms | Cold | Moderate |

*Note*: Other factors, especially seasonality, are also important.

## Forests are closed canopy systems dominated by trees

**Forests** are communities dominated by trees whose leaves touch one another, resulting in a closed canopy. If you were to travel southward in eastern North America from northern Canada to the Carolinas, Georgia, and northern Florida, you would pass through a series of very different forests. Forest composition is affected by both climate and soil; to focus on the effects of climate, we confine our example to the relatively rich soils found inland (away from the nutrient-poor coastal plain).

Immediately to the south of the tree line, you would find taiga or boreal forest, which is dominated by coniferous, evergreen needle-leaved trees. Next you would enter a region dominated by broad-leaved trees—deciduous angiosperms that drop their leaves in winter—although conifers would still be present. Finally, you would come upon broad-leaved evergreen forests dominated by angiosperms that retain their leaves all year. In summary, there is a transition from evergreen forests to deciduous forests and back to evergreen forests, although the type of evergreen tree changes from needle-leaved to broad-leaved, and from gymnosperm to angiosperm. These changes in tree type with latitude are largely driven by differences in seasonal temperatures. Of course, this description of our imaginary trip is very general; for the moment, we are ignoring many other details to concentrate on temperature effects. Similar patterns are found in eastern Asia as one travels from Korea and northeastern Russia to the south of China and beyond to Southeast Asia.

To understand why these changes occur, let's reverse our trip. Beginning in southeastern North America, growing seasons are longest in the broad-leaved evergreen forests found along the Atlantic and Gulf coasts. These

areas are classified as **subtropical** because they rarely experience temperatures below freezing. Tropical and subtropical forests can be dominated by either deciduous or evergreen angiosperms. The critical climatic factor determining which one occurs in a given area is seasonality of precipitation. Areas that receive rain throughout the year are evergreen, while areas with pronounced dry periods are deciduous. As always, there are exceptions to this general pattern. For example, many forests in the southeastern United States are dominated by a particular genus of needle-leaved evergreen gymnosperms: *Pinus* (pines). These pine forests are typically found on soils that are very low in nutrients and are subject to frequent fires. (The effects of soil nutrients on leaf life-span are considered in the context of plant strategies in Chapter 7.)

Moving north, the transition from broad-leaved evergreen to broad-leaved deciduous forests depends on the likelihood of freezing temperatures (Figure 18.3). Broad-leaved evergreen trees are vulnerable to frost through two of its effects. First, leaf tissues will die if they remain frozen. Second, the weight of ice or snow on leaves will cause them to be torn off or, worse, will break branches and even trunks. A very unusual October snowstorm in 1997 in the Midwest and Great Plains of the United States, which occurred before the leaves had fallen, caused major damage to deciduous trees as branches and trunks snapped under the weight of the snow-burdened canopies. Tall, narrow conifers such as spruces (*Picea*) and firs (*Abies*) shed snow far more effectively. So it is no surprise that the boundary of broad-leaved evergreen trees corresponds with that of freezing temperatures.

Farther north, the transition between broad-leaved deciduous and needle-leaved evergreen forests is related to growing season length. An evergreen tree can take advantage of short periods of favorable conditions as long as there is enough unfrozen moisture in the soil to support transpiration (a major limitation for these trees). Conifers are able to avoid leaf damage due to freezing by increasing the osmotic concentration of intracellular water (thereby lowering the cells' freezing point). Ice crystals form only between the cells (which does little damage to cell membranes); this acts to draw water out of the cells, which further increases the osmotic concentration inside them. In contrast, a deciduous tree requires time to produce new leaves in the spring before it can begin photosynthesizing. When the new leaves are emerging, they are particularly vulnerable to injury from freezing. Thus, leaf emergence does not occur until the chance of a late frost is very small. Following leaf emergence, the growing season must be long enough for the tree to accumulate enough carbon for its maintenance, growth, and reproduction. Thus, deciduous trees are able to survive only in regions with a relatively long growing season.

The slow leafing out of deciduous trees in the spring has another effect on the form of deciduous forests: These communities are much more complex in form and higher in species diversity than boreal forests. Because the canopy is deciduous, there are times of year when large amounts of light penetrate to the forest floor. Consequently, there can be up to four layers of vegetation in deciduous forests: herbaceous species along the ground, small shrubs, small trees and large shrubs, and

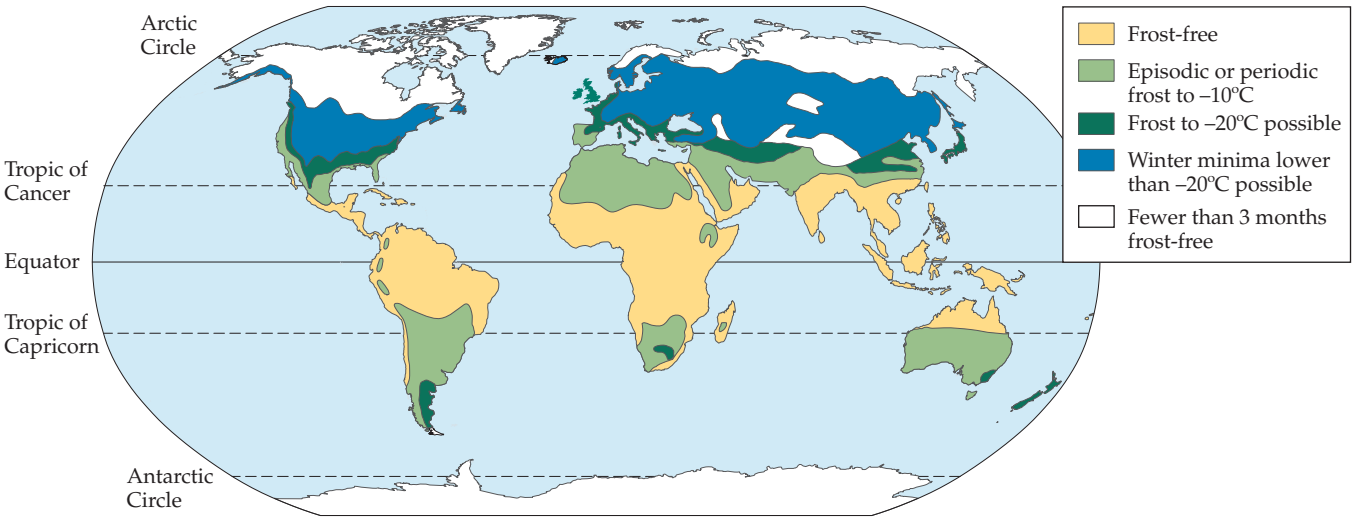



**Figure 18.3** The occurrence of subfreezing temperatures across the globe. (After W. Larcher and H. Bauer. 1981. In *Encyclopedia of Plant Physiology*, Vol. 12A, O. L. Lange et al. [Eds.], pp. 403–437. Springer: Berlin; data from G. Hoffmann. 1960. *Met Abh* 8: Heft 3. Reimer: Berlin.)

canopy trees. In the spring, the forests turn green from the ground up, with each layer leafing out slightly later. This ground-up effect occurs because the air is warmer near the ground, especially early in the season before the canopy leafs out and reduces convective cooling. The diversity of species and form in deciduous forests, therefore, depends on seasonal variation in temperature.

The evergreen needle-leaved trees that dominate the northern coniferous forests pay a cost for their ability to withstand freezing temperatures: They grow much more slowly than broad-leaved deciduous trees. In general, once the growing season is sufficiently long, deciduous trees outcompete needle-leaved evergreen trees. The border between the deciduous forest and the needle-leaved evergreen forest corresponds to a temperature line north of which the minimum temperature during the winter falls below –40°C (Arris and Eagleson 1989) (compare Figure 18.1 and Figure 18.3). Above this temperature and below 0°C, plant tissues can supercool, meaning that they can cool without ice formation. It is the formation of ice crystals in plant and animal tissues that is the primary source of damage and injury from freezing.

## Tree line defines the edge between treed and treeless landscapes

Temperature has a direct effect on the types of vegetation that can grow in a region. One of the most striking examples of this effect is the alpine **tree line**, a boundary where trees are replaced by low-growing vegetation that can be abrubt or diffuse (Figure 18.4) (Harsch and Bader 2011). The tree line is most obvious in high mountain ranges, such as the Sierra Nevada and the Rockies in North America, the Alps in Europe, the Andes in South America, or the Himalayas in Asia, all of which have forests that stop below the summit. If you were to hike up such mountains, you would reach a point where the forest would thin and the trees would get shorter. At some point, the trees would disappear entirely. A similar sight would greet you at the Arctic or Antarctic tree line on a hike toward the North Pole in North America or Eurasia, or toward the southern tip of South America.

Tree lines are caused by a complex of several factors, including the limits to supercooling mentioned above. Air temperatures are critical: alpine tree lines form where the mean growing season temperature falls below 6.5°C and there is a growing season length of about 90 days (Paulsen and Körner 2014). In addition, wood will form only when air temperatures are above 10°C. Microenvironmental differences in temperature can lead to interesting effects, such as the formation of **krummholz** (Figure 18.5), a peculiar tree form found at the tree line on mountains in different parts of the world.

Above the tree line on Mount Washington, New Hampshire—the highest peak in eastern North America—you might notice an odd-looking plant that looks like a horizontal woody vine growing along the ground (Figure 18.6). If you examined it closely, however, you would find that it is *Picea mariana* (black spruce, Pinaceae), the dominant tree species in the nearby forest. Seeds taken from these plants, if planted farther down the mountain, will grow into tall trees! This krummholz form of the plant comes about because the temperature right at ground level is just a little bit warmer than the air above it—just enough above 10°C for wood formation (Teeri 1969). The warmer temperature near the ground is a result of reduced convective cooling and longwave radiation by the ground of the energy it absorbs from sunlight.

## Grasslands and woodlands dominate in areas of lower precipitation

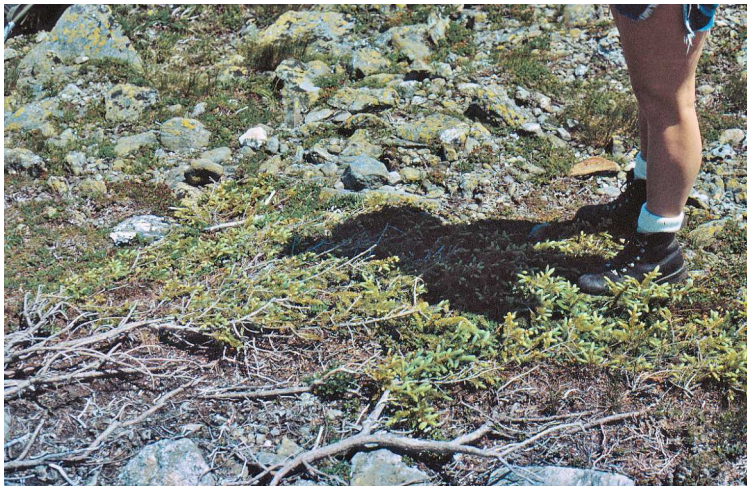
Traveling from west to east, starting in the rain shadow on the eastern face of the Rocky Mountains and

Courtesy of S. Scheiner

**Figure 18.4** Tree line at about 1200 m in Rocky Mountain National Park, Colorado, U.S.A.

Courtesy of S. Scheiner

**Figure 18.5** Krummholz vegetation at tree line in Rocky Mountain National Park, Colorado, U.S.A. Note the twisted trunk of the tree.



Courtesy of J. Teeri

**Figure 18.6** *Picea mariana* (black spruce, Pinaceae) growing in a "vine" form above tree line on Mount Washington, New Hampshire, U.S.A., the highest peak in eastern North America.

traversing the center of the United States, you will see another gradual transition in the physiognomy of the vegetation. Your journey begins in shortgrass prairie, in the midst of low-growing clumps of grasses, with bare patches interspersed between the clumps. As you travel eastward, the vegetation becomes taller and thicker as you reach the mixed-grass prairie. Here you may notice fewer grasses growing in clumps and more rhizomatous grasses, as well as many dicots.

By the time you reach Iowa, you are in tallgrass prairie. In addition to taller vegetation, these communities have greater biomass; they are also more productive and more diverse. Besides the obvious grasses, there are many dicot forbs whose stature matches that of the grasses. This shift in plant stature and diversity closely follows a pattern of increasing rainfall. As you continue eastward, more and more trees are evident. In the areas dominated by grasslands, trees are clustered along streams. Farther east, they begin to dot the countryside, becoming more dense. These areas are **woodlands**, dominated by trees, but without a closed canopy. Eventually, at about the Illinois–Indiana border, you reach forests.

This transition from grassland to woodland to forest is a function of changes in both the amount and seasonality of precipitation. Annual rainfall, for example, is about 80 cm both in the middle of Kansas and to the northeast in the middle of Minnesota, yet the former is in the center of the tallgrass prairie, while the latter is forested. The difference is that in Kansas, the rainfall is much more seasonal, with more of it occurring in spring and fall; the temperature is warmer; and the chance of drought is greater.

This seasonality of precipitation and propensity for drought underly another key factor: fire (see Chapter 13). Under natural circumstances, mixed-grass and tallgrass prairies experience fires every 3 to 5 years, depending on climatic and landform conditions. Each year, dry grasses and other plant materials accumulate until there is enough fuel to sustain a conflagration, and lightning (or human activity) ignites a fire. Extensive fires are rare in shortgrass prairies because there is insufficient plant material to support them. Prairie plants are adapted to these frequent fires. Their vulnerable meristems are located at or below ground level. A fast-moving fire will burn off the tops of the plants, but they quickly resprout, using resources stored below ground. Trees, however, have their meristems at the tips of their branches. Seedlings and young saplings are particularly vulnerable to fire. Grasslands begin to give way to woodlands where the frequency of fire drops low enough that some trees have sufficient time between fires to grow tall enough to become resistant to fire. These woodlands are dominated by species that have thick bark and are tolerant of fire. The woodlands become closed-canopy forests where precipitation levels are high enough to keep fire frequency low.

## Shrublands and deserts are found in very dry or cool regions

Shrublands, another widespread physiognomic form, are dominated by low-growing woody plants with multiple stems. The locations of these communities are determined by a combination of temperature and precipitation effects. They are primarily found in dry to very dry regions. Areas with mediterranean climates, for example, contain a mixture of shrublands and woodlands. Which type of vegetation predominates in an area depends on disturbance. Fire often favors shrubs, which can resprout readily after being burned, over trees. Annuals are also very common in shrublands and are especially prevalent in the years immediately following a fire. Other disturbances can also favor shrubs over trees. Today, grazing by domesticated animals is the most important factor spreading and maintaining shrublands in subtropical regions of the world.

The other major climatic regions with extensive shrublands are **deserts**, which are regions where potential evapotranspiration exceeds actual evapotranspiration. The amount of precipitation and its seasonality are among the most important factors in determining what kinds of plants thrive in very dry environments (Figure 18.7). Summer and winter precipitation represent very different sources of water for plants. At summer temperatures, rain only wets the soil surface, often for only a few days. At winter temperatures, precipitation can infiltrate deeper soil layers and remain available there throughout the year.

Wet winters with dry summers favor deep-rooted, long-lived plants with large leaf areas, such as trees. These plants use almost no summer rainwater, even

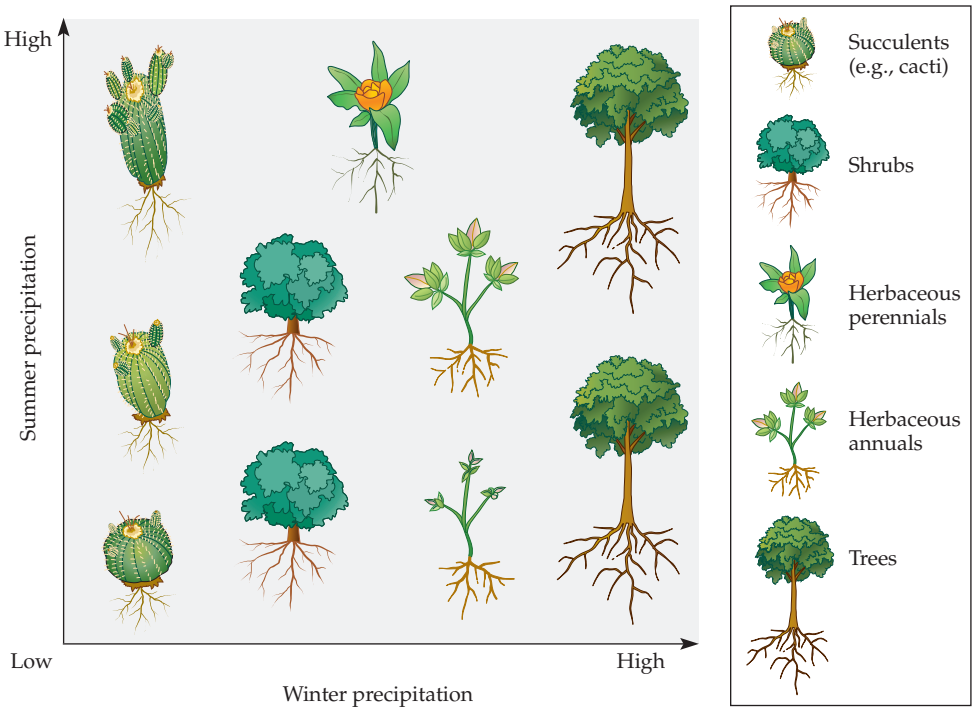



**Figure 18.7** Plant types favored by various combinations of winter and summer precipitation in temperate regions. (After S. Schwinning and J. R. Ehleringer. 2001. *J Ecol* 89: 464–480.)

when it is available. As the amount of winter precipitation declines, plants with greater shoot biomass relative to the amount of root biomass, such as shrubs, are favored. As winter precipitation decreases further, summer precipitation becomes increasingly important, favoring plants with shallower root systems and even less root material, such as herbaceous annuals and perennials. These plants are capable of extracting surface water quickly, before it is lost to evaporation. When there is no winter water stored deep in the soil, only succulent plants such as cacti, which store water internally and lose it slowly, can survive between summer rain events. The more often it rains in summer, the less need to store water, and the community becomes increasingly more herbaceous. If wet summers are accompanied by moderately dry winters, herbaceous, shallow-rooted plants, such as grasses or summer annuals, are favored.

These effects of the amount and seasonality of precipitation in deserts explain the differences in physiognomy among the hot deserts of western North America: the Mojave, the Sonoran, and the Chihuahuan. The Mojave Desert (mainly in California) gets almost exclusively winter precipitation. The Chihuahuan Desert, to its southeast (mainly in the Mexican states of Chihuahua and Coahuila), gets almost exclusively summer precipitation, whereas the Sonoran Desert, lying between the other two in Arizona, U.S.A., and Sonora, Mexico, gets both summer and winter precipitation. As a result, while all three regions have many shrubs, the Mojave Desert is dominated by low-growing shrubs (many of which are deciduous) and has a very large number of annual plant species, the Chihuahuan Desert is dominated by evergreen shrubs and grasses, and the Sonoran Desert has a mixture of many growth forms, including the spectacular columnar cacti and other succulent species. In all of these deserts, however, several growth forms coexist; no one growth form can preempt all of the available water.

## 18.2 Biomes with Similar Vegetation Forms May Be the Result of Convergent Evolution

By definition, a given biome (such as tropical deciduous forest) always contains species that look similar (such as deciduous trees). This similarity of form comes about through three processes, one ecological, one evolutionary, and one at the intersection of ecology and evolution. The first two both lead to similar patterns, but for different reasons. Under ecological **species sorting**, the same or closely related species are found in a given biome because they are well suited to that environment. For example, palms may be common in tropical woodlands around the world because they are well suited to the warm climate and are fairly drought tolerant. On the other hand, related species may occur in similar environments because they originated and have diversified there. For example, many hot deserts in North and South America contain cacti because the family Cactaceae arose in North America.

The process that combines ecology and evolution is **convergent evolution**, the independent evolution of similar features in unrelated taxa. Generally, these similar features come about through natural selection, which is driven by the ecology of the organisms (see Chapter 9). A striking example of convergent evolution is found among alpine plants in the families Campanulaceae and Asteraceae. *Lobelia rhynchopetalum* (Campanulaceae) grows in the mountains of Ethiopia, and *Espeletia pycnophylla* (Asteraceae) grows in the Andes of South America. They are both unusual-looking plants, standing up to 2 m tall, with a thick stem topped by a rosette of leaves (Figure 18.8). Their height is an adaptation to large daily temperature swings. At night, the temperature near the ground can drop below freezing because the ground loses large amounts of radiant energy to the very cold night sky. But the upper parts of the plants, at 2 m above the ground, stay above freezing due to longwave radiant energy input from the ground (see Chapter 3 and Chapter 16). Thus, the sensitive flowers and meristems never experience freezing temperatures. Both of these species evolved from herbaceous, ground-dwelling species growing at lower elevations.

While a given biome shows broad similarities of form in different places, the characteristics of the biome in one location may differ in detail from those of that same biome in another location due to evolutionary and historical processes. Consider the previously mentioned example of temperate evergreen forests. The reason that these forests are dominated by broad-leaved eucalypts only in Australia is that the genus *Eucalyptus* arose during the Cretaceous on the ancient continent that included Australia (see Figure 17.6B) and never dispersed far outside that continent (Ladiges et al. 2003). So biogeographic history can create both similarities and differences among biomes. One useful aspect of the biome concept is that it provides a framework for analysis and separation of the effects of species sorting, convergent evolution, commonality of descent, and biogeographic history.

Convergent evolution is an important piece of evidence that a given trait is an adaptation due to natural selection. One way to determine whether this is the case is to ask the question, Do we find the same form repeated in similar environments? Heavy-metal tolerance, for example, has been found repeatedly in plant populations growing on mine tailings in different locations (see Chapter 9). But how do we decide about adaptations if an entire species, or group of species, shows one form? In the Cactaceae, all species have CAM photosynthesis (see Chapter 2). Almost all cacti also grow in dry habitats or microhabitats. Is CAM photosynthesis an adaptation to low water availability? One alternative explanation is that plants in the Cactaceae all have CAM photosynthesis not because it is adaptive, but only because their ancestors had CAM photosynthesis. The adaptation of species in the Cactaceae to dry conditions might have nothing to do with their possession of CAM photosynthesis; this association could just be a coincidence.

How do we separate such chance associations from those due to adaptation by natural selection? We look for repeated associations such as convergent evolution. Because plants in many unrelated families found in arid habitats or microhabitats have CAM photosynthesis, and because CAM photosynthesis evolved independently in

(A)

© M. Harvey/Alamy Stock Photo

(B)

©Roland Knauer/Alamy Stock Photo

**Figure 18.8** An example of convergent evolution: two similar-looking alpine tundra species. (A) *Lobelia rhynchopetalum* (Campanulaceae), found in the Bale Mountains of Ethiopia. (B) *Espeletia pycnophylla* (Asteraceae), found in the Andes of South America.

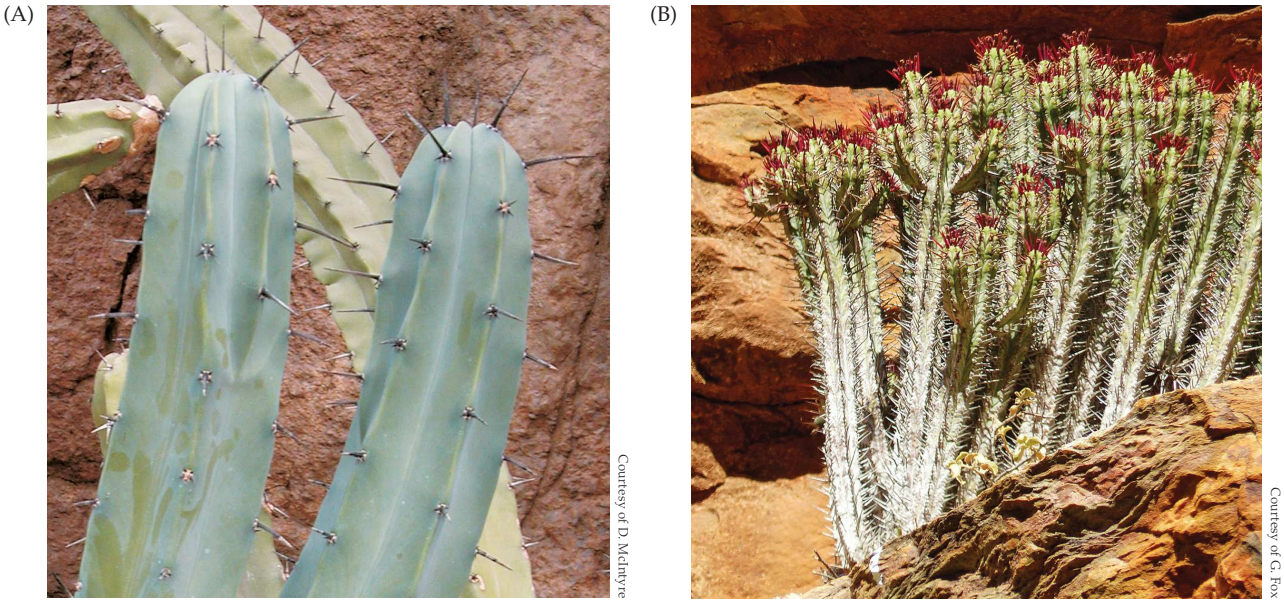



**Figure 18.9** An example of convergent evolution: two similar-looking desert-dwelling species. (A) *Myrtillocactus geometrizans* (blue candle, Cactaceae), found in the deserts of Mexico. (B) *Euphorbia enopla* (Euphorbiaceae), found in South Africa.

many of them, we begin to be convinced that CAM photosynthesis offers adaptive advantages in these environments. Because we know the physiological mechanisms involved, we understand that CAM photosynthesis results in restricted water use and greater water use efficiency in plants that have this photosynthetic pathway. These multiple lines of evidence make it more believable that CAM photosynthesis is an adaptation to dry conditions.

Another example of convergent evolution is found in the morphology of some of the cacti of the deserts of North and South America and plants in the family Euphorbiaceae growing in arid regions of southern Africa and south Asia. Consider *Myrtillocactus geometrizans* and *Euphorbia enopla*; these two plants look remarkably alike (Figure 18.9). Both are leafless, have green photosynthetic stems, are columnar in shape, have spines along the ribs of the stem, and have CAM photosynthesis. These species are not closely related, so their similar features must be due to independent evolutionary events. Other members of the genus *Euphorbia*, far more closely related to *E. lactea*, look very different; the genus includes plants with a wide variety of growth forms, ranging from small herbaceous annuals to shrubs to large plants (the widely cultivated poinsettia, *E. pulcherrima*, is a member of this genus). Another way we know that the striking similarities between *M. geometrizans* and *E. enopla* are a result of convergent evolution is that some of the features have different developmental origins. The spines on cacti evolved from leaves, which can still be seen on primitive cacti. In contrast, the spines on *E. enopla* evolved from **stipules** (appendages to petioles). The independent evolution of so many similar features by chance alone is very unlikely. Therefore, we can be confident in our conclusion that these features are adaptations to conditions in desert environments.

The first person to recognize what we now call biomes was Alexander von Humboldt (Figure 18.10), an eighteenth- and nineteenth-century explorer who is one of the founders of the discipline of biogeography (von

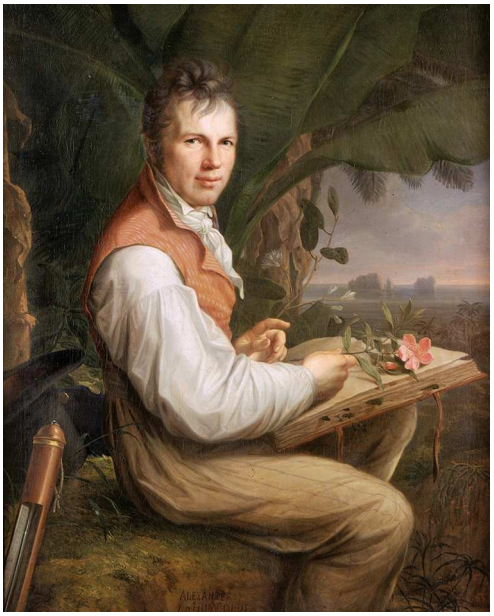


**Figure 18.10** Portrait of Alexander von Humboldt by Friedrich Georg Weitsch, 1806 (located in the Alte Nationalgalerie [Old National Gallery], Berlin, Germany).

Humboldt 1808). The first classification of biomes was in his book *Cosmos* (von Humboldt 1847). Humboldt recognized not only the great patterns of worldwide vegetation, but also the similarity of those patterns along latitudinal and elevational gradients and that they were caused by similar gradients of temperature and precipitation. In the remainder of this chapter, we look at these global patterns of vegetation as we examine the characteristics of the biomes.

## 18.3 Moist Tropical Forests

### Tropical rainforest

Tropical rainforest is one of the most diverse and productive biomes on the planet. Although it makes up only 11% of the Earth's land area, this biome accounts for over 30% of terrestrial net primary production (Table 18.2) and contains perhaps half of all living species. Its tree species diversity can be as high as 300 species per 0.1 hectare. Animal diversity is also high, especially that of insects and other invertebrates.

Tropical rainforests are found within a band extending between roughly 10°N and 10°S latitude at elevations below 1 km, in areas that get more than 250 cm of rainfall each year, with some rain falling in all months, although varying throughout the year. The most extensive rainforests are found in the Amazon basin of South America, the equatorial regions of western Africa, and parts of Southeast Asia. Moving away from the equator, precipitation becomes more seasonal, and this biome grades into tropical semievergreen forests dominated by broad-leaved trees, some of which retain their leaves year-round and others that lose their leaves seasonally.

Tropical rainforests are characteristically multilayered, with two, three, or even four layers of trees (Figure 18.11). Tallest are the **emergent trees**, which are widely spaced and often more than 40 m in height, with umbrella-shaped canopies that extend above the general canopy of the forest. Below them is a closed canopy of trees, typically 30–40 m high. Light is readily available at the top of this layer but greatly reduced below it. Under the closed canopy, there may be another layer of trees less than 20 m in height. Under these trees, there is often a shrub/sapling layer, and finally a sparse ground layer.

Because of these multiple layers, the floors of tropical rainforests can be very dark, with less than 1% of above-canopy light levels. While Hollywood movies may portray tropical rainforests as jungles with a tangle

(A)

© D. Perlman/EcoLibrary.org

(B)

Courtesy of J. Thomson

**Figure 18.11** (A) The interior of the Rara Avis Rainforest in Costa Rica. (B) Epiphytes are symbiotic commensals (see Box 11A) that grow on the surfaces of other plants and are common in most tropical rainforests. Numerous epiphytic bromeliads (Bromeliaceae) can be seen on the trunks of several of the trees in this forest in Puerto Rico.

**TABLE 18.2** Primary productivity of selected major biome types

| Biome | Area ($\times 10^6$ km$^2$) | | Net primary productivity per unit area (g/m$^2$/yr) | | Global net primary productivity ($10^{12}$ g/yr) | | Biomass per unit area (kg/m) | | Total global biomass ($10^{12}$ g) | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Amount | % of total terrestrial area | Normal range | Mean | Amount | % of total terrestrial productivity | Normal range | Mean | Amount | % of total terrestrial biomass |
| Tropical rainforest | 17.0 | 11.4 | 1000–3500 | 2200 | 37.4 | 32.5 | 6–80 | 45 | 765 | 41.7 |
| Tropical deciduous forest | 7.5 | 5.0 | 1000–2500 | 1600 | 12.0 | 10.4 | 6–60 | 35 | 260 | 14.2 |
| Temperate deciduous forest | 7.0 | 4.7 | 600–2500 | 1200 | 8.4 | 7.3 | 6–60 | 30 | 210 | 11.4 |
| Temperate evergreen forest | 5.0 | 3.4 | 600–2500 | 1300 | 6.5 | 5.6 | 6–200 | 35 | 175 | 9.5 |
| Woodland and shrubland | 8.5 | 5.7 | 250–1200 | 700 | 6.0 | 5.2 | 2–20 | 6 | 50 | 2.7 |
| Taiga | 12.0 | 8.1 | 400–2000 | 800 | 9.6 | 8.3 | 6–40 | 20 | 240 | 13.1 |
| Temperate grassland | 9.0 | 6.0 | 200–1500 | 600 | 5.4 | 4.7 | 0.2–5 | 1.6 | 14 | 0.8 |
| Tropical savanna | 15.0 | 10.1 | 200–2000 | 900 | 13.5 | 11.7 | 0.2–15 | 4 | 60 | 3.3 |
| Desert | 18.0 | 12.1 | 10–250 | 90 | 1.6 | 1.4 | 0.1–4 | 0.7 | 13 | 0.7 |
| Tundra | 8.0 | 5.4 | 10–400 | 140 | 1.1 | 1.0 | 0.1–3 | 0.6 | 5 | 0.3 |
| Extreme desert, rock, sand, and ice | 24.0 | 16.1 | 0–10 | 3 | 0.07 | 0.1 | 0–0.2 | 0.02 | 0.05 | 0.03 |
| Cultivated land | 14.0 | 9.4 | 100–3500 | 650 | 9.1 | 7.9 | 0.4–12 | 1 | 14 | 0.8 |
| Swamp and marsh | 2.0 | 1.3 | 800–1500 | 2000 | 4.0 | 3.5 | 3–50 | 15 | 30 | 1.6 |
| Lake and stream | 2.0 | 1.3 | 100–1500 | 250 | 0.5 | 0.4 | 0–0.1 | 0.02 | 0.05 | 0.003 |
| Total terrestrial | 149.0 | | | 773 | 115 | | | 12.3 | 1837 | |
| Total world | 510.0 | | | 333 | 170 | | | 3.6 | 1841 | |

*Source*: After R. H. Whittaker and G. E. Likens. 1975. In *Primary Productivity of the Biosphere*, H. Lieth and R. H. Whittaker (Eds.), pp. 305–328. Springer: Berlin.

of vines and understory plants, the understory is generally sparsely vegetated with highly specialized plants capable of tolerating low light levels (see Chapter 2). The interior of a rainforest was compared by early European explorers to a dimly lit, soaring green cathedral with an expansive space reaching to a high, vaulted roof. "Jungles" are generally the consequences of disturbed, high-light environments on the margins of the forest, along roads and rivers, and in gaps or where logging has thinned the canopy.

This vertical light distribution within the rainforest results in much of the productivity being concentrated high in the canopy. Plants with growth forms that are not present in most temperate forests greatly contribute to the diversity of tropical forests. **Lianas** are woody vines that are rooted in the ground. **Epiphytes**, which grow on other plants and are not rooted in the ground (Figure 18.11B), include many species of ferns, orchids, and bromeliads. Lianas and epiphytes are common in tropical forests because of the very low light levels near the ground. They use an alternative strategy to being a tree—grow high up on a tree to gain access to light.

Most epiphytes and lianas have commensal relationships (the host neither benefits nor is harmed) with the trees on which they grow unless the burden of these plants becomes too large, weakening the tree and possibly pulling it down. Some species are not so benign, however. Strangler figs (*Ficus* spp., Moraceae) begin as either a single individual or multiple, separate individual plants growing on a host tree. They send roots downward to the forest floor, develop thick trunks and canopies, and gradually envelop and kill their host (see Figure 10.3A). In some instances, the separate individual plants constituting the strangler can ultimately merge to become a functional individual tree. Thus, this plant has a combined strategy of beginning life high in the canopy as an epiphyte, becoming rooted in the ground, which allows more rapid growth, and eventually becoming a tree and outcompeting other species for light.

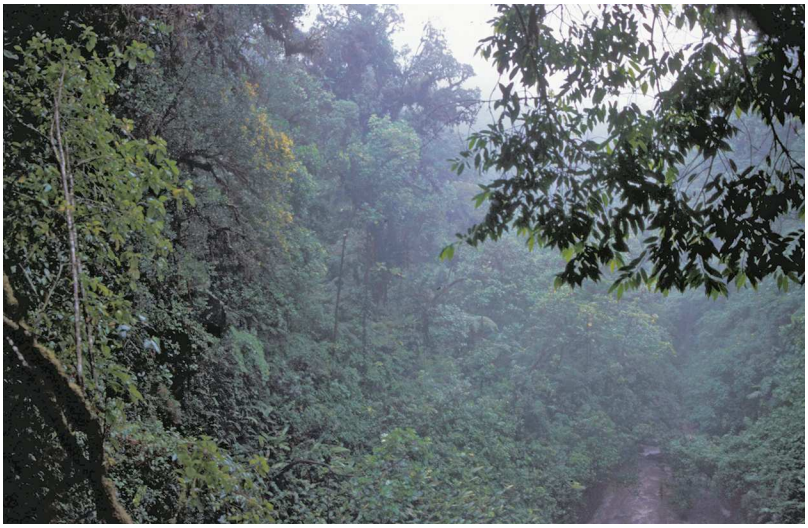
Moist tropical forests typically have highly infertile soils. In tropical rainforests, warmth and abundant moisture contribute to an extraordinary amount of respiration in the soil—not only by roots, but also by many kinds of soil dwellers, particularly decomposing microorganisms. When $CO_2$ is dissolved in water, it produces carbonic acid (that is why natural rainwater is always slightly acidic). The huge amount of soil respiration in tropical rainforests releases large amounts of $CO_2$, much of which goes into solution in the soil. The net effect is that negatively charged nutrient ions like nitrate ($NO_3^-$), which are not adsorbed onto negatively charged clay and organic matter soil particles, enter the soil solution. As a result, they are either taken up almost instantaneously by plants or bacteria, or washed away by rainfall (see Chapter 4).

If their soils are so low in nutrients, how is it that tropical rainforests are so productive? Factors contributing to their high productivity are year-round water availability, warm temperatures, and long days that vary little throughout the year. As for the mineral nutrients needed for growth, there are large quantities of nutrients in these forests, but they are mainly found in the living biomass, rather than in the soil. Because decomposition is extremely rapid in these environments, nutrients cycle very rapidly. Like plants in many areas where nutrients are limiting, those in tropical forests are very good at acquiring and retaining nutrients. Particularly important in this regard are mycorrhizal associations (see Chapter 11).

Economic, population, and political pressures have contributed to the destruction of a large proportion of the world's rainforests (see Chapter 19). Exploitation comes not only from within the tropical countries themselves, but also in great measure from individuals and corporations based in North America, Europe, China, Australia, New Zealand, and Japan. Today all tropical rainforests are threatened by logging, and even more by the clearing of land for a range of purposes, including subsistence slash-and-burn agriculture, corporate-owned cattle ranching, and massive palm oil plantations.

Logging and the clearing of land for cattle grazing, palm oil plantations, or human habitation can have numerous effects on rainforests. Because most of the nutrients are tied up in biomass, logging, burning, and agriculture remove substantial amounts of nutrients from the system. The remaining soils are likely to lack the nutrients required for plant growth, and depending on the soil texture, with the plant roots gone, they may quickly achieve a cement-like texture or become severely eroded. These changes can make agriculture impossible in a very short time—in other words, the purposes for which the forest was removed are not sustainable. Removal of rainforest from large enough areas can also drastically affect regional rainfall levels. When the forest is present, rainfall is quickly taken up and transpired back into the atmosphere. When the forest trees are no longer there, the rainfall runs off, which leads to increasing erosion, particularly on slopes, with great reductions in the water vapor returned to the atmosphere and a slowing of the regional water cycle (see Chapter 16). Even when the destruction is not so dramatic, as in selectively logged forests, non-native plant species can invade disturbed forests, threatening the regeneration of native species. For example, Kerry Brown and Jessica Gurevitch (2004) concluded that logging of tropical rainforests in Madagascar has facilitated the establishment of invasive species, which suppress native species, resulting in reduced diversity of native trees even 150 years after the logging was carried out.

While habitat destruction is the key factor threatening tropical biodiversity, other kinds of human exploitation—from the illicit trade in rare species such as orchids and parrots brought into the United States, Europe, and China, to the hunting of chimpanzees and other primates for bushmeat for African cities—also endanger animal and plant species within these forests. These forms of forest destruction are not particularly new; wealthy nineteenth-century orchid collectors in England and Germany destroyed entire tropical forests to ensure that their



Courtesy of S. Scheiner

**Figure 18.12** A tropical montane forest in the La Paz Waterfall Gardens Nature Park, Costa Rica.

competitors could not collect orchids from them. However, the scale of the damage occurring today dwarfs that of the past, and, of course, the destruction is cumulative.

### Tropical montane forest

Tropical montane forests—also called elfin forests or cloud forests at the highest elevations—are higher-elevation neighbors of rainforests (Figure 18.12). They are called cloud forests because they are often swathed in clouds, and much of the moisture available to plants comes from condensation from fog. They are cooler than tropical rainforests because they are at higher elevations. The trees are typically shorter, growth is slower, and there are many more epiphytes festooning the branches of trees. Overall diversity is lower than in rainforests, however, because there are fewer lianas. These forests are found at middle elevations in the mountains of Africa, South America, Central America, Mexico, and New Guinea.

## 18.4 Seasonal Tropical Forests and Woodlands

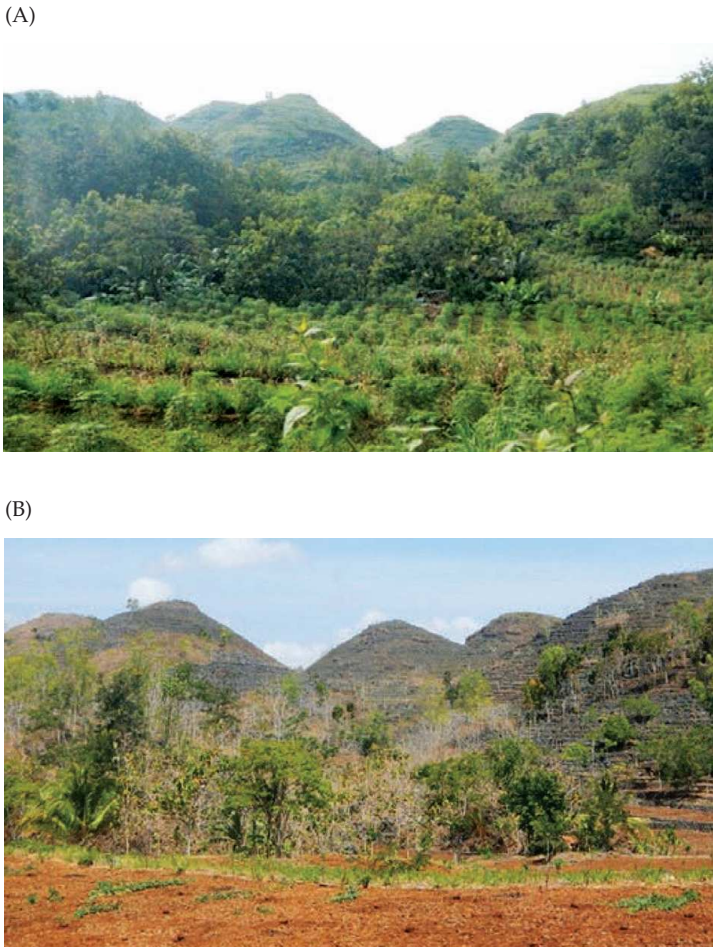
Many tropical regions are considerably drier than rainforests, and many of these forests are also rich in species. Seasonal tropical forests, with their associated species, are being lost very rapidly. In fact, these forests are disappearing—or have already disappeared—even more rapidly than tropical rainforests, as they are often easier to clear and are on soils that are much better for farming. Dry tropical regions are home to a range of biomes that grade from deciduous forest to thorn forest and woodland.

Several geographic factors determine the locations of seasonal tropical forests and woodlands. First, they tend to occur relatively close to the Tropics of Cancer and Capricorn, rather than near the equator. This pattern is a result of the shifting throughout the year of the **intertropical convergence zone (ITCZ)**—the latitude at which the Hadley cells meet on Earth's surface and the trade winds converge (see Figure 16.10). At the ITCZ rainfall is concentrated by the adiabatic cooling of air rising to the top of the troposphere, driven by the energy of the direct rays of sunlight at this latitude. The ITCZ usually passes over these latitudes only once a year, whereas it passes over the equator twice a year. As a result, rainfall in these latitudes is more seasonal, whereas rainfall closer to the equator is more evenly spread throughout the year (see Figure 16.19). Second, dry forests and woodlands often occur in the rain shadows of mountain ranges (see Figure 16.14) and on the western slopes of continents. For example, the eastern (Caribbean) side of the mountain range in Costa Rica is tropical rainforest, whereas the western (Pacific Ocean) side is tropical deciduous forest.

### Tropical deciduous forest

Tropical deciduous forests, also called tropical dry forests, include a range of community types, from wholly deciduous to semievergreen (Figure 18.13). The trees are **drought-deciduous**, meaning that they lose their leaves during the dry season. These communities are found in tropical areas that have pronounced wet and dry periods, and they were once especially extensive in India and Southeast Asia; they also occur in Central and South America and elsewhere. Depending on the latitude, there might be a single wet and a single dry season, or two of each (Figure 18.14). Although most photosynthetic activity and plant growth occurs during the wet season, the dry season is often when flowering takes place. There may be advantages to flowering in the dry season for both animal-pollinated and wind-pollinated plants. The lack of leaves makes the flowers more conspicuous to pollinators, and the pollen and nectar rewards may be one of the few food sources available to them. Similarly, wind-dispersed pollen can travel with fewer obstructions.

Tropical deciduous forests are similar in overall form and structure to rainforests, with a species richness lower than but approaching that of tropical rainforests in some areas. Localities with lower amounts of rainfall generally have shorter canopies and fewer




E. Haryono et al. 2016. *Acta Carsologica* 45: 173–185. DOI: 103986/ac.v4512.4541. Courtesy of Sutanto Trijuni Putro.

**Figure 18.13** (A) A tropical deciduous forest bordering on agricultural lands on the island of Java, Indonesia, during the rainy season. (B) The same view in the dry season, when many of the trees are leafless.

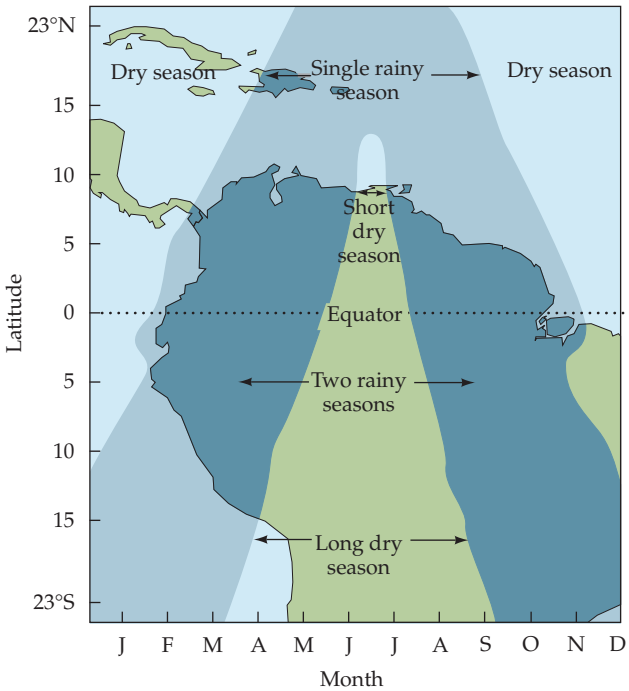



**Figure 18.14** The length and number of rainy seasons depends on latitude. Near the equator, there are two wet and two dry seasons each year, while farther north and south there is a single rainy season in summer and a single dry season in winter. The timing of the seasons is reversed north and south of the equator.

species. Human populations are often very high in these biomes, and they face many of the same threats as tropical rainforests.

## Thorn forest

Drier tropical areas support thorn forests (also called thornwoods), named for the menacing armaments of many of the plant species there (Figure 18.15). These armaments are generally the result of convergent evolution. Thorn forests are common from South America to Mexico, including those on limestone soils in the West Indies and Central America, as well as in Asia (Myanmar, India, and Thailand), continental Africa, and Madagascar. They have pronounced wet and dry seasons, and the total yearly rainfall may be half that in tropical deciduous forests. Species of *Acacia* and other woody members of the pea family (Fabaceae) often dominate these communities. These plants have leaves that are divided into many small leaflets, which helps to minimize water loss. Diversity is considerably lower than in tropical rainforests and tropical deciduous forests but can still be quite high compared with that in temperate systems. As in tropical deciduous forests, the trees are typically drought-deciduous. Succulents are also common, especially in drier areas. As the climate becomes even drier, these communities grade into either desert scrublands or tropical savannas.

## Tropical woodland

Forests are dominated by trees whose canopies touch each other. In contrast, the trees in woodlands are spaced apart, although the density of those trees can vary from closely spaced to sparse. Tropical woodlands are found in areas with strongly seasonal climates that are drier or have less favorable soils than areas with tropical deciduous forests. Tropical woodlands occupy areas of Brazil and interior southern Africa (miombo) and also occur in northern South America, the West Indies, northern Australia (Figure 18.16), and Myanmar and other parts of Southeast Asia.

© Nick Garbutt Natural Selection

**Figure 18.15** A thorn forest on the island of Madagascar, off the east coast of Africa. This photograph was taken after heavy rains in December. From August through November—the dry season—the trees are parched and brown.

The trees in tropical woodlands typically range in height from 3 to 7 m. Most trees and shrubs have large, tough, long-lived, sclerophyllous semievergreen leaves and thick, fire-adapted bark. Many of the trees are members of the Fabaceae, as in thorn forests. Palms are also common, whereas succulents and spiny plants are rare. African tropical woodlands—as well as savannas—are rich in large mammal species, including herds of elephants, giraffes, wildebeests, and their predators, such as lions and hyenas (see Figure 11.18).

Courtesy of S. Scheiner

**Figure 18.16** A tropical woodland in Uluru-Kata Tjuta National Park, Australia.

## 18.5 Temperate Deciduous Forest

Temperate deciduous forests are found in three disjunct areas of the Northern Hemisphere: eastern North America, eastern Asia, and Europe (Figure 18.17). These forests grow on young soils formed since the most recent glaciation; the reason for the lack of such forests in the Southern Hemisphere is unclear but may be related to the much smaller available area in southern South American, Australia, and New Zealand. Temperate deciduous forests are found in areas with continental or maritime climates with warm, moist summers and cold, sometimes snowy winters, with a range of 50–250 cm of rainfall annually, distributed throughout the year. Diversity and productivity span a fairly wide range but are generally moderate. In North America and Asia, this biome is known for the brilliant red, orange, and gold colors of its leaves in autumn, before they are shed for winter. The dominant trees are 18–30 m tall. Growing seasons range from very short in the northernmost and high-altitude parts of this biome to very long in its southern portions.

The genera of the dominant trees common to temperate deciduous forests on all three continents include *Quercus* (oak, Fagaceae), *Acer* (maple, Sapindaceae), *Fagus* (beech), *Castanea* (chestnut, Fagaceae), *Ulmus* (elm), *Tilia* (basswood or linden, Tiliaceae), *Juglans* (walnut), and *Liquidambar* (sweet gum). This sharing of genera among different continents is due to dispersal of these genera across the Northern Hemisphere and is an example of biogeographic history shaping biome composition. Different species of these genera occur on each continent. The diversity of local forest types on all three continents may be very great, depending on the details of topography, climate, soils, and disturbance regimes. In North America, for instance, the eastern deciduous forest consists of many different forest subtypes (Braun 1950; Greller 1988).

At higher latitudes and elevations, temperate deciduous forest sometimes grades into mixed deciduous-evergreen forest. An example is the hemlock–white pine–northern hardwood forest that is found in the northern United States (especially around the Great Lakes) and extending south in the upper elevations of the Appalachian Mountains. The dominant trees in this forest are a mixture of evergreen needle-leaved and deciduous broad-leaved species. Elsewhere in the temperate deciduous forest biome, patches of pine-dominated communities can be found where the soil is especially shallow or infertile or there is a high frequency of fire. If such areas are large enough, they may be separately classified as temperate evergreen forest.

Tree species diversity in temperate deciduous forests is generally moderate, with 5–30 dominant tree species in any region. However, the

temperate deciduous forests of the southeastern United States are a center of plant diversity, with the highest tree species diversity in North America as well as a very high diversity of ferns and deciduous and shrubby angiosperms. For example, there are over 120 species of trees and over 3000 plant species in the southern Appalachian region. These forests escaped glaciation, serving as a refuge for many species during the last ice age (see Chapter 17), and some areas have been continuously vegetated for as long as 200 million years. The dominant genus here is *Quercus* (oak, Fagaceae). Other common canopy trees include numerous species of *Carya* (hickory, Juglandaceae), *Acer* (maple, Sapindaceae), *Aesculus* (buckeye or horse chestnut, Sapindaceae), and *Magnolia* (Magnoliaceae), as well as *Tilia americana*, *Liriodendron tulipifera* (tulip poplar, Magnoliaceae), and *Betula lenta* (black birch, Betulaceae). Common in the understory are trees such as *Cornus florida* (flowering dogwood, Cornaceae) and large shrubs such as *Kalmia latifolia* (mountain laurel, Ericaceae) and *Rhododendron* spp. (evergreen rhododendrons and deciduous azaleas, Ericaceae). In wetter areas such as coves or stream banks, *Tsuga canadensis* (eastern hemlock, Pinaceae) is found, while pines are common on drier ridges. The formerly dominant *Castanea dentata* (see Figure 11.24) was devastated by chestnut blight during the early part of the twentieth century and has largely been replaced by various oak species in its former habitats.

Temperate deciduous forests are threatened by a variety of factors. Some of the world's largest cities (such as London and New York) occur on land that was once temperate deciduous forest, although large areas of this biome still exist. In Europe, air pollution, including acid precipitation and nitrogen deposition, has been a major cause of physiological stress in forest trees and general forest decline (see Chapter 19). Worldwide, logging has been a major cause of forest destruction for hundreds of years, as has development for farmland or residential use. In some places the forests have been destroyed, while in others, such as the eastern United States, secondary forests have regenerated over large regions. In some areas of the world (such as parts of Japan and Mexico), the only forests to escape cutting occur in places that are inaccessible or very steep. Road building and expanding suburbanization are renewing threats to many secondary forests in North America. Invasion by exotic plant species is a threat to this and many other biomes worldwide.

Courtesy of S. Scheiner

**Figure 18.17** A temperate deciduous forest in Michigan, here dominated by American beech (*Fagus grandifolia*, Fagaceae) and sugar maple (*Acer saccharum*, Sapindaceae).

## 18.6 Other Temperate Forests and Woodlands

### Temperate rainforest

Temperate rainforests are found in southern Chile and along the northwestern coast of North America, from northern California through British Columbia, Canada, and up into Alaska (Figure 18.18). Smaller areas of temperate rainforest occur on the western coast of New Zealand, on the southeastern and southwestern coasts of mainland Australia, and on the island of Tasmania.

Courtesy of S. Scheiner

**Figure 18.18** An old-growth temperate rainforest in the Muir Woods National Monument, California, U.S.A.

© M. Boulton/Alamy Stock Photo

**Figure 18.19** The Highlands of Scotland were denuded of their temperate rainforests centuries ago and maintained as grasslands by the grazing of cattle and sheep. The forest stand seen here is the result of a reforestation effort begun several decades ago.

Fragments remain in Norway; in other parts of Europe, such as Scotland and Ireland, these forests were once extensive but were completely destroyed several centuries ago (Figure 18.19). Temperate rainforests are all similar climatically: Under maritime influences, temperatures stay cool, and there is abundant winter rainfall and summer fog. Productivity and decomposition are extremely high in some of these forests, largely because of the mild temperatures and extended rainy period, which result in a very long growing season. They differ from tropical rainforests by having cool temperatures and low species diversity. In North America, these forests are dominated by needle-leaved conifers. In contrast, in Chile, New Zealand, and Australia, temperate rainforests are dominated by broad-leaved trees, such as *Nothofagus* (Nothofagaceae), although conifers such as those in the family Podocarpaceae also play an important role in some temperate rainforests in the Southern Hemisphere.

Temperate rainforests contain some of the tallest trees in the world, including the giant redwoods (*Sequoia sempervirens*, Cupressaceae) of California and some of the eucalypts of Australia, with heights exceeding 100 m. Not all forests of this type contain trees of such enormous stature, however; the coastal Sitka spruce forests of Alaska have trees similar in height to those of temperate deciduous forests. Some of these forests (including the redwood forests) depend on occasional fires for regeneration.

As the smallest of the global biomes, temperate rainforests are among the most threatened ecological systems, because of their economic value in many places, and by human development in some others. Because of their enormous standing crop of wood, many of these forests on the western coast of Canada and the United States have been clear-cut for lumber, resulting in high rates of erosion and losses of species diversity. Because of the high potential growth rates of the trees, some of these areas have been replanted with tree monocultures for lumbering, with the result being low-diversity tree plantations in place of natural old-growth temperate rainforest. The remaining stands of old-growth forest have been the focus of extended and bitter fights between environmentalists and logging interests. Old-growth temperate rainforests in North America are disappearing more rapidly than tropical rainforests, almost entirely as a result of commercial lumbering operations (Figure 18.20).

Courtesy of S. Hillebrand/U.S. Fish and Wildlife Service

**Figure 18.20** This old-growth temperate rainforest in Oregon is being clear-cut. All the trees will be felled for commercial lumber.

## Temperate evergreen forest

Temperate evergreen forest is a catchall category that includes forests of very different types. All of them are located in areas that have pronounced seasonal climates and are drier than areas with temperate deciduous forests, due to either low rainfall or soils that do not retain moisture well (Figure 18.21). These forests generally experience relatively mild temperatures but have strong wet-dry cycles (such as in the wetter parts of the Mediterranean region, southern California, and southern Australia) or have greater temperature ranges and drier continental climates

(such as in the mountains of the western United States). Temperate evergreen forests often grade, at their drier margins, into temperate woodlands or shrub-dominated communities such as chaparral. The trees can be broad-leaved or needle-leaved, but their leaves are retained throughout the year. Trees with **sclerophyllous** (tough, evergreen, long-lived, and thick) leaves are common, especially in areas with milder temperatures. In cooler regions, needle-leaved forests often predominate, including the pine and fir forests at middle altitudes in the Sierra Nevada of California and the mixed needle-leaved and broad-leaved forests of conifer and beech in Chile and New Zealand. These Southern Hemisphere forests might be classified with temperate deciduous forests, depending on the relative proportions of needle-leaved evergreen trees and broad-leaved deciduous trees.

Needle-leaved evergreen forests are found in areas of the eastern United States and western Europe where low soil fertility, or low soil water due to high sand content or shallow soil on rocky outcrops, favors pines over deciduous trees. Examples are the "pine barrens" and "piney woods" of the eastern and southern coastal plains of the United States. These forests are dominated by pitch pine (*Pinus rigida*) in parts of the eastern coastal plain and on rock outcrops along the East Coast, and by longleaf pine (*P. palustris*), slash pine (*P. elliottii*), or loblolly pine (*P. taeda*) from Virginia to East Texas, along with various tree and shrub oaks (*Quercus* spp.) and low-growing ericaceous shrubs. All of these forests experience frequent fires and, to varying extents, depend on fire for their maintenance and regeneration. Many of the species in these forests have numerous adaptations to fire (see Chapter 13). Productivity and decomposition rates are typically low in these forests. Diversity can vary extensively; while many have low diversity, the understories of longleaf pine forests include some of the most species-rich systems in North America.

In the southeastern United States, pine-dominated forests are currently more extensive than they were several hundred years ago. Areas that were originally broad-leaved forests were cleared for farming, which depleted the soil of nutrients. When the farms were abandoned during episodic convergences of bad economics and bad weather, from the nineteenth century to the Great Depression of the 1930s, the forests returned, but they are now dominated by pines and differ substantially from the original forests. Today, because various pine species can be grown quickly for pulp, many pine forests are maintained in this state by logging and forestry practices, with none of the natural diversity or structure remaining. Periodic fires in unmanaged areas may also help to maintain pine dominance. The remaining natural pinelands, especially those dominated by *Pinus palustris*, are among the most endangered ecosystems in the United States. Many of these areas are now being urbanized.

Courtesy of S. Scheiner

**Figure 18.21** In their natural state, temperate evergreen forests are one of the most endangered habitats in the world. Shown here is Hartwick Pines State Park in central Michigan. The last remaining stand of old-growth white pine forest in Michigan, it is home to the largest white pine tree (*Pinus strobus*, Pinaceae) in the world.

## Temperate woodland

Temperate woodlands, like their tropical counterparts, occur on the drier margins of forests and often form transition zones between forests and either grasslands or shrublands. Diversity and productivity are lower in woodlands than in deciduous forests. Temperate woodlands range from having nearly continuous tree canopies to being very open with scattered trees (Figure 18.22). The tree species can range from needle-leaved evergreens to broad-leaved sclerophylls to broad-leaved deciduous trees with heights from 3 to 7 m. In North America, coniferous woodlands are extensive in the area between the Rocky Mountains and the Sierra Nevada, where they are dominated by several pinyon pine (*Pinus*) and juniper (*Juniperus*) species. Oak (*Quercus*) and oak-pine woodlands occur from Washington State to the Central Valley of California and across areas of the southwestern United States and northern Mexico, east to Texas. They are also common around the Mediterranean Sea (where they are again oak dominated) and in South America, Africa, and Australia.

Courtesy of S. Scheiner

**Figure 18.22** A pinyon-juniper woodland in northern Arizona, U.S.A.

## 18.7 Taiga

The taiga is a broad swath of primarily coniferous forest that covers an enormous region of the circumpolar subarctic (Figure 18.23). Similar vegetation is found in many subalpine regions globally. The name *taiga* comes from the Russian word for "boreal forest," which is another name for this biome. The taiga encompasses approximately 12 million km$^2$, occupying 8% of the Earth's land surface, makes up about a third of the world's forested area, and stretches from Alaska and Canada to Scandinavia, Russia, China, Mongolia, and parts of Japan and the Korean peninsula. In North America, boreal forests can also be found at high elevations in the Appalachian Mountains, the Rocky Mountains, and the Sierra Nevada; they are found in similar mountainous regions across Eurasia. These high-elevation forests are typically referred to as subalpine forest, with *taiga* restricted to high-latitude, low-elevation forests. At their northern or upper-elevation margins, these forests open up into woodlands, muskeg (northern bogs), or wind-swept shrublands.

Taiga climates are largely continental, with long, bitterly cold winters (up to 6 months below freezing) and short, warm summers (with only 50 to 100 frost-free days). Precipitation is low (38–50 cm per year), but the climate is humid due to very low evaporation rates. In contrast, subalpine forests can have high levels of winter precipitation, especially on the upwind sides of mountains (see Chapter 16). Large parts of this region were covered with ice during the most recent glaciation. Recurring disturbances are an important feature

Courtesy of the U.S. Fish and Wildlife Service

**Figure 18.23** Taiga in Koyukuk National Wildlife Refuge, Alaska, U.S.A.

of the landscape and include forest fires, treefalls due to storms, and defoliation of large areas by insect attack. Productivity and decomposition rates tend to be very low, and soils are generally very acidic.

Species diversity in boreal forests is low, and these forests are generally dominated by just a few species of coniferous trees. Four tree genera dominate much of the circumpolar taiga: *Picea* (spruce), *Abies* (fir), *Pinus* (pine), and *Larix* (larch or tamarack). (Larches are one of the few deciduous conifer species.) The understory of a boreal forest is typically shady, with few species. Broad-leaved deciduous trees and shrubs are common dominants in the early stages of both primary and secondary succession, with *Alnus* (alder), *Betula* (birch), and *Populus* (aspen) particularly common. In much of Canada and Alaska, for example, the dominant species are *Picea glauca* (white spruce) and *Picea mariana* (black spruce), with lesser numbers of *Abies balsamea* (balsam fir), *Larix laricina* (larch), *Pinus banksiana* (jack pine), *Pinus contorta* (lodgepole pine), *Populus tremuloides* (quaking aspen), *Populus balsamifera* (balsam popular), and *Betula papyrifera* (paper birch). Large mammals including various bear species, wolves, moose, and caribou make this biome their home in the Northern Hemisphere, and many

birds migrate here in summer to breed and raise their chicks. Insects, notably black flies and mosquitoes, can have very large populations in summer.

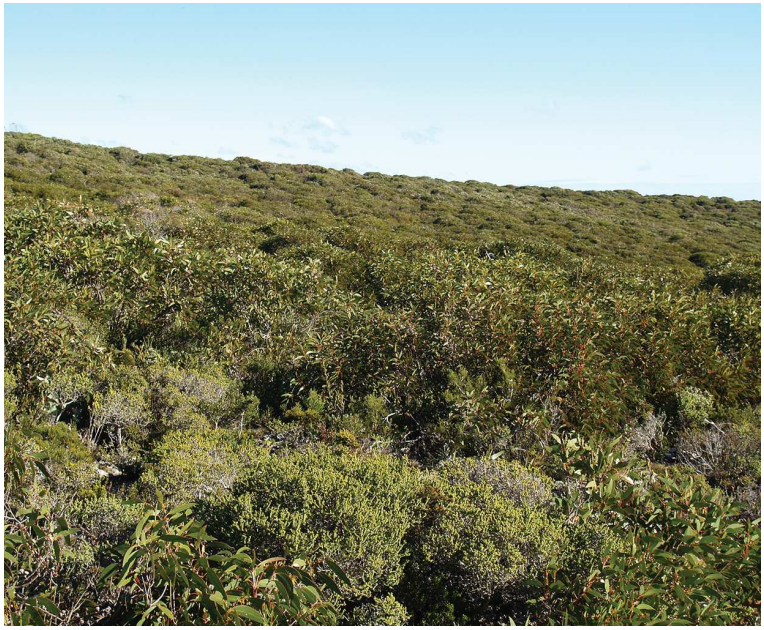
Because low decomposition rates lead to the accumulation of peat (see Chapter 5), the taiga is a "net carbon sink," playing a critical role in the global carbon balance by absorbing about 700 million tons of carbon per year, about equal to the annual emissions of Germany. Large areas of taiga remain as old growth, or at least are intact and relatively undisturbed. The taiga of Russia and Canada constitutes more than half of the world's remaining old-growth forest. Because human populations in the taiga are low, the primary threat to these forests is commercial lumbering for timber, pulp, and paper. In Sweden, for example, large-scale commercial logging and replanting with monocultures of timber species for over a century has transformed many boreal forests into low-diversity timber plantations. Climate change, mining, and oil and gas drilling also threaten many taiga ecosystems.

## 18.8 Temperate Shrubland

Temperate **shrublands** are found in regions with mediterranean climates, which have cool, wet winters and warm to hot, dry summers (Figure 18.24). These regions include the lands surrounding the Mediterranean Sea itself plus lands from 30° to 40° latitude on the western coasts of continents that are adjacent to cold ocean currents. These regions are highly isolated from one another on different continents, resulting in many endemic species. Names for this biome differ around the world: garrigue and maquis around the Mediterranean Sea; fynbos, Karoo, and others in the Cape region of southern Africa; matorral in central Chile; kwongan in southwestern Australia; and chaparral on the southern coast of California. Species diversity ranges from low to exceptionally high, most notably in the fynbos (see Box 19A) and kwongan. All of these different types of temperate shrubland are dominated by sclerophyllous evergreen shrubs (Figure 18.25). Many members of the shrub flora are highly aromatic, particularly those from the Mediterranean proper, which include many members of the mint (Lamiaceae) family (e.g., sage, rosemary, thyme, and oregano). Many contain highly flammable oils.

Total annual precipitation in temperate shrublands is 40–100 cm per year, but precipitation is highly concentrated in a few months of the cool season. Temperatures are cool to hot, moderated by proximity to the ocean and by fog caused by the cold ocean currents. There is a short, predictable growing season when soil moisture and temperatures are sufficient for growth; most plant growth and flowering occurs from late winter to spring. Many of the plants are adapted to drought. Fires are frequent, as the residents of California and Australian suburbs encroaching into this biome have been dismayed to discover. In addition to the dominant shrubs, annuals are very common, especially in the years immediately following a fire.

Temperate shrublands have been affected by humans for millennia due to livestock grazing and the use of fire. The Mediterranean region itself, as we know from classical Greek literature, was formerly forested with evergreen oaks, pines, cedars, wild carob, and wild olive. Many scientists believe that the shrublands of California, likewise, became much more extensive as a result of deliberate burning by Native Americans as a management tool, as well as livestock grazing by Spanish settlers (Woodward 1987). There is wide variation among continents in the current human population size in, and the nature of human effects on, temperate shrublands. Human populations are not large or rapidly growing in most shrubland regions, although in California, suburbanization and development coupled with rapid population increases are major threats to chaparral. The shrubland of Australia has a low human population, but in some places it has been cleared for agriculture. The fynbos region is particularly threatened by invasive plant and animal species.



Courtesy of S. Scheiner

**Figure 18.24** Temperate shrubland on Kangaroo Island, Australia.

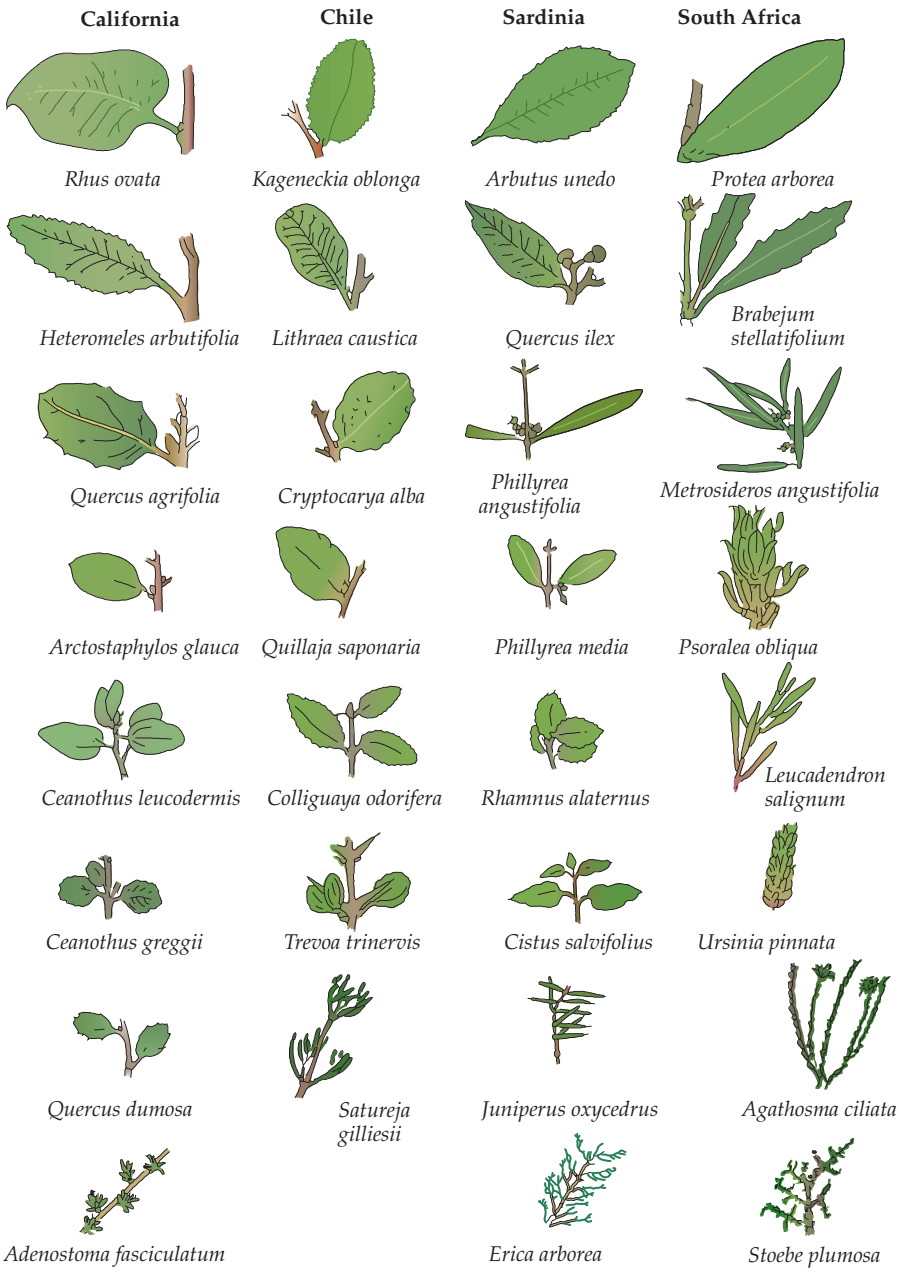



**Figure 18.25** Examples of how convergent evolution, common descent, and biogeographic history have led to similarity of leaf morphology in four widely separated areas with temperate shrubland biomes. Leaves in temperate shrublands tend to be sclerophyllous: tough, evergreen, relatively small, rounded, and thick. The leaves illustrated are from southern California in North America (chaparral), northern Chile in South America (matorral), the island of Sardinia in the Mediterranean Sea, and South Africa (fynbos). In some cases, different species that are members of the same genus are found in different areas (such as *Quercus* in California and Sardinia), an example of the effects of common descent and biogeographic history. In other cases, unrelated species found in the same area have similar leaves (as in California with species in the genus *Quercus* that are members of the family Fagaceae and species in the genus *Ceanothus* that are members of the family Rhamnaceae), an example of convergent evolution. (After M. L. Cody and H. A. Mooney. 1978. *Annu Rev Ecol Syst* 9: 265–321.)

## 18.9 Grasslands

### Temperate grassland

Temperate grasslands (Figure 18.26) are known by many names around the world, such as prairie (North America), steppe (vast parts of north-central Eurasia), veldt (South Africa), pampas (South America), and puszta (Hungary). They were once extensive but are now largely converted to croplands, although grazing lands and rangelands preserve some of their original species and character. The grassland biome is generally found in moderately dry continental climates with cold, dry winters and warm to hot summers with higher rainfall. Fire is an essential part of this biome. Grasslands are the largest of the North American biomes; at one time, they made up 266 million ha (35%) of the United States. The grasses can be short, especially in drier areas, or tall. *Andropogon gerardii* (big bluestem), a dominant species of the tallgrass prairie of central North America, can grow as tall as 2.5 m. The tallest grasses and greatest productivity are in the easternmost part of the North American grassland biome, with stature and productivity declining to the west.

The invention of the steel plow by John Deere in 1837 resulted in the conversion of almost all of the tallgrass prairie in North America, and much of the mixed-grass prairie to its west, to cropland. Today only a few prairie remnants are preserved. In Illinois, for example, less than 1% of the original prairie still exists. The grasslands that have been converted to croplands for growing grain

© Robert and Jean Pollack/Biological Photo Service

**Figure 18.26** Tallgrass prairie in Nebraska, in the center of the United States. It is easy to see why this biome is sometimes poetically referred to as "a sea of grass."

worldwide are often described as "the breadbasket" of their particular country.

Rainfall is strongly seasonal in most temperate grasslands, leading to seasonal drought. In some grassland areas, the total precipitation would be enough to support trees if it were more evenly distributed throughout the year. In many grasslands, fire is the key factor in the exclusion of trees and shrubs. All of the U.S. prairie regions east of the Mississippi River are maintained by fire; evidence suggests that Native Americans used fire to maintain these grasslands in areas that would otherwise be temperate deciduous forest.

Human effects on temperate grasslands are extensive. Fire suppression has led to the invasion of many grasslands by woody plants. Human hunting of large native grazers, such as bison in North America, has also favored woody plants such as shrubs that are vulnerable to grazing. High densities of domestic cattle have also had major effects. One major difference between bison and cattle is that while bison were always on the move, allowing the grasses to recover, cattle are kept in one place for long periods. The introduction and spread of invasive non-native plant species in North America has also resulted in dramatic changes in grasslands at the community, ecosystem, and landscape levels.

Enormous regions of temperate grassland have been converted to irrigated cropland. Such intensive agriculture is often not sustainable over the long term. For example, much of the North American Great Plains is irrigated with water from the Ogallala Aquifer that underlies the southern part of the region. This water will not be available forever; the aquifer is being used up much faster than it is being recharged. Over many years, irrigation leads to salinization of soil because the water leaves behind concentrated mineral deposits when it evaporates. Moreover, the costs of pumping the water increase over time as the water table drops. Irrigated agriculture has already become untenable in parts of the Great Plains, and the problems being experienced there will expand to other regions in the future. Meanwhile, the native grassland is gone and cannot replace the irrigated cropland—and the change has occurred within just two or three generations. The efforts of restoration ecologists to restore grassland ecosystems have met with a variety of results, from what optimistically appear to be long-term successful restorations to complete failures. Successful restoration of prairies and other grasslands to their original condition is costly, labor-intensive, takes a long time, and depends on a range of conditions; it is therefore unlikely to provide workable solutions for anything other than very limited areas. On the other hand, the return of ecosystem structure and function, albeit to a very different community, is much easier to accomplish (see Chapter 13).

Anthropogenic disturbances in grasslands can have interactive effects, as seen in the disappearance of the valley grasslands of California. The Central Valley of California was once a vast perennial grassland, with extensive wetland areas and fertile riverbeds. Spanish explorers wrote that, in wet years, it was impossible to ride a horse across the Sacramento Valley due to the density and height of the vegetation, as well as the widespread wetlands. This vegetation is now entirely gone, having been replaced by short Mediterranean annual grasses (and some exotic forbs). Conversion to irrigated cropland, while important, was not the main reason for this change. Interactions between drought, high densities of non-native domestic cattle and sheep, and the introduction of non-native plants all contributed to the demise of the perennial grasslands. The clumped growth form of the native grasses provides open spaces that allow the establishment of invading plants. The native grasses also do poorly under heavy or close grazing. When the United States took possession of California from Mexico, and grazers were introduced in the 1850s and 1860s, the balance began to tip toward the successful incursion of non-native plant species. The fact that the area was then undergoing a prolonged drought also played a role in giving the invasive annuals an advantage.

## Tropical savanna

Tropical savannas are perennial grasslands that may have an open tree canopy or scattered shrubs. *Savanna* is a Hispanicized variation of a Native American word for "plains." Like temperate grasslands, savannas occur in areas with distinct dry and wet seasons that would be seasonal forests or woodlands if soil conditions or disturbances did not limit the establishment of canopy trees. Rainfall is 75–125 cm per year, with a dry winter season that lasts at least 5 months. Disturbances are caused largely by grazing or frequent fires, or occasionally by other factors such as soil saturation (due to impermeable soils). The most extensive tropical savannas are found in Africa and in Brazil (Cerrado); others are found in Central America, Australia, and Southeast Asia.

Savannas are dominated by tall perennial grasses 1–2 m high, typically interspersed with drought- and fire-resistant trees or shrubs. Palm, eucalyptus, and acacia savannas are examples of this vegetation type, each named for its scattered overstory trees. Savannas have a wide range of characteristics. Most are fairly high in species diversity. Decomposition rates are lower than those in tropical forests but are still generally very high, and soils are quite variable. The acacia savannas of eastern and southern Africa are a mosaic of local communities created and maintained by fire and grazing and are preserved in Kenya, Tanzania, Zimbabwe, Botswana, Namibia, and South Africa. The famous Serengeti Plains of Tanzania are acacia savanna growing on nutrient-rich volcanic sands, although most savanna vegetation occurs on nutrient-poor soils. The llanos of the Orinoco basin of Venezuela and Colombia are grass savannas maintained by the annual flooding of the Orinoco River, which results in long periods of standing water that prevent the establishment of trees. The Cerrado of Brazil has an open canopy of short trees with a distinctive twisted form; its acidic, heavily leached soils are high in toxic aluminum ions and low in nutrients. This region has many endemic species, and its overall plant species diversity approaches that of tropical rainforest. The pine savannas of Central America occur on dry, low-nutrient quartz sands. The savannas of Southeast Asia are generally considered to be created and maintained by humans through frequent fires.

Grazing can combine with fire to establish or maintain savanna vegetation. In East Africa, elephants debark trees and, more dramatically, knock them over, opening woodlands to invasion by grasses. The presence of grasses increases the attractiveness of the area to other grazing animals, such as antelopes and zebras, which eat and trample tree seedlings and prevent woodland regeneration. Only thorny trees and shrubs can become established in this environment, leading to the creation of acacia savannas (Figure 18.27).

The world's greatest diversity of ungulates (hoofed mammals) is found on the savannas of Africa; the species include buffalo, giraffes, eland, impalas, gazelles, oryxes, gerenuks, wildebeests, plains zebras, rhinoceroses, elephants, and warthogs. The vast herds of herbivores support carnivores that include lions, leopards, cheetahs, jackals, and hyenas. In contrast, few Neotropical (Central and South American) mammals or birds are restricted to savannas, although capybaras (very large, semiaquatic rodents) are typical of llanos, and anteaters are common in South American savannas. Termites are especially notable residents of all tropical savannas; these detritivores are important in soil formation (Figure 18.28). The termite-eating aardvarks and pangolins that live in the savannas are among the oddest mammals on Earth.

Like many other biomes, tropical savannas are threatened by human population growth. In Africa, the greatest challenge in preserving the large herbivores is to find ways in which they can coexist with humans. Beyond the problem of poaching, elephants, for example, are at risk when they wander off the nature reserves and people shoot them to prevent them from destroying vital food crops. Conservation plans that include economic incentives and participation in decision-making by the people living in the region, not just by outside conservation experts, are likely to be the most successful strategies for preserving these and other natural systems.

© D. Perlman/EcoLibrary.org

**Figure 18.27** Savanna in Tarangire National Park, Tanzania. Grazing by animals such as zebras maintains the grassland by reducing regeneration of trees.

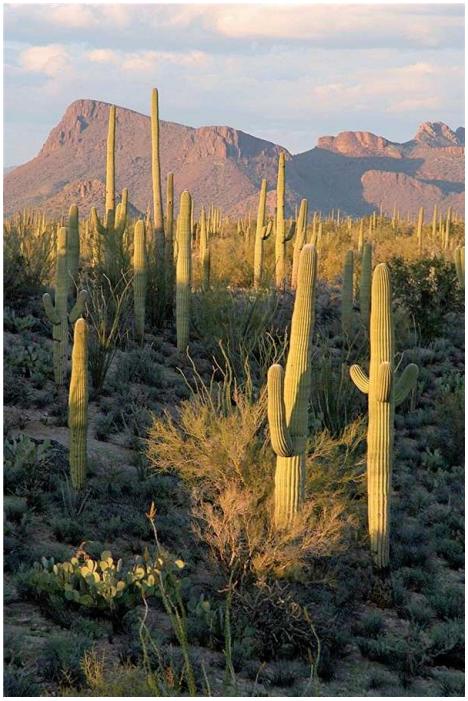
Courtesy of S. Scheiner

**Figure 18.28** Some tropical savannas and dry tropical woodlands support large colonies of termites, whose mounds can visually dominate the landscape. Here, a mound built by *Nasutitermes triodiae* (cathedral termite) in the "top end" region of the Northern Territory, Australia. Mounds of this grass-feeding termite can grow to 5–6 m in height. Mound-building termites are often the dominant herbivores in nutrient-poor tropical regions with highly seasonal rainfall.

## 18.10 Deserts

### Hot desert

Deserts and semideserts occupy large areas of western North America, Australia, Asia, and Africa. Deserts occur where potential evapotranspiration exceeds actual evapotranspiration—in other words, where water can be lost to the atmosphere at a greater rate than it is normally made available to plants. Typically this means that desert regions receive less than 25 cm of precipitation per year, but such numbers can be deceptive; for example, the coastal regions of southern California receive less rainfall than this, but they are not deserts because cold offshore currents act to lower temperatures and increase relative humidity. Thus, San Diego, California, which has a mediterranean climate, typically receives less annual rainfall than Tucson, Arizona, in the Sonoran Desert.



Courtesy of G. Fox

**Figure 18.29** The Sonoran Desert outside Tucson, Arizona, U.S.A.

Hot deserts (also called "warm deserts") are found in two belts centered at 30°N and 30°S latitudes. The hot desert biome includes the Mojave, Sonoran, and Chihuahuan Deserts of North America; the Sahara and Namib and the Kalahari Desert of Africa; the Arabian Desert of Asia; the Atacama Desert of South America; and the deserts of Australia. These areas all have pronounced seasonal wet-dry cycles, although some of the driest deserts may experience periods of many years without any measurable rainfall.

Precipitation in deserts typically comes during one time period of the year. In North America, the Mojave Desert receives almost all of its rain in winter, while the Chihuahuan Desert receives mainly summer rains. The Sonoran Desert is unique among the deserts of the world in having two periods of rainfall, a winter and a summer rainy season, with very dry autumn and spring seasons separating them. As a result, the Sonoran is the greenest of the world's deserts, with the highest standing biomass and productivity, and is one of the most diverse in the number and forms of species (Figure 18.29).

As a rule, the lower the average annual rainfall, the greater the variation in both the amount and timing of annual rainfall. The "rainy season" in most deserts is a time when rain is most likely, although it is not uncommon to have very little rain in any given rainy season. Consequently, desert organisms must be adapted not only to limited moisture and high temperatures, but also to the lack of predictability in rainfall (see Chapter 3 and

Chapter 6). At one extreme of adaptation, annual plants spend most of their lives as desiccation-resistant seeds, living only briefly as photosynthetic organisms when conditions are most favorable. At the other extreme, succulents and evergreen shrubs have photosynthetic tissue that is always exposed to potential desiccation. Drought-deciduous shrubs have an intermediate strategy, exposing little moisture-losing tissue during the long dry periods. In part because of the unpredictability of rainfall, succulents are less abundant or absent in the driest deserts; they are characteristic of deserts with relatively greater and more predictable rainfall. In Death Valley, California, which is extremely hot and dry, annual plants are the dominant growth form. The other common growth form in Death Valley is drought-deciduous shrubs. In contrast, succulent cacti, agaves, and yuccas are abundant and diverse in the Sonoran Desert. The world's driest desert is the Atacama Desert on the west coast of South America, where some places may not receive rain for 10 years (see Chapter 16).

Because annual species predominate in hot deserts, their seeds are usually relatively abundant in the soil, and consequently there are many **granivorous** (seed-eating) animals, particularly rodents, ants, and birds. Southern Arizona, for instance, has the highest diversity of rodents in the world, and almost the highest diversity of seed-eating ants. Although the seeds of desert annuals are abundant, they do not usually disperse well. Many desert plant species have adaptations that limit their ability to disperse. Areas from which plants have been cleared often are not recolonized for decades or even centuries.

Deserts are important areas for agriculture, grazing, and mineral exploitation, and they are becoming increasingly important for recreation as well. For almost all of human history and prehistory, human populations in deserts have been very low. In the last few decades, however, there have been rapid increases in human populations in some desert areas. The explosive growth of desert cities, such as Phoenix, Arizona, has meant that human activities are having massive effects on natural systems in such areas. Underground aquifers are being depleted by urban dwellers and by agricultural use, and surface waters (such as the Colorado River in the southwestern United States) face very heavy demands. Irrigated agriculture can lead to soil salinization in deserts, just as it does in grasslands. This process is thought to have caused the collapse of several ancient civilizations in the American Southwest. Environmentally damaging mining and oil extraction techniques are additional sources of severe disturbance to biotic systems in many deserts. Overgrazing, by both livestock and feral ungulates, has widespread effects in many deserts. Heavy grazing disturbs the soil and increases evapotranspiration rates by reducing cover. Use of recreational off-road vehicles, common across the American desert landscape (as well as in other regions), is highly destructive to the fragile vegetation. Because the vegetation is responsible for holding the soil in place, the loss of vegetation through grazing and other human activities leads to soil loss and severe erosion.

## Cold desert

Cold deserts (also called "cool deserts") are found in rain shadow areas in the centers of continents, the largest being the Kazakho-Dzungarian and Gobi deserts of central Asia and the Intermountain or Great Basin desert of North America (Figure 18.30). Unlike hot deserts, cold deserts experience freezing temperatures for part of the year. All deserts can be quite hot in summer, however. In cold deserts, precipitation is generally concentrated in winter. Cold deserts are dominated by shrubs and grasses and have few of the succulents found in some hot deserts, as most succulents are vulnerable to damage by ice crystal formation. Cold deserts also have few annual plants because moisture is available mainly as snowmelt, precisely when temperatures are too cool for rapid growth. There may also be historical reasons for the dearth of annuals. In the Great Basin of North America, for example, there are few native annual species, but the successful invasion of the Old World annual grass *Bromus tectorum* is evidence that annuals can do well there. Most of the anthropogenic threats to hot deserts also threaten cold deserts.

Courtesy of J. Gemme

**Figure 18.30** Sagebrush (*Artemesia*, Asteraceae) dominates this cold desert landscape in Wyoming, U.S.A.

## 18.11 Alpine and Arctic Vegetation

### Alpine grassland and shrubland

Alpine shrublands extend upward from the tree line on high mountains and are often interspersed with alpine grasslands (Figure 18.31). The vegetation in these areas can include both what we typically think of as shrubs—multistemmed woody plants of short stature—and bizarre-looking plants that have a rosette form and can grow 3–5 m tall (Figure 18.8). Alpine shrublands are found on all of the continents except Australia and Antarctica. In South America, they constitute a large part of the high-altitude vegetation type called paramo (Figure 18.32). **Heath** is the name for a general vegetation type that is dominated by low-growing shrubs, and in some parts of the world alpine shrublands consist to a large extent of heath vegetation, although not all heaths are alpine nor are all alpine areas heaths. In the northern temperate zone, the shrubs are typically *Salix* (willow, Salicaceae) and *Betula* (birch, Betulaceae).

Above the tree line in mountainous areas or in areas with cold air drainage, alpine grasslands are the major community type (Figure 18.33). These grasslands are often dominated by **sedges**, plants in the family Cyperaceae that are closely related to and similar in growth form to grasses. These communities can also contain scattered shrubs, dwarf trees, and small, herbaceous perennial plants that are shaped like little cushions, thus grading into alpine shrublands and tundra. Grasslands can also be found in alpine or Artic regions where trees are suppressed by grazing (Figure 18.34).

Courtesy of S. Scheiner

**Figure 18.31** Alpine shrubland on the island of Tenerife, Canary Islands, Spain.

Temperate and tropical alpine areas differ in some important ways. Temperate alpine regions have long, cold winters, during which the ground is frozen and deep snow covers much of the vegetation for long periods. Snowmelt generally marks the start of spring growth. In contrast, tropical alpine regions, especially in very high mountains, have a strong daily temperature cycle, as if it were winter every night and summer every afternoon. Growth continues year-round, and the soil does not freeze (at least not deeply or for long periods of time). All alpine regions have very high solar energy input because the atmosphere is thinner than at sea level, so there is less atmosphere to absorb and scatter incoming sunlight. The sunlight is particularly intense in tropical alpine regions because the angle of incoming solar radiation is most direct over the tropics. The resulting high levels of ultraviolet radiation can damage biochemical processes and DNA in leaves and meristems (and can cause severe sunburn and skin cancer in humans).

The total area occupied by alpine vegetation of all types is low. Precipitation is generally limited in these areas, and productivity and species diversity are also low. The harsh climate, steep slopes, and low productivity of alpine regions have combined to keep human population sizes low from prehistory to the present, although several distinct cultures, such as those of the Incan and Tibetan peoples, have developed and thrived in high mountain regions. Even small human populations, however, can have significant effects on alpine habitats. The cold temperatures mean that many processes, such as plant growth, soil formation, and decomposition, are very slow. As a result, human activities such as mining and grazing (the primary form of agriculture in these regions) can have very long-lasting effects.

Pedro Szekely/CC BY 2.0

**Figure 18.32** *Azorella compacta* (yareta, Apiaceae) growing in Bolivia. These plants grow extremely slowly, approximately 1 mm/year, so this plant may be as much as 3000 years old.

Courtesy of S. Scheiner

**Figure 18.33** Alpine grassland in Rocky Mountain National Park, Colorado, U.S.A.

## Tundra

**Tundra** is found in areas where temperatures are too cold for tree growth. There are vast areas of Arctic tundra in northern North America and Eurasia. Because there is little land in the Southern Hemisphere at high latitudes, there is little tundra there—only a few areas on the Antarctic Peninsula and neighboring islands are referred to as tundra. Conditions at the tops of high mountains, including wind and extreme cold, create alpine tundra.

The Arctic tundra has long, dark, and often bitterly cold winters and short, cool summers with up to 24 h of sunlight. The soil thaws each summer to a depth of only 25–30 cm, resulting in **permafrost**, a permanently frozen layer of soil below that depth, which severely limits root growth. Trees are excluded from the Arctic tundra by a combination of factors, including permafrost and extreme cold. Temperatures below –40°C kill almost all living aboveground plant tissue, preventing long-lived plants such as trees from becoming established. Another factor limiting tree growth is that emergent vegetation above the snow cover in winter would experience not only extreme cold, but also high winds, especially in alpine areas. Tundra soils are generally young and poorly differentiated because most Arctic tundra sites were glaciated in the recent past. Litter decomposition is generally slow due to the cold temperatures and consequent low microbial activity and, secondarily, to the anaerobic conditions in many places.

Alpine and Arctic tundra have similarities and differences. Both have short growing seasons, and alpine tundra (at least in temperate areas, although not in the tropics) is subject to severe cold in winter. In both types of tundra, most precipitation is in the form of snow, so there are long periods during which plants are unable to take up moisture; in many settings (such as the alpine tundra at mid-latitudes in western North America), drought is common during the summer. Arctic and alpine tundra differ greatly in other factors, including day length, atmospheric pressures, partial pressures of gases such as $CO_2$, exposure to ultraviolet radiation, and soil conditions.

Many genera and species found in Northern Hemisphere alpine tundra are also found in Arctic tundra. *Oxyria digyna* (mountain sorrel, Polygonaceae) is a plant found both in the Arctic and in alpine areas of the western and eastern United States and Europe. A comparative study of Arctic and alpine populations of this plant by Hal Mooney and Dwight Billings (1961; Billings and Mooney 1968) concluded that the populations are each adapted to specific local conditions. Compared with the Arctic plants, the alpine plants had a lower leaf chlorophyll content, a higher respiratory rate, and a lower light saturation point. Alpine tundra species are often evolved from local lowland species (see Figure 18.8). In the high Andes, for example, there are many species of high-altitude bromeliads, a group we normally associate with tropical forests.

The growth forms of tundra plants include low shrubs, taller shrubs alongside rivers, herbaceous perennials, grasses and sedges, mosses, and lichens. Different plant communities exist within the tundra,

Courtesy of V. Vandvik

**Figure 18.34** Arctic grassland at Ervik, Stadt, Norway. This area is maintained as grassland by grazing. The species flowering in the foreground is *Primula veris* (cowslip, Primulaceae), a characteristic species of these species-rich coastal grasslands.

depending on topography, soils, and other factors. Tussock tundra is dominated by grasses and sedges (notably by the sedge *Eriophorum vaginatum*), deciduous shrubs, and evergreen shrubs in roughly equal abundance. Riparian bottomlands are dominated by relatively tall (1–4 m) willows and other shrubs. Wet sedge tundra, dominated by short (<20–30 cm) rhizomatous sedges, covers flat areas with restricted drainage, forming a broad expanse on the Arctic coastal plain of Alaska as well as in other areas. Soils typically thaw to only 25–30 cm here, and the freeze-thaw cycle often creates a characteristic wedge-shaped pattern in the soil. The soils of wet sedge tundra have a very high organic matter content and may be covered with water during the summer because the permafrost prevents drainage.

Plant species diversity in tundra is very low but can include a variety of growth forms. Many genera (such as *Salix*) and even some species (such as *Dryas octopetala*, *Eriophorum vaginatum*, and *Oxyria digyna*) are common across the circumpolar Arctic (Figure 18.35), and similar communities are found over wide ranges. Animal population densities and species richness are also low. However, the Arctic tundra provides critical habitat both for species that are characteristic of this region, such as lemmings, and for others, such as wolves, that have been reduced to low numbers elsewhere. Brown bears, grizzly bears, polar bears, snowshoe hares, lynx, moose, and caribou (called reindeer in northern Europe) are some of the other large mammals that are notable components of the Arctic tundra and taiga fauna, and marine mammals such as walrus and various seal and whale species are native to, or rely heavily on, Arctic habitats. The Arctic tundra serves as critical nesting area for many bird species, particularly those associated with wetlands.

Human population sizes and densities in the tundra are extremely low (under 1 person/km$^2$), although humans have lived in, traveled through, and exploited Arctic environments since prehistoric times. In spite of these low population sizes, humans have managed to pollute Arctic environments in disproportion to our numbers, and the cold temperatures mean that the pollution persists for a long time. Small soil disturbances can last a very long time. In the Arctic tundra, disturbance may result in the melting of permafrost, which leads to subsidence and soil compaction. Perhaps the most threatening of all anthropogenic effects is global warming, which is already having profound effects in polar regions (see Chapter 16 and Chapter 19).

Courtesy of Lawrence R. Walker

**Figure 18.35** The vegetation of this Arctic tundra along Glacier Bay in southern Alaska, U.S.A., is dominated by species of *Dryas* (Rosaceae).

## Summary

- Biomes are large regions that differ from one another in the physiognomy of their vegetation.
- Biomes are strongly determined by climate—temperature, rainfall, and seasonality—and, to a lesser extent, by soil properties.
- While the boundaries of biomes are defined arbitrarily, the concept of the biome remains a useful one so long as it is clear that it is a descriptive categorization. There is a broad consensus among scientists on the classification of biomes.
- The biome concept provides a framework for analysis and separation of the effects of species sorting, convergent evolution, commonality of descent, and biogeographic history.
- Biomes are classified based on whether they are dominated by trees (forests and woodlands), shrubs, or grasses and other herbaceous species.
- It is the combination of the mean climate of a region, the variability of that climate, and the predictability of that variation that determines the types of plants that will grow there.

(*Continued*)

## Summary (*continued*)

- Trees can grow only where temperatures exceed 10°C for sufficient periods of time.
- Forests are further classified by whether the trees are evergreen or deciduous, broad-leaved or needle-leaved. Warm temperatures and long growing seasons favor broad-leaved deciduous trees over needle-leaved evergreen trees.
- The borders between shrublands, woodlands, and grasslands are often controlled by rates of disturbance, especially the frequency of fire, which in turn is controlled by patterns of temperature and precipitation.
- Where conditions become too dry for trees, grasses predominate; grasslands can be tropical, temperate, alpine, or Arctic.
- Shrubs dominate under a variety of conditions—including those that are too hot or dry for grasses, those with very seasonal precipitation, and some with very cold temperatures—or as a secondary result of overgrazing or fire suppression.
- Some biomes, such as temperate deciduous forests, can recover relatively rapidly from human disturbance because many of their species disperse readily and grow rapidly. Others, such as hot deserts, recover poorly because their plants disperse poorly and grow slowly. Full recovery in still others, such as tropical forests, is difficult because of soil destruction and species extinctions in these highly diverse communities.
- Recovery of all damaged or degraded biomes is limited by continued encroachment, habitat destruction, and global climate change.

Courtesy of J. Gurevitch

19

# Global Biodiversity Patterns, Loss, and Conservation

You may have noticed that some places have many different kinds of plants, while others have far fewer species. As you travel to different parts of the world, you become aware that not only does species diversity change from one place to another, but some places have similar taxa—such as maples (genus *Acer*) in eastern North America, northern Europe, and the mountains of Japan—while others have plants found nowhere else. In this chapter, we consider global and regional patterns of diversity and the explanations that have been proposed for those patterns. We also consider species loss. Human activities have resulted in a high rate of extinctions of species of all kinds. While we are more likely to hear about the extinction of birds or mammals, we are also losing plants many times faster than would occur without human activities. Almost 600 plant species have become extinct over the last 250 years, and the extinction rate has been accelerating (Vellend et al. 2017; Humphreys et al. 2019). Where are species being lost most rapidly? What can we do about it, and what is being done? We will take a look at these critical issues in this chapter.

Above: A fynbos landscape dominated by plants with grass-like forms in the Restionaceae, with scattered Erica shrub species and plants in the Asteraceae. Western Cape, South Africa.

## 19.1 Biodiversity Varies Enormously across the Earth

One aspect of biodiversity is variation among organisms, which encompasses a range of kinds of variability and spatial scales. In practice, most studies of biodiversity focus on species richness, and differences in species composition from one place to another. At local and regional scales, we may often be concerned with variation in abundances as well as numbers of species, but at the largest spatial extents almost all studies focus on species richness, in part because it is the only information available. When we measure biodiversity as species richness, we assume implicitly that the number of species is highly correlated with other aspects of biodiversity. Biodiversity also includes genetic variation, variation in plant functional trait characteristics, differences in physiognomy, community or ecosystem characteristics, and variation at higher taxonomic levels, such as the number of genera or families. Measures of some of these types of diversity are detailed in Chapter 12. In this chapter we focus on patterns, drivers, and consequences of species richness at large spatial scales and on threats to biodiversity, but we will touch on approaches to quantifying other aspects of biodiversity.

People have been aware of variation among plants since before we were human, because our survival depended on recognizing and predicting which plants were useful and which were toxic. Many societies and cultures developed formal means for understanding plant diversity, but the modern scientific understanding of plant species began in the 1700s with the taxonomic methods of Carl Linnaeus. Linnaeus was primarily interested in classification, not ecology or evolution. Awareness that biodiversity differed across the globe slowly grew in the 1700s, but it was the work of the scientific explorer Alexander von Humboldt (see Figure 18.10) that captured the popular (and scientific) imagination. In 1799, when he was in his 30s, von Humboldt set out on a 4-year trip to the Caribbean and South America with Aime Bonpland, a plant taxonomist, through rainforests, wetlands, grasslands, and the high Andes. They collected over 60,000 plant species and made scientific measurements of geography, geology, climate, and meteorology. Von Humbolt's wide-ranging ideas established the foundation for the scientific study of the geographic patterns of plant biodiversity, setting the stage for modern biogeography and ecology.

How many different kinds of plants are there? While some species may not yet have been described by scientists, there has been a huge effort in recent years to accurately classify the species that are known. Current estimates are that there are between 250,000 and 350,000 species of angiosperms (flowering plants), with between 200 and 400 families, and up to 14,500 genera. In contrast, there are fewer than 1000 species of gymnosperms in 14 families and 81 genera (although conifers, the largest group of gymnosperms, dominate huge areas ecologically; see Chapter 18). There are about 13,000 species of ferns and fern allies in 48 families, and 20,000 species of mosses and other bryophytes. There are all together 351,000 plant species with taxonomically accepted names, in 642 plant families and 17,000 genera (good sources are The Plant List and the World Checklist of Selected Plant Families). Angiosperms are far and away the most diverse group. Among angiosperms, the most numerous families are the Asteraceae and Orchidaceae, each with well over 25,000 species; the Rubiaceae and Poaceae follow, each with over 10,000 species. One of the smallest familiar plant families is a gymnosperm—the Ginkgoaceae, with one genus, *Ginkgo*, and one species, *Ginkgo biloba*.

Some parts of the world have a great many more species than others. Some tropical forests have as many as 1500 species of flowering plants (including 750 species of trees) in 1000 ha (10 $km^2$). In Indonesian Borneo, researchers recorded more than 400 tree species in a 0.75 $km^2$ plot! In contrast, the northern regions of Canada and Russia may have only a few dozen plant species spread across hundreds of square kilometers. Why do some regions of the world have so many more plant species than others?

In this chapter, we describe regional and global patterns of species richness, and we discuss progress in understanding their causes by drawing together concepts introduced in other chapters. In Chapter 12 we introduced the ideas of species diversity and species-area curves and considered patterns of diversity within and among communities. Landscape patterns and processes were examined in Chapter 15. Here, we look at patterns at the scale of regions, continents, and the entire globe. At each scale, we examine some of the hypotheses that have been proposed to explain observed patterns. Some of these hypotheses have now been ruled out, others explain observed patterns at least in part, and still others are being debated vigorously by ecologists. No one explanation can account for all the observed patterns of species diversity. Rather, ecologists try to evaluate the relative importance of the various factors contributing to biodiversity in different places and over time. In Chapter 18 we saw how one of those factors—climate—determines patterns of community physiognomy and the distribution of biomes. As von Humboldt discovered, climate and associated factors also have a major influence in determining patterns of plant diversity.

Understanding patterns of plant biodiversity is essential for conserving that biodiversity. An ever-increasing proportion of the planet is occupied by

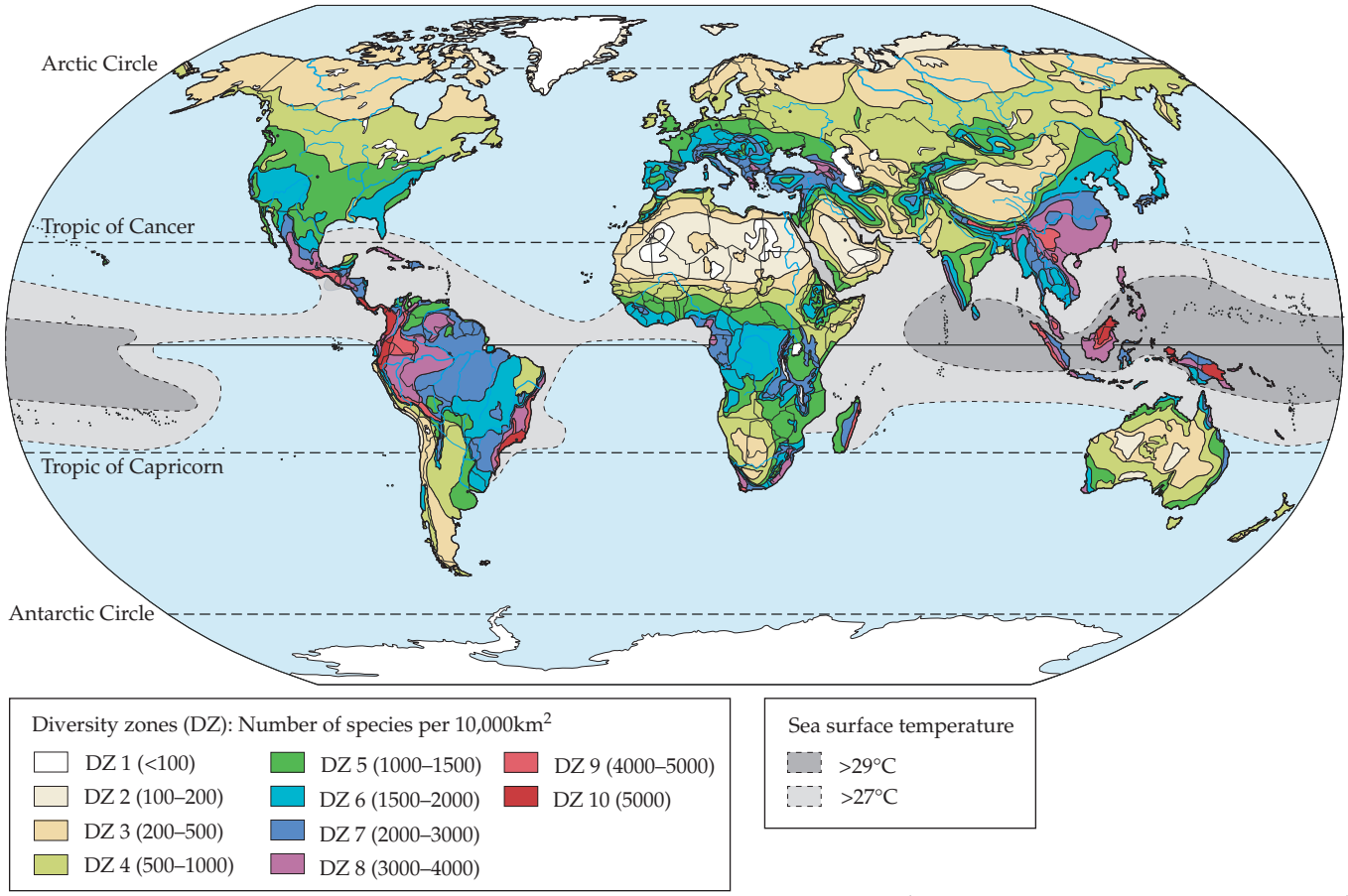



**Figure 19.1** Global patterns of vascular plant species richness (number of species per 10,000 km²). The greatest species richness is found in tropical rainforests and tropical seasonal forests; the lowest species richness is found in deserts and tundra. (After W. Barthlott et al. 1996. *Erdkunde* 50: 317–327; J. Mutke and W. Barthlott. 2005. *Biologiske Skrifter* 55: 521–531.)

human-dominated environments such as farms, homes, and industrial areas. Our domination of landscapes globally leaves wildlands more diminished every year and alters those that remain. Old-growth forests and other intact plant communities are shrinking rapidly in extent and becoming more fragmented; climate change is influencing habitats globally (see Chapter 16). These changes are influencing species distribution, abundance, and richness. If we wish to preserve biodiversity, we need to understand existing patterns of diversity and the processes that sustain those patterns.

## Global biodiversity increases toward the tropics

Scientists began observing patterns of global biodiversity in the nineteenth century and have been more accurately and quantitatively documenting these patterns ever since. Tropical regions are typically very rich in species, polar regions are very species poor, and temperate regions have intermediate numbers of species (Figure 19.1; Barthlott et al. 2005). For example, Brazil has over 56,000 named plant species, the United States has about 18,000, and Canada has about 4200. This pattern is very old geologically, going back at least to the Cretaceous period (Crane and Lidgard 1989; see Chapter 17) and is similar to gradients in other taxa (Willig et al. 2003). Species richness in tropical regions is probably underestimated because the scientific study of the ecology of parts of these areas is especially incomplete. Even in North America, which has been studied by botanists for hundreds of years, new plant species are being named at a rate of about 60 per year. These plants occasionally are found right under one's feet—for instance, in 1997 a new species, the mustard *Lesquerella vicina* (Brassicaceae), was discovered by a botanist, James Reveal, growing in his neighbor's yard in Montrose, Colorado. A "remarkable orchid," *Disa remota* (Figure 19.2), was discovered in 2006 in the very well studied Cape flora of South Africa by Hans Peter Linder and A. Hitchcock (2006).

But how do we actually know that plant biodiversity increases toward the tropics and decreases toward the poles on a global scale? Many researchers have documented latitudinal diversity gradients for a wide range

Courtesy of A. N. Hitchcock and H. Linder

**Figure 19.2** *Disa remota* (Orchidaceae), South African Cape region.

of different organisms. Helmut Hillebrand (2004) analyzed all of the published latitudinal diversity gradients that had been documented up until 2002, confirming von Humboldt's observations quantitatively; plant numbers decrease from the equator toward the poles. (He found similar results for other organisms, although differing in details from those of plants.) Nicole Kinlock and collaborators (2018) updated Hillebrand's analysis with data from 2003 to 2015, confirmed his conclusions, and showed that the pattern was particularly strong for forests. They also found that the latitudinal diversity gradient varied by continent and is stronger in North and South America than in Africa and Eurasia.

What accounts for the latitudinal diversity gradient? Ecologists and biogeographers have been seeking to explain it as long as it has been known. The fundamental problem is that there is only one Earth, but many factors are associated with the gradient. Disentangling them is not simple.

## 19.2 What Explains Global Biodiversity Patterns?

While it would be satisfying to be able to identify a single, unifying explanation for the latitudinal diversity gradient for plants (or even for some groups of plants), it is likely that multiple factors have acted together to create these patterns. Even considering those multiple factors, we do not know if there is a general explanation across many groups of plants and across the Earth, or if explaining specific patterns for particular groups of species in each region will require delving into the details of the ecology, evolution, and geologic histories of each group in each region. Many models have been put forward to explain both geographic and ecological species richness patterns; we examine some of these hypotheses and models here.

### Explanations for the latitudinal diversity gradient include energy, water, and environmental heterogeneity, but all explanations have limitations

One ecological factor that has been consistently shown to affect latitudinal gradients of plant species richness is available energy (see the discussion below, and see Scheiner and Rey-Benayas 1994, O'Brien 1998, Hawkins et al. 2003, and Field et al. 2005). Holger Kreft and Walter Jetz (2007) carried out a global assessment of species richness of vascular plants (including ferns, gymnosperms, and angiosperms) based on 1032 areas with biodiversity data. They found that climate and environmental heterogeneity stood out as important factors for explaining species numbers. Potential evapotranspiration (see Chapter 5), the number of wet days per year (which includes amount and seasonality of rainfall), and measurements of topographical and habitat heterogeneity were the most important factors associated with the latitudinal diversity gradient. Spatial heterogeneity can reflect both evolutionary and ecological factors influencing plant richness. There was one striking exception that could not be explained by their analyses—the astoundingly high species richness of the Cape region of South Africa (Box 19A).

Eleven different evolutionary hypotheses have been proposed to explain the higher diversity of the tropics (Mittelbach et al. 2007); those hypotheses can be summarized as involving higher speciation rates, lower extinction rates, or more time for diversification to occur. A simple model, the area hypothesis, proposes that the tropics are more species rich because the total tropical land area is greater than the total temperate land area, allowing more species to evolve and to coexist (Terborgh 1973; Rosenzweig 1995). However, the study by Kreft and Jetz discussed above found that area explained very little of the latitudinal diversity gradient of plant species, although it could be an important contributor to regional species richness. Another simple explanation for latitudinal species richness patterns involves differences in opportunities for growth. In the most extreme environments, conditions are so unfavorable that few species are physiologically capable of persisting there. Antarctica has only two native angiosperms, both grasses. The climate, with very cold air and little precipitation, means that the soil is poorly developed, the growing season is very short, and photosynthesis is limited. An analogous argument attempts to explain why species diversity is so low in cold, arid Patagonia, Argentina, in southernmost South America, and so much greater

## BOX 19A The Fynbos and the Cape Region of Africa Have Some of the World's Highest Plant Diversity

The fynbos is a small region in the southernmost part of Africa that has some of the highest levels of plant diversity on Earth. Its vegetation is so distinctive that a few scientists have elevated it to the status of a unique biome. Its unique character was recognized in the early twentieth century by the German botanist Adolf Engler, who classified it (including the surrounding plant communities in the Cape region of Africa; see Figure 19.15) as the Cape Floristic Region, the smallest of the world's six great floristic divisions.* The fynbos is home to about 8000 species, some 5000 of which are endemic. It has an extremely high concentration of plant species, with 1300 species per square kilometer and up to 121 species per square meter. The latter statistic points to a truly remarkable feature of the fynbos: its high diversity at small to medium spatial scales.

Derived from the Dutch phrase *fijn bosch* ("fine bush"), fynbos (pronounced "fane boss") is named for the narrow-leaved, grasslike plants and sclerophyllous evergreen shrubs with fine, small, or rolled leaves that are dominant there (see Figure 18.25). It has a mediterranean climate, with cool, wet winters and hot, dry summers. Like other mediterranean vegetation, the fynbos is adapted to frequent fires (which occur predominantly in late summer). The soils are infertile, acidic, sandy, and very ancient. The topography is diverse, with sandstone mountains, steep and dissected valleys, limestone outcrops, shale flats, and coastal plains.

Fynbos is defined by the presence of plants in the Restionaceae (Cape reeds). Restios are somewhat grasslike in form and their ecological role replaces grasses in the fynbos. The Ericaceae (the heaths—generally low-growing, shrubby plants typically found on acidic soils) is another important fynbos plant family; the genus *Erica* reaches a peak of diversity here, with about 600 species, whereas there are fewer than 30 species in the entire rest of the world. Other characteristic plants of the fynbos are members of the Proteaceae and the Asteraceae; over 1000 fynbos species belong to the Asteraceae, more than half of them endemic. The Proteaceae have their evolutionary origin in the supercontinent Gondwana (see Chapter 17). Many of them have truly spectacular flowers. There are many plants in the Iridaceae (the irises) with particularly beautiful flowers, including many that are grown horticulturally. Unlike vegetation in most other places, the fynbos typically has no one dominant species or group of species.

Various distinct local plant communities occur as part of the fynbos and the surrounding Cape region, including the Succulent Karoo and Nama Karoo, the Mountain Fynbos, and the Renosterveld. Animals are far less diverse in the Cape region than plants, although there are rare and endemic species in a number of animal taxa, including fish, reptiles, mammals, birds, and many insect groups. Most dramatic of all may be the highly specialized animal-plant interactions on which many fynbos plant species depend for reproduction and seed dispersal. *Protea acaulis* and many other proteas are totally dependent on mice for pollination. Their flowers are held near the ground and produce a strong odor that mimics yeast. Mice drink the pungent nectar, carrying pollen on their noses to the next plant. Ants are important seed dispersers for many fynbos plants, such as those in the genus *Callistemon* (bottlebrush, Myrtaceae). The seeds of these plants are attached to special

*The six floristic divisions categorized by Engler are the Cape, Antarctic, Australian, Neotropical (Western Hemisphere tropical regions), Paleotropical (African and Asian tropical regions), and Holarctic (the entire Northern Hemisphere north of the tropics).

Courtesy of G. Fox

An upland fynbos landscape at Jonkershoek Nature Reserve, Western Cape, South Africa shown here in early spring. Many of the species in the highly species-rich flora are endemic to this tiny region. There are a multitude of species in the Proteaceae, Restionaceae, Ericaceae, Iridaceae and Asteraceae, including shrubs, spring and summer flowering bulbs, and many herbaceous plants.

(*Continued*)

**BOX 19A** (*continued*)

structures called elaiosomes, which offer a food reward to the ants for carrying the seeds off.

The great diversity of the fynbos is a geologically recent phenomenon. Until some 5 million years ago, the Cape region was covered with subtropical forests. As the climate became cooler and drier, mediterranean shrublands spread into the Cape from the north. Much of the diversification of the flora of the fynbos occurred during the large climatic fluctuations of the past 2 million years, as glacial periods alternated with warmer periods.

The fynbos faces many serious threats. Some 1700 plant species are threatened with extinction in the Cape region. Many fynbos species are highly localized, and there are also many very local centers of species diversification and endemism. The city of Cape Town is located on two such centers of endemism, and its expansion threatens several hundred species. Agriculture and the spread of fertilized commercial crops threaten lowland fynbos. Anthropogenic changes in fire regimes threaten species throughout the fynbos.

The invasion of exotic species and climate change pose major threats to the Cape flora. The silky hakea (*Hakea sericea*) is a fire-adapted member of the Proteaceae from southeastern Australia that was brought to South Africa as a hedge plant and for firewood around 1830. It soon escaped cultivation and has become a major threat to fynbos vegetation. Clumps of this shrub or small tree develop into dense, impenetrable thickets that eliminate native species, alter fire regimes, and deplete soil water reserves. Other devastating invaders include *Pinus radiata* (Monterey or radiata pine, Pinaceae; see Chapter 14) from coastal California, other pine species from Europe, and *Acacia cyclops* (Fabaceae) and *Acacia saligna* from Australia. An invasive, highly aggressive ant species from South America, *Iridomyrmex humilis* (the Argentine ant), is threatening the native ant species on which many fynbos plants depend for seed dispersal, compromising the regeneration of many plant species.

Global climate change is especially threatening to the survival of highly restricted ecosystems, such as those of the Cape region, and to species with very limited distributions. The fynbos is located at the very southern edge of the African continent; if the climate warms and species are driven poleward, there is nowhere for them to go but the salty deep.

Both photos courtesy of G. Fox

*Protea lepicarpodendron* (black-beard sugarbush, Proteaceae) and one of the important pollinators of plants in this genus, the Cape sugarbird (*Promerops cafer*). Protea flowers are also pollinated by beetles, bees, and other insects. Inset photo shows the bud of *Protea lepicarpodendron*.

in the Amazon basin near the equator, and why species richness is lower in the Canadian tundra than in the rainforests of Central America. Unfortunately, this simple explanation does little to explain the rest of the latitudinal diversity gradient, such as the diversity of deciduous forests at mid-latitudes.

Another explanation proposed for the increase in species richness toward the equator is that communities in the tropics are older than those closer to the poles (Fischer 1960). Many communities closer to the poles have experienced repeated glaciations and recolonization of their plant species (see Chapter 18). If communities in the tropics have existed for a longer time, then the organisms in those communities should have had more opportunity to diversify into many species. Given enough time, temperate and polar communities might evolve greater numbers of species, or experience more migration of species from low to high latitudes, and the latitudinal diversity gradient might slowly disappear. A related idea is that tropical communities have experienced

fewer extinctions because they have experienced fewer climatic fluctuations and, in particular, have never been glaciated (Sanders 1968). These explanations or models, which invoke time and climatic stability, seem persuasive at first but have a number of serious weaknesses. For example, there is much evidence that there has been sufficient time for many tropical plant species to migrate into temperate and even polar regions, given the right geography (see Chapter 17). Perhaps more compellingly, there is strong evidence that tropical communities were highly affected by climatic fluctuations during recent glaciations, even though they were not covered by glaciers; large areas of tropical forest were grasslands at some points in the past, for instance. There is another serious problem with this as a general explanation for the factors that create and maintain gradients in diversity. Gradients in species diversity also occur as one moves from low to high elevation on mountains (see below). The diversity gradients on tropical mountains are not readily explained by this hypothesis.

In contrast to the argument that climatic stability in the tropics has led to increased species richness, some ecologists have proposed the opposite, suggesting that speciation and species coexistence are more likely to occur with more frequent disturbances (see Chapter 13). One reason to believe that such a relationship might exist is that disturbance may reduce the chance for competitive exclusion, allowing more species to coexist. However, there is no evidence that tropical communities differ from temperate ones in frequency of disturbance, nor is there any evidence that higher rates of disturbance support more species at large spatial scales, and the hypothesis is vague about the scale or type of disturbance that might accomplish this (see Chapter 13). The disturbance hypothesis cannot explain the large difference in species richness between Brazil and the United States.

Geographic isolation is known to promote speciation (see Chapter 9). However, both temperate and tropical species should have been subjected to similar kinds of isolation during recurrent periods of glaciation and warming. This pattern of speciation would result in species within a genus that have adjacent or overlapping ranges, a pattern that we find in the distributions of species in northern boreal and polar regions (Figure 19.3).

There is evidence, however, that climate variation is linked to the latitudinal diversity gradient. Roland Jansson and T. Jonathan Davies (2008) found that the highest rates of speciation were not in lowland tropical rainforests, where the greatest number of species are found. Rather, speciation was greatest in adjacent regions such as the slopes of the Andes, where climates were more

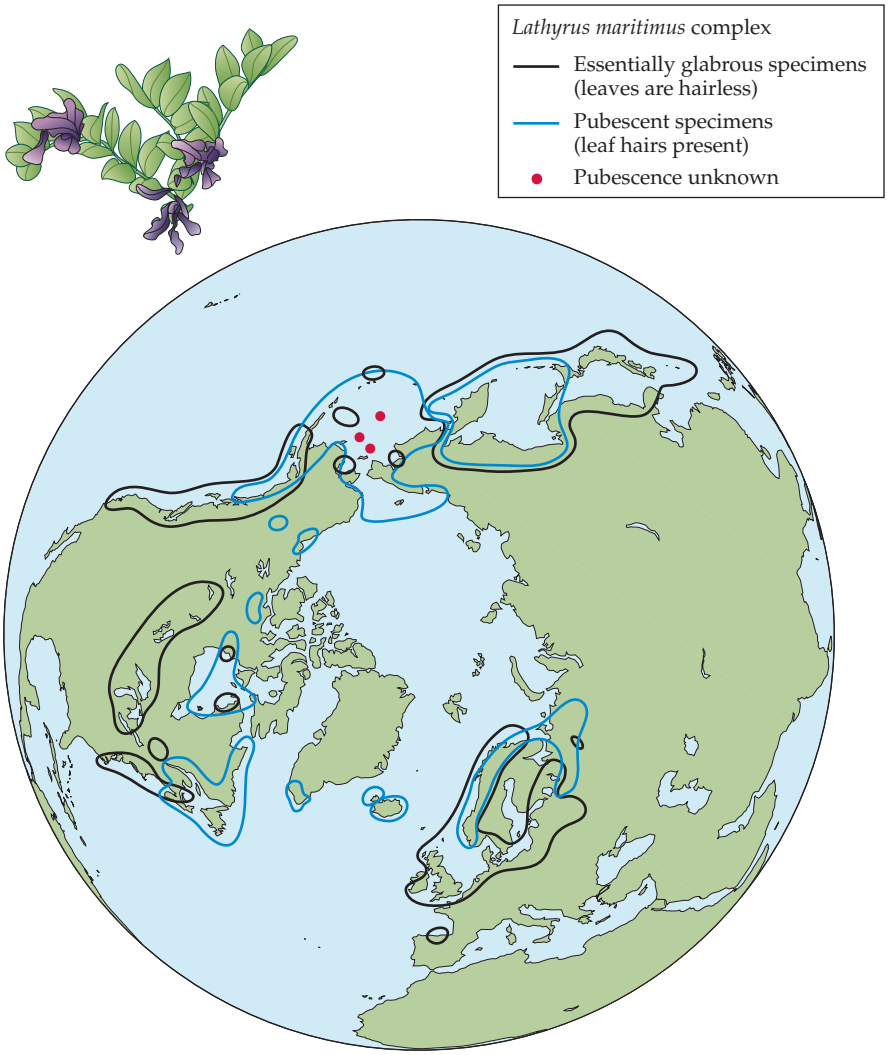



**Figure 19.3** Distribution of the *Lathyrus maritimus* (Fabaceae) species complex around the Northern Hemisphere polar region. Populations of this plant in different regions differ morphologically, particularly in whether and where their leaves have hairs, and are likely to differ genetically. The circumpolar distribution of this species complex is typical of many polar and boreal species and genera. (Adapted from the Linnaeus Server, Swedish Museum of Natural History (NRM), Stockholm. Last update: 10 October 1998. [http://linnaeus.nrm.se/].)

variable. Perhaps high tropical rainforest diversity might be due in part to species immigrating from nearby slopes with more variable climates.

In all cases, there are seemingly insurmountable challenges in extracting simple explanations for the latitudinal diversity gradient: there is only one Earth, with one history, and with many different factors occurring together. That means that there are no true replicates, and there are limited chances to separate the many factors that occur together (in statistical terms, they are confounded). Nevertheless, we can piece together some of the most likely explanations.

### There are also regional and global patterns of β-diversity

Beta (β) diversity is a measure of how species composition changes as you travel from one area to another (see Chapter 12). Patterns of α-diversity and β-diversity may be very different in different places, and the reasons for high or low α- or β-diversity may also differ. We do not have good explanations for all of these contrasting patterns, or even general documentation for the circumstances that cause different patterns. As with α-diversity, different patterns of β-diversity may have both ecological and evolutionary explanations, and it is plausible that several factors together may be important in any one situation. The great diversity of the Cape flora of South Africa, for example, is due both to very high numbers of species in a small area—high α-diversity—and to very different groups of species from one place to another—high β-diversity.

In a study of global patterns of species richness, Samuel Scheiner and José Rey-Benayas (1994) found that the greatest β-diversity occurred at 30° N and S latitude, where the great deserts of the world exist (Figure 19.4; see Figure 16.10 and Figure 16.16). While tropical regions have the greatest species richness at a local level, the same species are often found across an entire region. Within the tropics, however, β-diversity varies. Richard Condit and associates (2002) found that β-diversity was substantially higher in tropical forests in Panama than in western Amazonia. Across North America north of Mexico, Hong Qian and Robert Ricklefs (2007) showed, β-diversity decreases as you move northward; these patterns of spatial turnover with distance hold not only at the species level, but for plant genera and families. Northern regions are more spatially homogeneous at the species level and also at the level of genera and families.

Interestingly, two very different environments with low productivity—hot deserts and cold Arctic tundra—have very different effects on β-diversity. In hot deserts, β-diversity is high, meaning that species composition can be very different over very short distances. In contrast, in Arctic tundra, the same few species are found everywhere over very large regions. Scheiner and Rey-Benayas

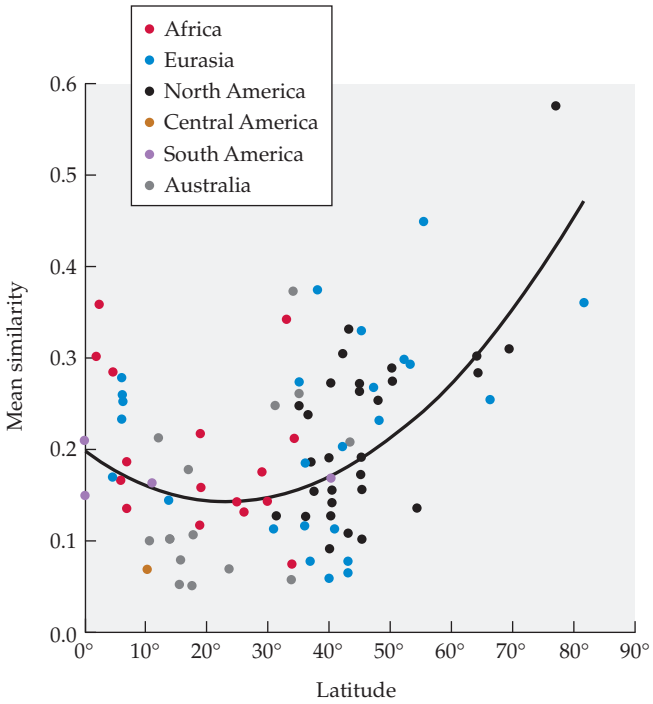



**Figure 19.4** A global pattern in β-diversity of vascular plant species. Each point represents the mean similarity across sites in a particular landscape. Latitude is given in absolute degrees, so landscapes in the Northern and Southern Hemispheres are superimposed. Lower mean similarity indicates that different sites in a landscape are more different in species composition. (Data from S. M. Scheiner and J. M. Rey-Benayas. 1994. *Evol Ecol* 8: 331–347.)

speculated that although both environments have limited resources, their limiting environmental factors differ in two important respects: in the extent of their environmental heterogeneity and in the extent to which plants compete for limiting factors. Water (the limiting resource in hot deserts) can vary greatly in availability over very short distances and can be the object of intense competition (see Chapter 3 and Chapter 10). In contrast, low temperatures and a short growing season (the limiting conditions in Arctic tundra) are the same everywhere and are not resources for which plants can compete.

## 19.3 There Are Distinctive Regional and Continental Patterns of Plant Biodiversity

In addition to the latitudinal gradient, many other patterns of diversity have been observed at regional and global scales. Mountain ranges often have been observed to have the largest number of species at mid-elevations. Within continents, there are longitudinal gradients of species richness; in the boreal forest of North America, species richness within sites is higher in central Canada

than it is in either western or eastern regions (Figure 19.5). The regional species richness of a particular taxonomic group may be very similar to or extremely different from the species richness of that group in similar habitats in other parts of the globe (Schluter and Ricklefs 1993). For example, the diverse and globally widespread plant family Ericaceae (see Figure 11.29) has common species in northern Europe and North America but relatively low species numbers in both regions, and it has very high numbers of much rarer species in southwestern China and southern Africa.

## Continents at the same latitudes differ in species diversity

If we compare similar regions in eastern Asia, North America, and Europe, we find that eastern Asia has the most plant species, while Europe has the fewest. This pattern holds over a range of taxa. For instance, if we look at the number of tree species across all plant families, eastern Asia has the greatest number of species and Europe has the fewest. These patterns also hold for orders, families, and genera of trees (Latham and Ricklefs 1993).

What can explain these patterns? One process involves patterns of extinction. There were differences in rates of species extinction among the continents over the repeated glaciations and warming periods of the Pleistocene. A key difference between Europe and the other two regions has to do with an accident of topography. The Alps, the main mountain range in Europe, run in an east-west direction, whereas the principal mountain ranges in eastern Asia and North America run north-south. During periods of glaciation, climate zones are pushed southward. At the height of the most recent glaciation in North America, for example, the southern edge of the ice sheet was in Wisconsin, and boreal forests extended as far south as the Gulf Coast (see Chapter 17). As climate zones moved southward, the plants adapted to those climates followed them by dispersing southward. The southern portions of the United States provided refuges for many species. In Europe, however, the Alps created a barrier to southward migration. Species that were adapted to warmer conditions were unable to migrate over the tops of the mountains to escape from the cold, and many became extinct. As the glaciers retreated, the species in North America and Asia that had retreated southward dispersed northward again. But in Europe, those species had disappeared and were not available to recolonize the newly ice-free temperate areas. Extinction rates in eastern Asia were lower because glaciers there in the Pleistocene were limited to high mountains and plateaus, and the glaciated area was far smaller than in North America.

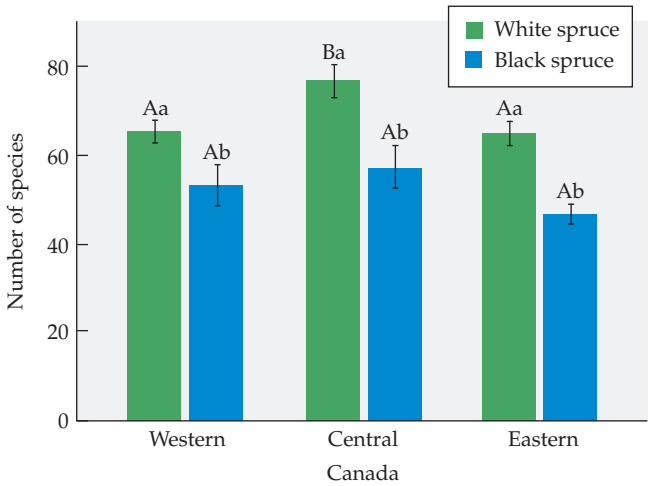



**Figure 19.5** Number of species (mean ± 1 standard error) in 30 × 30 m² plots in boreal forest communities of Canada. Species richness is higher in the central region, especially in white spruce communities. Bars with different capital letters indicate statistically significant differences among regions within community types. Bars with different small letters indicate statistically significant differences within regions among community types. (Data from H. Qian et al. 1998. *Plant Ecol* 138: 161–178.)

The end result of these events, repeated during each of the Pleistocene glaciation and warming periods, is the greatly reduced species diversity of Europe relative to North America (Table 19.1); eastern Asia had the fewest extinctions. Northern temperate and boreal communities in North America contain about 50% more vascular plant species than equivalent European ones (approximately 18,000 versus 12,000: Niemelä and Mattson 1996). Annik Schnitzler and colleagues (2005) compared temperate floodplain forests in the Mississippi River and Rhine River valleys and found fewer species, genera, and families in Europe. Fossil evidence supports this explanation: European forests before the Pleistocene had much greater species richness than they do now (Latham and Ricklefs 1993). This explanation is an example of historical contingency: the pattern of diversity is the result of a particular sequence of historical events in a particular place, not of a general or fundamental process that explains all such patterns.

The difference in richness between the floras of eastern North America and eastern Asia is also thought to be the result of evolutionary history, but one operating on a longer time scale and depending on rates of evolution—the rise of orders, families, and species—rather than extinction (Latham and Ricklefs 1993; Qian and Ricklefs 1999; Ricklefs et al. 2004). The temperate regions of eastern North America and eastern Asia are similar in geographic area and climate. They share many of the same plant families and even some of the same genera. Yet China contains 60% more species of vascular plants than the United States (29,188 vs. 17,997). Much of this difference is caused by the fact that plant families in

**TABLE 19.1** The number of woody species of various genera that went extinct in Europe during the Pleistocene but are still found in North America

| Genera | Number of species |
|---|---|
| Gymnosperms | |
| *Chamaecyparis* | 3 |
| *Sequoia* | 1 |
| *Taxodium* | 2 |
| *Thuja* | 2 |
| *Torreya* | 2 |
| *Tsuga* | 4 |
| Angiosperms | |
| *Asimina* | 3 |
| *Carya* | 11 |
| *Diospyros* | 2 |
| *Lindera* | 2 |
| *Liquidambar* | 1 |
| *Liriodendron* | 1 |
| *Magnolia* | 8 |
| *Morus* | 2 |
| *Nyssa* | 3 |
| *Persea* | 1 |
| *Robinia* | 4 |
| *Sabal* | 3 |
| *Sapindus* | 2 |
| *Sassafras* | 1 |

*Source*: P. Niemelä and W. J. Mattson. 1996. *Bioscience* 46: 741–753. Based on data from E. L. Little. 1979. *Handbook nr 541*, pg 354–359. Washington, DC: USDA Forest Service.

China generally have more species than do the same families in the United States.

Hong Qian and colleagues (2017) found that both species diversity and phylogenetic diversity were higher in east Asia than in North America, and this pattern was true not only for angiosperm trees, but also for herbaceous and shrubby angiosperms. One major factor explaining these differences in species richness and phylogenetic diversity is a result of continental-scale geographic relationships. The southern United States is separated from tropical regions by large bodies of water to the east and south, and by very dry conditions in the arid west. The Isthmus of Panama connecting North and South America has only been above sea level for about 3 million years. In contrast, southern China has been directly connected to a large expanse of tropical and subtropical regions for tens of millions of years. The temperate regions of eastern Asia, Qian and colleagues hypothesize, were repeatedly colonized by species from adjacent subtropical and tropical areas. These species then underwent extensive adaptive radiation—speciation and the evolution of higher taxonomic levels—in the temperate regions. Qian and colleagues also suggest that some taxa with northern affinities may have dispersed across Beringia (the area between Siberia and Alaska that was above sea level during parts of the Pleistocene) from eastern Asia to North America, explaining the congruence among many angiosperm tree genera in these two regions. What makes the "out of the tropics" argument especially compelling is the finding that the greatest differences in species richness between the two continents are found in older taxonomic groups with strong tropical affinities. Europe suffers from the same limitation as North America in this regard because the Mediterranean Sea and the arid expanse of North Africa create a barrier between it and tropical and subtropical floras. Thus, a coincidence of geography may have been responsible for differences in species evolution over very long periods as well as extinction rates on a shorter ecological time scale. Other explanations have also been suggested. Eastern Asia has greater climatic and topographic heterogeneity than eastern North America, which might have fostered evolutionary divergence and ecological adaptation to different habitats and conditions. Over a long time scale, it is possible that when the Indian tectonic plate collided with Eurasia, it brought with it new taxonomic groups to Asia; the collision itself led to mountain building and created a complex topography.

However, in a study that incorporated both modern climate and glacial-interglacial climate change, tree species endemism (species and phylogenetic endemism; see below) was found to be highest where long-term climates were more stable, with warmer and wetter modern conditions (Feng et al. 2019). Understanding the factors giving rise to higher richness of endemic species is another way to consider the evolution of diversity in different environments.

As with global patterns, different continental patterns of species richness exist for different functional groups of plants. For trees in North America, the highest number of species is found in the southeastern United States, but not the most southern area, with steady decreases to the north and west (Figure 19.6). (The general pattern of tree species richness echoes the pattern for endemism discussed below.) Other patterns relate to species distributions and ranges. For instance, Jason Pither (2003) showed that North American tree species with the largest ranges are also those found where January minimum temperatures are the coldest (Figure 19.7). The minimum environmental temperature a species can tolerate is strongly related to its resistance to freezing. A similar,

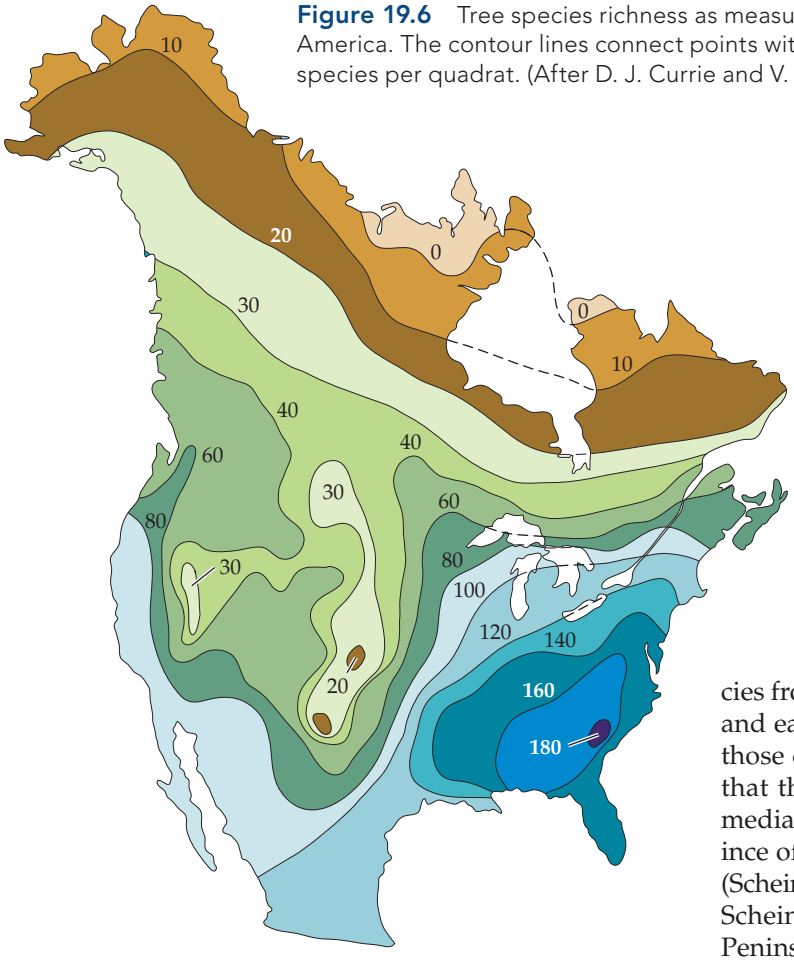



**Figure 19.6** Tree species richness as measured in 2.5° × 2.5° quadrats across North America. The contour lines connect points with the same approximate number of species per quadrat. (After D. J. Currie and V. Paquin. 1987. *Nature* 329: 326–327.)

but smaller, effect was found for monthly precipitation: species able to withstand dry conditions tend to have larger ranges.

## Transition zones may have higher diversity due to overlaps in species' distributions

Sometimes ecologists have found increased species richness in **transition zones**, areas of contact between communities, regions, or biomes. This increase occurs because the transition zone contains species from the areas on either side of it (Shmida and Wilson 1985). For example, the area around the Great Lakes in North America is home to the hemlock–white pine–northern hardwood forest (see Figure 18.17 and Figure 18.21), which contains species from both the eastern deciduous forest to the south and east, and the boreal forest to the north. A study of those communities in northern lower Michigan found that the region was more species rich than areas immediately adjacent, to the north in the Canadian Province of Ontario, and to the south in southern Michigan (Scheiner and Istock 1994). José Rey Benayas and Samuel Scheiner (2002) found a similar pattern in the Iberian Peninsula, where there are two transition zones (Figure 19.8A). One transition zone occurs along the Atlantic coast and the Pyrenees mountains, where the Eurosiberian flora abuts that of the Mediterranean region. The other is found in the southeast, in the region adjacent to Africa. These two areas have more species than others in Iberia because this is where the most species have overlapping ranges (Figure 19.8B); the causes of these peaks in range overlap are unknown.

Within the boreal forest of North America, there is a longitudinal pattern of species richness, with the highest richness in the central portion of the region, intermediate richness in the west, and the lowest richness in the east (Qian et al. 1998). The differences in species composition along this longitudinal transect appear to be related to dispersal patterns following the most recent glaciation (see Chapter 17). The greater species richness in the middle of the continent may be due to an overlap

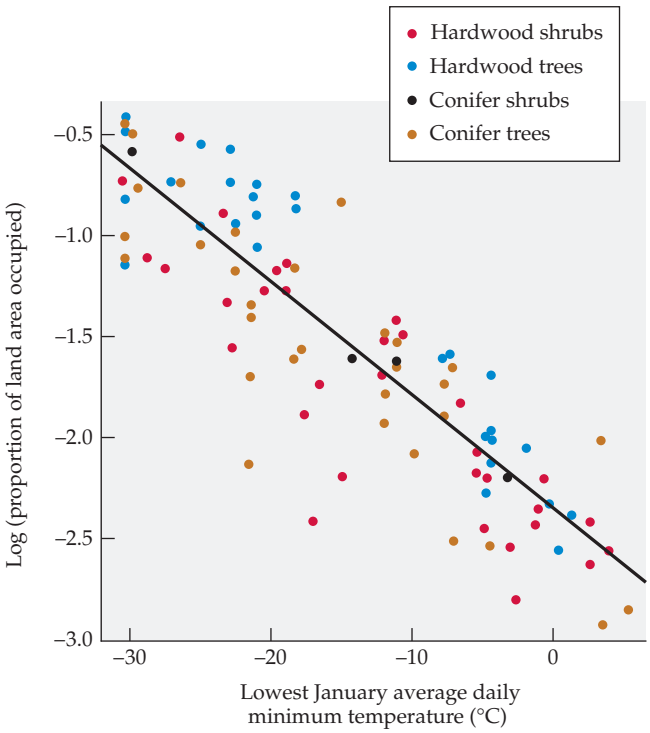



**Figure 19.7** The relationship between the lowest value of January average daily minimum temperature experienced within a species' geographic range and its range size as indicated by the proportion of area occupied for hardwood shrubs, hardwood trees, conifer shrubs, and conifer trees in North America. (After J. Pither. 2003. *Proc R Soc London* 270B: 475–481.)

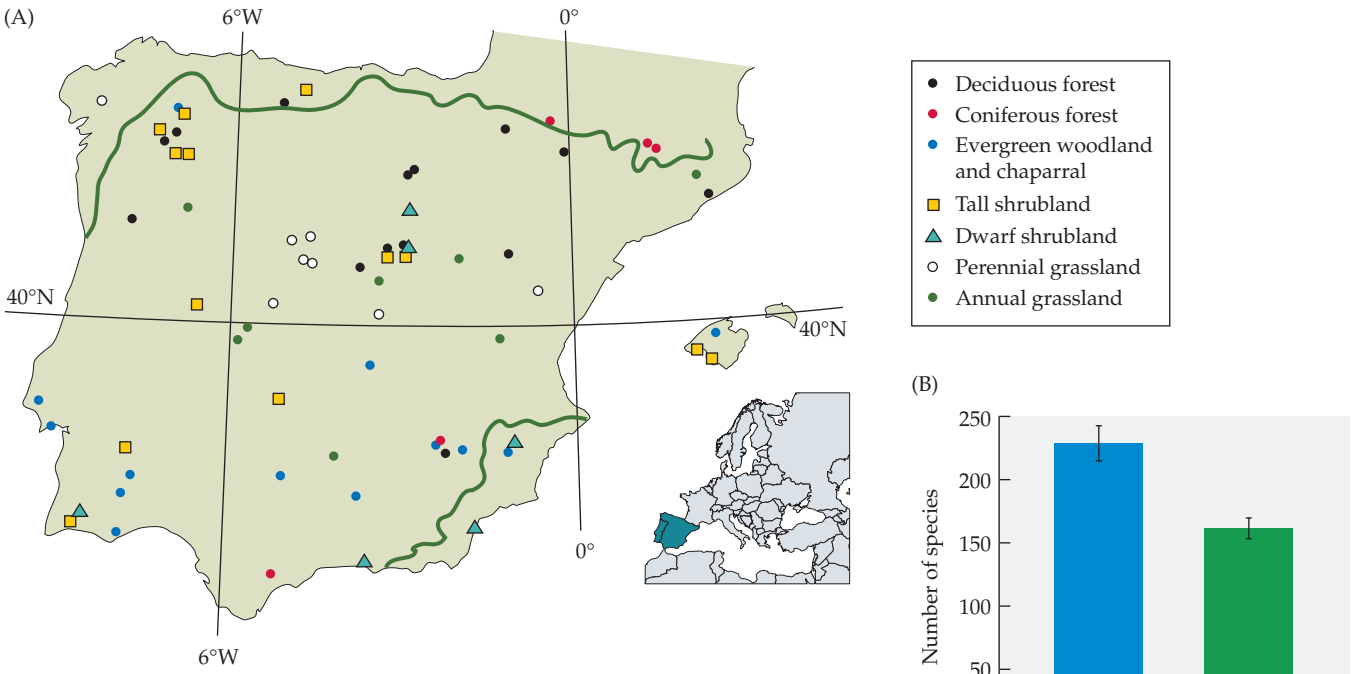



**Figure 19.8** (A) Map of Iberia, showing the locations of the two transition zones (areas along green lines). Each point is the center of a large sampled landscape. (B) Landscapes in the transition zones contain more species (mean ± 1 standard error) than those elsewhere in Iberia. (After J. M. Rey Benayas and S. M. Scheiner. 2002. *J Veg Sci* 13: 245–258.)

of species ranges caused by the different pathways from the east and west taken by plants expanding their ranges northward from unglaciated refuges in the southeast and southwest after the glacial retreat.

Another kind of overlap is temporal; it occurs in landscapes that have highly seasonal continental climates. This seasonality allows for the coexistence of species with different life history patterns (see Chapter 2 for a discussion of spring understory species in North American temperate forests); the heterogeneity creates opportunities in time, rather than in space.

## Mountains and mountainous regions have distinct but complex patterns of species diversity

Mountains and mountain ranges present intriguing patterns of plant diversity. Many ecologists have found that there are typically more species at the mid-elevations on mountains than either higher or lower on a gradient from the base to higher up (Figure 19.9). There are also many cases where this has not been found. It has been hypothesized that this mid-domain peak in richness occurs because of the greater overlap of species ranges near the midpoint of a mountain due to geometric constraints. Ji-Hua Wang and colleagues (2018) studied

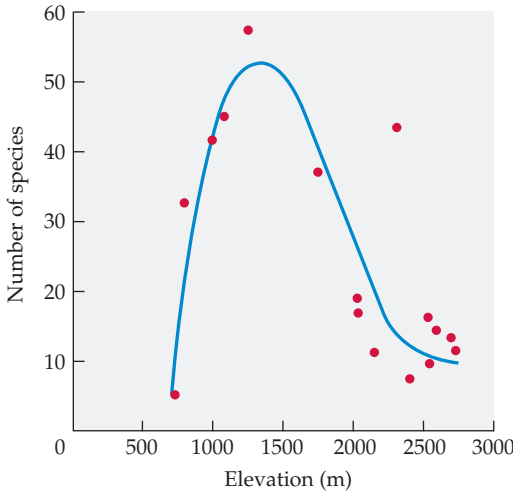
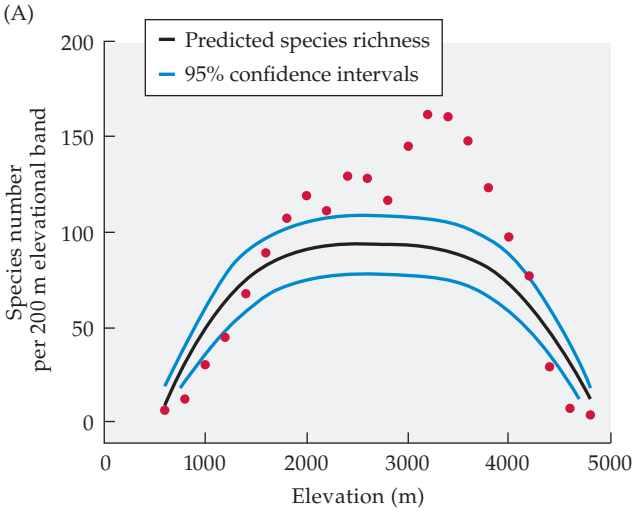



**Figure 19.9** Plant species richness along an elevational gradient in the Santa Catalina Mountains in southern Arizona. Both the low- and high-elevation communities—the Sonoran Desert and boreal forests, respectively—are in extreme environments, resulting in a peak in diversity at intermediate elevations. (Data from R. H. Whittaker and W. A. Niering. 1975. *Ecology* 56: 771–790; after J. H. Brown. 1988. In *Analytical Biogeography*, A. Myers and R. S. Giller [Eds.], pp. 57–89. Chapman and Hall: London.)



From C.E. Hughes. 2017 *PNAS*: 114, 4275–4277. © Jian Huang

**Figure 19.10** (A) Number of species per 200m elevational band in the family Ericaceae across the provice of Yunnan, China. Red dots indicate observed species richness. Black lines represent predicted species richness and blue lines are 95% confidence intervals based on a randomization model. The higher numbers of species at 3000–4000 m elevation are largely due to very high species richness of shrubs at these elevations. (B) *Rhododendron traillianum* (Ericaceae, with pink flowers) and *Picea likiangensis* (Pinaceae) at 4300 m elevation, close to tree line at the Shangri-La Great Snow Pass, in the Hengduan Mountains, Yunnan, in southwestern China. (A after J. H. Wang et al. 2018. *Forests* 9: 511. CC BY 4.0)

species richness in the Ericaceae, a species-rich group, across the province of Yunnan in southwestern China, an area of very high global species and topographical diversity (Figure 19.10A). They found that species richness in this family was stongly hump-shaped, with a peak about two-thirds up the elevational gradient, and was lower below and above that elevation (Figure 19.10B). Temperature and moisture were the dominant factors determining diversity patterns, rather than geometric constraints (see below), and they inferred that competition also contributed to the pattern of species richness with elevation.

Mountains in different parts of the world have very different patterns of gradients in vegetation at different elevations. Tropical mountains often have tropical evergreen forests (see Figure 18.11) at their lowest elevations, transitioning to colder and drier conditions as elevations increase, with cloud forests (see Figure 18.12) and then paramo (Figure 19.11) at very high elevations. Mountains in desert regions have desert vegetation at their base and, with rainfall increasing as elevation increases, woodlands and coniferous forests higher. Mountains in northern temperate and boreal regions may have coniferous forests replaced with tundra at higher elevations (see Figure 18.4). Diversity varies, of course, as the form of the vegetation changes. If mountains are sufficiently high, even in the tropics, they can have glaciers at their peaks, with no terrestrial plant life and no plant diversity at all.

Different plant functional groups and different taxa may show distinct patterns of diversity along elevational gradients. On an elevational gradient from 30 m to almost 3000 m in Costa Rica, James Watkins and his collaborators (2006) found 263 species of fern, including ones growing on the ground (terrestrial species), low on tree trunks (low-trunk epiphytes), and at the tops of trees (canopy epiphytes). They found that the overall species richness of all of the ferns had a hump-shaped distribution with a peak at 1000 m elevation; this was also true for the canopy and low-trunk epiphytic species. Terrestrial ferns increased in species richness from low to middle elevations, but remained at high diversity at higher elevations and did not have a hump-shaped distribution.

© Martin Tychl/Laidbacktrip.com

**Figure 19.11** The paramo is a high-elevation ecological system; the tall plants are *Espeletia* spp. (frailejones, Asteraceae). The site is in Los Nevados National Park, Colombia.

Mountainous areas contain many different habitats and microhabitats, from broad lowland valleys and riparian environments to slopes at different elevations; microhabitats may include different aspects (e.g., north- or south-facing slopes, which differ in moisture, light, and other characteristics) and different soils. This environmental heterogeneity can result in differences in interactions with pollinators, herbivores, and other organisms. The varied terrain of mountains creates opportunities for the physical isolation of populations. All of these conditions may foster speciation and niche divergence (see Chapter 14) resulting in higher diversity in some mountain regions compared with adjacent areas.

## 19.4 Regional Diversity and Local Diversity Can Influence One Another

How do regional and local processes interact to determine local and regional diversity? Ecologists have made progress in understanding the mechanisms for the connection between the diversity of a region and the diversity of a community, although our understanding is still incomplete. Is local species richness set by the number of species available in the regional species pool, or is it determined mainly by processes acting locally? We can test this by sampling a series of regions, measuring both the total number of species in each region and the number of species in a small subset of that region, and then plotting local species richness as a function of regional species richness (Figure 19.12A). If local species richness is determined by the size of the regional species pool, then we should see a straight-line relationship, in which a larger regional species pool leads to more species locally. On the other hand, if local processes, such as competition, limit the local number of species, the relationship should show a slower rise or level off at some point, adding more species to the regional pool should have little or no effect on local species richness. In an analysis of plant communities in Estonia, Meelis Pärtel and colleagues (1996) found that local species richness was a function of regional species richness, with no leveling off of the relationship at the highest regional species richness (see Figure 19.11B). Similarly, in a study of palm species richness (family Arecaceae) in the western Amazon, Thea Kristiansen and associates (2011) found that regional richness was the single largest explanatory factor. One mechanism that would explain this relationship is if the regional species pool contributes a continuing supply of propagules to local communities. To test this idea, Jonathan Myers and Kyle Harms (2009) performed a meta-analysis of 28 studies that artificially increased seed input into local communities and found that species richness increased on average by 81%, with even higher numbers in plots with some sort of disturbance.

To examine the mechanisms that link local and regional species richness, Jonathan Shurin and Emily Allen (2001) asked how local species interactions (competition and predation) contrast with regional patterns to determine species richness. They found that predation expands the range of environments in which competitors can coexist at a regional level, promoting higher levels of regional

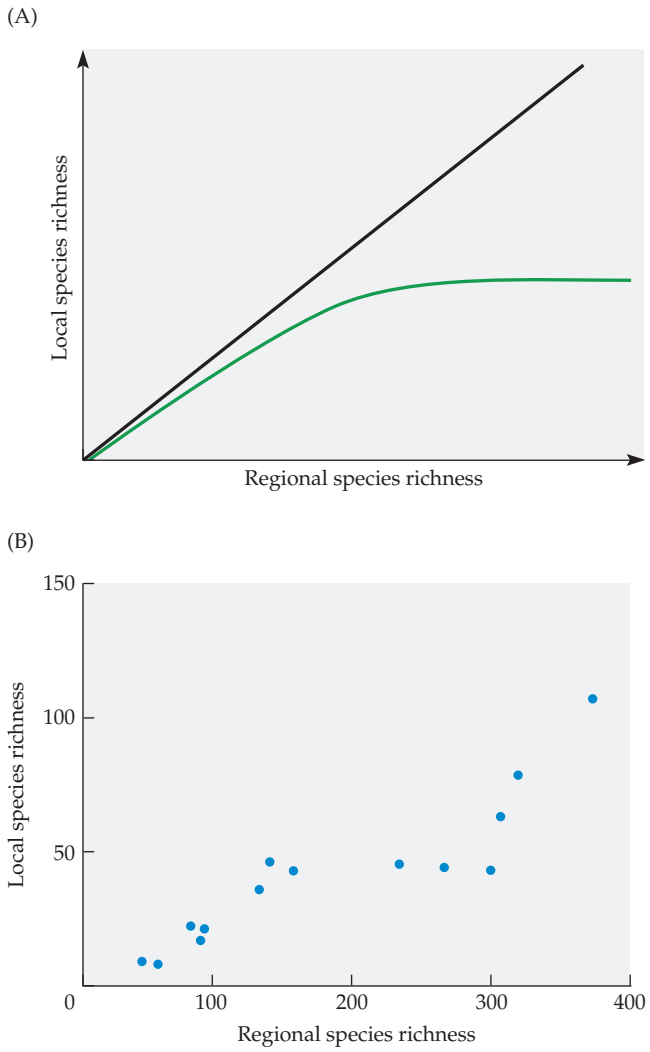



**Figure 19.12** (A) Two possible relationships between the size of the regional species pool and local species richness. If local richness depends only on the size of the regional pool, then local richness should increase with regional richness (black line). If local richness depends in addition on local processes, then at some point, adding species to the regional pool should not result in any more species locally (green line). (B) The relationship between regional and local species richness for 14 plant communities in Estonia, including grasslands, shrublands, and forests. (After M. Pärtel et al. 1996. *Oikos* 75: 111–117.)

species richness by preventing competitive exclusion in some locations (see Chapter 10), and reported a reciprocal relationship between local and regional processes. Their study used simulations of animal communities; no such comparison has been carried out for plant communities, although the results might prove valuable.

## Endemism, isolation, and global biodiversity hotspots

**Endemic species** are those found only in a single, limited geographic area. Typically that area is small, but it can be a larger region. For example, the cactus *Carnegiea gigantea* (see Figures 3.10A, 13.17) is endemic to the Sonoran Desert. Usually the area in which an endemic species occurs is where it originally evolved, but not always. Species that were once widespread but have become extinct in most places may only exist as relicts in areas different from their places of origin. When we say that a species is endemic, it is best to specify the area to which it is endemic, which usually is implied to be a limited area.

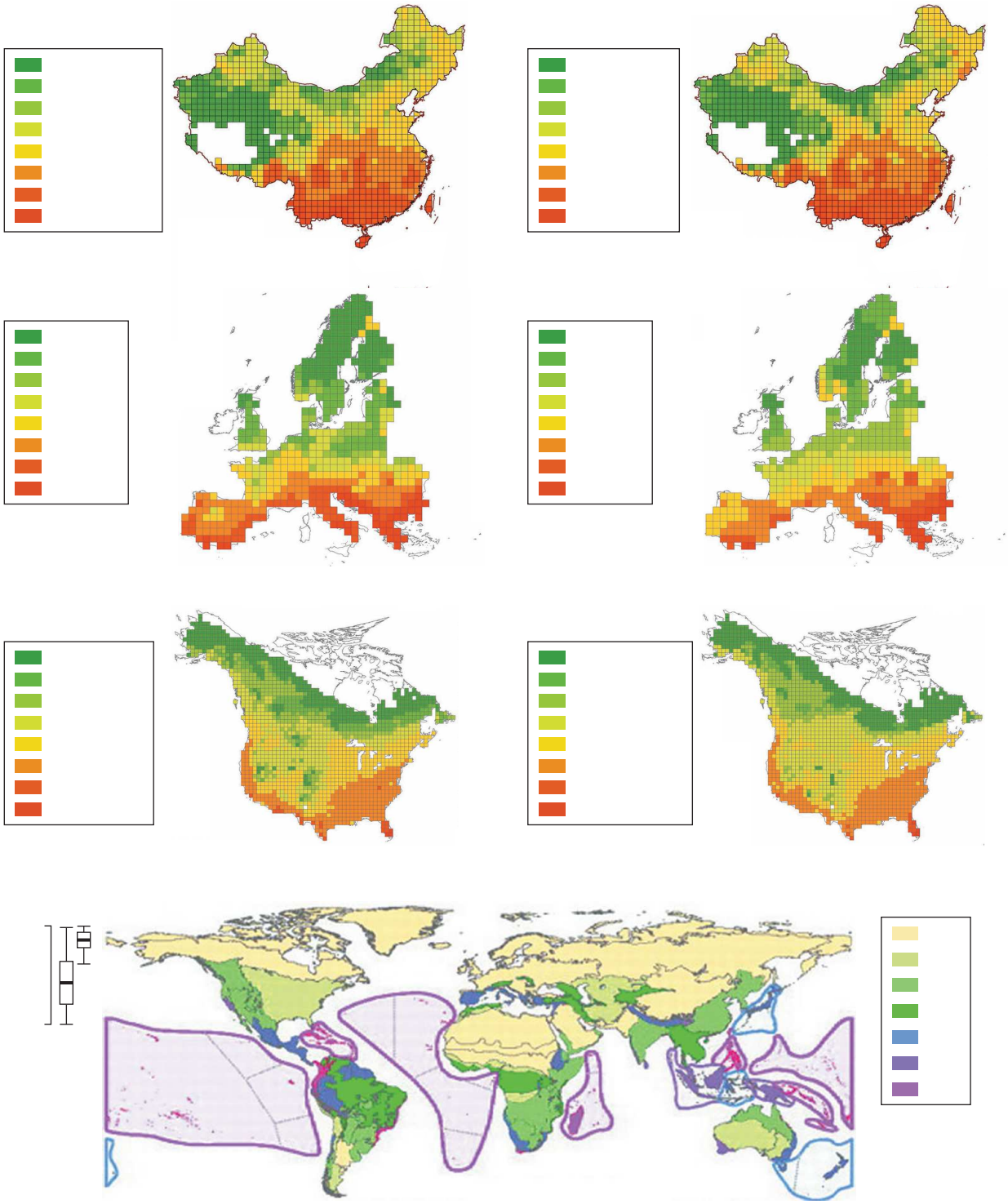
Because endemic species have limited ranges and numbers of individuals, they are of high priority for conservation, and regions that have high numbers of endemic species can be priorities for conservation. Such areas can be identified using weighted species endemism (SE). Species with smaller ranges found in that area are "weighted" or counted more heavily in this index than species with larger ranges, so the value of the index becomes greater for the area being evaluated as the overall ranges of at least some of the species in it decreases. In an area where most species are widespread the value of SE will be low, and an area where most of the species are endemic to that area will have a high SE value. For example, if an area contains ten species and all are endemic to that area, SE = 10; if only one species is endemic and the other nine are widespread, and the area being considered is only 1% of their range, SE = 1.09. (This weighting is similar to the way that abundances are weighted in calculating the Hill diversity index; see Chapter 12.) Another measure of endemism is phylogenetic endemism (PE), which weighs species by their evolutionary distinctiveness (phylogenetic branch lengths; see Figure 12.5) as well as the geographic area occupied by each clade (Kosman et al. 2019). These measures can have consequences for conservation policy (discussed at greater length later in this chapter). For example, if a species without close living relatives became extinct, this would cause a greater loss of phylogenetic diversity than the extinction of a recently evolved species from a diverse lineage (Laity et al. 2015). The Australian tree *Wollemia nobilis* (Wollemi pine, Araucariaceae), the sole living descendent of a 150- to 200-million-year-old lineage (and not a pine at all), is such a case, as it is critically endangered and facing likely extinction.

Gang Feng and collaborators (2019) calculated the continental-scale patterns of SE and PE for angiosperm trees across the Northern Hemisphere to look at the role of past and present climates and climate variability in shaping those patterns. They found that endemism was influenced by current climate conditions, with both SE and PE greatest where modern conditions are warmer and wetter and where glacial-interglacial climates were more stable. Evolutionary difference, measured by PE, was highest in southwestern China and at the tip of Florida in the United States, while it was much lower in Europe (Figure 19.13A–F).

In the U.S., Florida, Texas, and California have the highest numbers of endemic species. For example, Florida has 2840 native and naturalized vascular plant species, of which 215 are endemic. These plants are found largely in xeric, fire-prone habitats. In some cases, a species may be endemic to a local area because it is the product of recent speciation and has not spread to other areas. If speciation rates are high in a region and species cannot spread, that region will become species rich and have a large number of endemic species. Gerold Kier and his collaborators (2009) found a strong correlation between endemism richness of plants and vertebrates. There are many continental areas with endemic species, and some areas have very large numbers of endemic plants. Continental areas with large numbers of endemic species include the Western Ghats of India, Baja California in Mexico, southwestern Australia, the Cape region of South Africa, and parts of the Mediterranean (Figure 19.13G). Most of these areas are partially or fully isolated in various ways, such as being bounded by deserts or oceans.

The patterns we have discussed so far involve gradual changes in species diversity over large spatial extents or ecological gradients. Diversity can also vary in a patchy manner, in which some areas have more species than other nearby areas. Endemic plants may be restricted to unusual habitats, such as serpentine barrens (see Box 4A). Sometimes that greater richness is due to speciation fostered by geographic isolation (see Chapter 9). Oceanic islands are an obvious case of extreme isolation (Table 19.2). The Pacific islands of Hawaii, New Zealand, and New Caledonia have been isolated for millions of years, and all have very high species richness and very high proportions of endemic plants. All of these islands are also topographically highly complex, with many different habitats and climatic conditions, factors which likely increased speciation and fostered greater endemism.

Madagascar is a large island off the southeastern coast of Africa. It separated from that continent and has been isolated from all other landmasses for about 88 million years. It supports some of the highest plant diversity on Earth, with over 10,000 native plant species, more than



80% of which are endemic to Madagascar (Figure 19.14). There is also a high degree of phylogenetic diversity there, with five endemic plant families and numerous endemic genera. Madagascar also has some of the world's most threatened habitats. Although forest destruction began there hundreds of years ago or more, it accelerated greatly in the twentieth century. Between 1960 and 1990, almost half of the forests of Madagascar were destroyed, and today only a small remnant remains.

The general problem of explaining variation in speciation rates is one of the greatest challenges in evolutionary biology. Other than geographic isolation, there

◀ **Figure 19.13** Maps of tree species endemism (SE) (A, C, E) and phylogenetic endemism (PE) (B, D, F) in China, Europe, and North America in 100 km grid cells (see text for explanation of SE and PE). Note that the scales are very different for each continent. (G) Global pattern of the endemism richness (ER) of vascular plant species in range equivalents per 10,000 km$^2$. The greatest concentrations of endemic species are found on islands, but other areas such as Central America and the Iberian Peninsula of Europe also have large numbers. Circled areas indicate endemism richness for the region circled as a whole. (A–F after G. Feng et al. 2019. *Global Ecol Biogeogr* 28: 1393–1402; G after G. H. Kier et al. 2009. *Proc Natl Acad Sci USA* 106: 9322–9327.)

is little agreement about the factors resulting in speciation. Factors that have been hypothesized to cause high levels of speciation or endemism in plants include protection from herbivory, low productivity, the existence of many local topographic barriers, habitat diversity, a low frequency of major disturbances, and isolation. Explaining speciation patterns entails understanding not only ecological factors, but the geologic, geographic, and evolutionary history of the lineage and the area in which it lives and has evolved.

The extremely high levels of plant diversity and endemism found in the Cape region of South Africa illustrate many of these issues (see Box 19A). Such centers of diversification are of critical importance in conservation.

Courtesy of J. Gurevitch

**Figure 19.14** Ranomafana National Park, in the southeastern part of Madagascar, supports a high diversity of native and endemic plants and animals.

The evolutionary patterns in the Cape region of South Africa are complex, and much is still being learned about them. The fynbos in the Western Cape (also called the Cape floristic region) is particularly rich in species in the families Proteaceae, Ericaceae, Restionaceae, and Iridaceae. While some of the fynbos genera and species have arisen more recently, some of the radiations arose in the Oligocene (Verboom et al. 2009). The evolutionary radiations that gave rise to the exceptionally high diversity

**TABLE 19.2** Diversity of native (endemic + nonendemic) and endemic vascular plant species and introduced/invasive vascular plant species on selected islands and island archipelagos

| | Introduced | Native | Endemic | Location |
|---|---|---|---|---|
| Madagascar | 517 | 8000–14000 | 5000–12000 | Indian Ocean |
| Cuba | 577 | 6514 | 3229 | Caribbean Sea |
| Philippines | 628 | 9200 | 3500 | Philippine Sea/western Pacific Ocean |
| New Caladonia | 1400 | 3000–3400 | 2500–2650 | Pacific Ocean |
| New Zealand | 411–2500 | 2357–2371 | 1944 | Pacific Ocean |
| Hawaiian Islands | 869 | 956–1039 | 600–940 | Pacific Ocean |
| Fiji | 521 | 1769 | 812 | Pacific Ocean |
| Cyprus | 143 | 1738 | 162 | Mediterranean Sea |
| Seychelles | 247 | 250–1640 | 250 | Indian Ocean |
| Sri Lanka | 365 | 3368 | 900 | Indian Ocean |
| Andaman and Nicobar | 33 | 2270 | N.A. | Indian Ocean |
| Galápagos | 754 | 447–543 | 211–236 | Eastern Pacific Ocean |
| Cook Islands | 286 | 180–278 | 3 | Pacific Ocean |
| Cocos | 71 | 192 | 31 | Indian Ocean |
| Ascension | 200 | 10–20 | 11 | South Atlantic Ocean |

*Note*: Data are estimates; some data were not available (N.A.). Data from various sources ranging from 1992 to 2018, compiled by Qinnan Zhu, unpublished MA thesis from Stony Brook University.

of the fynbos have been attributed to the development of a mediterranean climate or to the development of extremely low-nutrient soils in the Cape mountains as sandstone was exposed after the breakup of Gondwana (van Santen and Linder 2020).

Félix Forest and his associates (2007) looked at the evolutionary relationships of the plants across the Cape region and considered what they mean for conservation strategies. They point out that the evolutionary history and ecological context are very different for the Western Cape and Eastern Cape. The Western Cape has a mediterranean climate, with winter rainfall (relatively recently in geologic time), while the Eastern Cape has rainfall all year. The number of species is about twice as high in the Western Cape, and within this region the highest species numbers and numbers of endemic genera are found in the mountainous areas; the region is adjacent to the Succulent Karoo, which also has very high species richness and endemism. The Eastern Cape is adjacent to another area of very high biodiversity, Maputaland, and there are overlaps in species between the two areas. The diversity and endemism of the taxa in the fynbos of the Western Cape are due to a smaller number of families and genera with a spectacular evolutionary radiation of species in those genera. In contrast, the eastern region is composed of genera that are not so closely related to one another. Both areas are critically important for conservation; the western region is more species rich, while the eastern region has a greater potential for preserving unique evolutionary lineages. If one only considered species numbers or numbers of endemics, one might overlook the greater number of lineages in the Eastern Cape. Both species richness and phylogenetic diversity are important for conservation, including consideration of the potential for future adaptations to climate change.

Twenty-five **biodiversity hotspots**—threatened areas with very high species diversity and endemism—have been identified worldwide. These regions are identified based on plants, animals and other organisms, not plants alone. These areas together contain 133,000 named plant species, or over half those known globally. Figure 19.15 shows the locations of these biodiversity hotspots. The five most species rich of these hotspots support 20% of all plant species. Taken together, the 25 hotspots occupied 2.1 million km$^2$ in the year 2000 (1.4% of the Earth's dry land surface); originally, these diverse regions covered an area of 17.4 million km$^2$. But human development and habitat destruction have greatly diminished them. Together they had already lost 88% of their original vegetation cover by the time Myers and his colleagues published the original paper on this in 2000 (Myers et al. 2000). More recent assessments of biodiversity losses estimate that human land transformation has reduced local and regional biodiversity below critical safe limits for

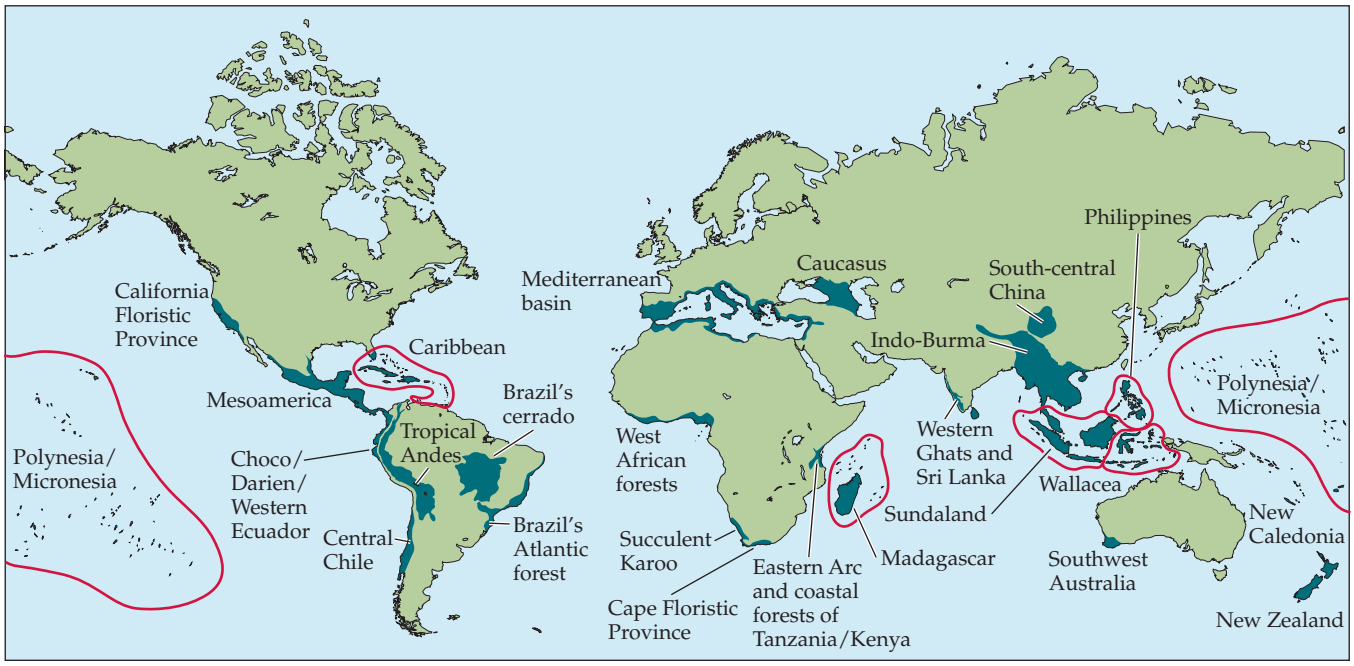



**Figure 19.15** Twenty-five regions of the world identified as biodiversity hotspots have exceptional densities of species diversity and endemism. The hotspots include both plants and animals but are defined as including at least 1500 endemic plant species. The hotspots are located within the blue regions but may make up only parts of those regions. (After N. Myers et al. 2000. *Nature* 403: 853–858.)

resilient provision of ecosystem functions for the planet as a whole (Newbold et al. 2016). We will return to the causes of biodiversity loss and to the conservation and restoration of biodiversity later in this chapter.

## 19.5 Patterns of Species Diversity May Be Explained in General Terms

How have ecologists attempted to explain the causes of variation in species diversity along geographic or environmental gradients? The processes that explain a pattern may differ with the temporal and spatial scales of that pattern (Table 19.3). At short to medium time scales—up to centuries or millennia—ecological processes dominate. Physiological properties of each species and ecological interactions such as competition and predation result in species sorting among communities within a region. For example, dry- and wet-adapted species might become restricted to uplands and wetlands, respectively. Over longer time scales, these properties and interactions, as well as the historical circumstances experienced by each species help determine their distributions at regional and continental scales. At even longer time scales, geological and evolutionary factors affect the distributions of species, as climates change, sea levels change, islands rise and subside, continents connect and separate, and lineages radiate or become extinct. Evolutionary processes can operate at shorter and longer time scales. At time scales from years to centuries, local adaptation and differentiation can create variation among populations within species. Over centuries to hundreds of thousands or millions of years, evolutionary processes of genetic isolation and speciation predominate. Some processes, including range expansions and contractions and extinctions of populations and species may occur over a long time or very quickly. Humans and their effects on the environment have greatly accelerated these processes for some species.

The mixes of processes governing diversity also differ depending on spatial scale. At local scales—within communities—the physical complexity of the habitat and interactions among species are of primary importance in determining species richness. These interactions include competition, herbivory, and mutualisms (see Chapter 10 and Chapter 11). As the scale expands to sets of communities, landscape-level processes may become increasingly important. These processes include migration among communities and extinction within communities (see Chapter 15). Finally, at continental to global scales, migration and extinction in response to climate change become important. We discuss one example, migration following the most recent glacial retreat, below and in Chapter 17.

Evolutionary processes also operate at all of these scales. At small spatial scales, local adaptation is very common (see Chapter 9). In response to competition with other species, a population may evolve to occupy a slightly different niche, thereby reducing competition and increasing local diversity (see Chapter 14). Similarly, a population may evolve a defense against an herbivore that allows it to occupy a new habitat. These processes of adaptation can lead to the formation of a new species. If the new species is different enough, it may be able to coexist in the same habitat or landscape with the species from which it was derived, increasing species richness in that landscape. At continental to global scales, as new species appear, they may migrate to new regions, increasing species richness over an entire set of regions.

Some proposed explanations for patterns of species diversity assume that species distributions are in equilibrium with current ecological conditions, while other explanations do not make that assumption. Whether a given explanation assumes that diversity is at equilibrium depends, in part, on its scale. Within a single community, for example, disturbance can increase species richness (see Chapter 13); this is a nonequilibrium explanation. But at a larger scale, across communities in a landscape or region, disturbance in particular communities balanced by migration among communities may determine species richness (see Chapter 15). This is an equilibrium explanation, although the equilibrium is dynamic.

**TABLE 19.3** A hierarchical framework for processes influencing biodiversity

| Spatial scale | Temporal scale | Ecological scale | Examples of important factors |
|---|---|---|---|
| Local | 1–100 years | Within communities | Microenvironment, biotic interactions, disturbance |
| Landscape | 100–1000 years | Among communities | Soil type, topography, elevation, metapopulation dynamics, disturbance |
| Region | 1000–10,000 years | Biome | Climatic gradients, recent glacial history, migration |
| Continent | 1–10 million years | Biogeographic province | Climatic cycles, mountain building, speciation, extinction |
| Global | 10–100 million years | Globe | Plate tectonics, sea level change |

*Source*: Modified from R. J. Whittaker et al. 2001. *J Biogeogr* 28: 453–470; K. J. Willis and R. J. Whittaker. 2002. *Science* 295: 1245–1248.

## Null models and the neutral theory of biodiversity and biogeography pose a different approach to explaining patterns of species diversity

A useful starting point for evaluating hypotheses is to build a **null model** (Harvey et al. 1983). Used throughout the sciences, null models describe the patterns that would emerge if only a specified set of processes were operating. A null model differs from a null hypothesis. Null hypotheses are used in statistical tests and are often based on the assumption that only random processes affect a system. A null hypothesis is typically expressed as there being no difference between two or more groups of observations, for instance, a control group and a group exposed to some experimental treatment. Null models generally are more elaborate. A null model asks, What would we expect to see if the process of interest were *not* operating? The null model only evaluates the effect of that one specified process.

One model, called the bounded ranges model, was proposed by Robert Colwell and George Hurtt (Colwell and Hurtt 1994; Colwell et al. 2004) and separately by Michael Willig and S. Kathleen Lyons (1998). Developed as a possible explanation for the latitudinal gradient of animal species richness, this model is based on a simple geometric argument. What if the only explanation for the latitudinal diversity gradient was that Earth is spherical, with more room in the middle and less room at the poles? Species ranges are necessarily bounded by the North and South Poles. In this model, species are placed at random and can appear anywhere, with the limitation that they are bounded at the poles (you can't go farther north or south than the poles because you are on a sphere). Random placement of species ranges on a uniform globe would result in many ranges overlapping along the equator, thus creating a peak in species richness in tropical regions and decreases in species richness toward the poles. Because the bounded ranges model includes no ecological processes—only geometry—it serves as a useful baseline against which to compare observed patterns. A null model provides a useful starting point for discovering where interesting patterns lie and what about those patterns requires explanation. No analysis of this kind has yet been done with plants. The bounded ranges model can also be used at scales other than the entire globe. It has also been applied to species richness patterns within continents and could be applied to any system that has hard boundaries at which species' ranges are truncated by an abrupt change in the environment, such as where a continent meets an ocean, or at the top of a mountain, as with the study of Ji-Hua Wang and colleagues (2018) described above. A fundamental assumption of this null model is that a species' range can be treated as a solid object that can be laid randomly on a globe or continent. This assumption is reasonable as long as the grain of the analysis is large, for example, when the prediction concerns species richness at scales of tens of square kilometers (it does not work at a scale of square meters).

An expansion of the use of null models was the **neutral theory of biodiversity and biogeography**, proposed by Stephen Hubbell (2001). The neutral theory assumes that all species are ecologically equivalent, so communities are not determined by either niche differentiation or limiting similarity (see Chapter 14). It predicts patterns of local and regional species richness and abundance using simple assumptions of rates of speciation, dispersal, and random extinction. The theory was derived from the theory of island biogeography (see Chapter 15) and models of genetic drift (see Chapter 9). Like the bounded ranges model, the neutral model of biodiversity can serve as a baseline against which to look for the presence of other ecological processes such as competition or adaptation to particular environmental conditions. The neutral theory predicts that β-diversity will increase along spatial gradients due to dispersal limitation, as species fail to reach other communities. Because the neutral model assumes that all individuals are ecologically equivalent and that niches do not exist, the model also predicts that β-diversity will not change in predictable ways along environmental gradients.

The observation that a model seems to fit a data set well does not always mean that it accurately describes the underlying processes generating the data, because many different processes may be able to generate similar patterns. For example, the neutral model accurately predicts the rank-abundance curves observed in a number of communities, but many other, non-neutral models make very similar predictions (McGill 2003; Purves and Pacala 2005). Other work suggests that the niche and neutral perspectives may be complementary, representing pieces of a more complete whole that includes both determinism (factors such as adaptation to environmental conditions, competition, and predation) and stochasticity (random dispersal, colonization, and extinction) that interact to create community structure (Vellend 2010; Chase and Myers 2011). For example, Akira Mori and colleagues (2013) examined plant species richness patterns within and among plots along an elevational gradient in northern Japan. At high elevations where productivity was low, community composition was mostly determined by competition. In contrast, at low elevations, communities varied in ways that were consistent with random processes.

## Other explanations have been posed to explain variation in biodiversity, but patterns are scale dependent

A primary cause for patterns of plant species richness at regional and global scales, including the latitudinal

diversity gradient, appears to be available energy, or productivity as a measure of energy captured in biomass (Hawkins et al. 2003). The available energy hypothesis is based on a simple idea: increasing productivity means that more energy is available for growth and reproduction, and therefore more individuals can live in the same area (Connell and Orias 1964; Wright 1983; Wright et al. 1993; see Chapter 14). Three mechanisms—random placement, local extinction, and speciation—are proposed to lead from more individuals to more species. Random placement occurs if local species richness is determined by random sampling from a regional species pool or metacommunity (Coleman 1981; Coleman et al. 1982; Hubbell 2001). As the number of individuals in a local area increases, the number of species should also increase, because the likelihood of including a rare species increases due to chance alone. Local extinction—the assumption that any local population will persist only above some minimum size (see Chapter 8)—is a component of the theory of island biogeography (see Chapter 15). At larger spatial extents, some speciation mechanisms assume a positive relationship between the number of individuals and the rate of speciation (VanderMeulen et al. 2001).

Many authors have considered explanations for how plant species richness varies along gradients of productivity, or gradients of fertility and water availability associated with productivity (Box 19B). There was a long-standing belief that species richness along productivity gradients should be low at the lowest productivities, increase to a peak at intermediate productivity, and decrease again at high productivity (Grime 1973). The explanation made sense to many ecologists: at the lowest productivity, on infertile soils or in arid conditions, fewer species are adapted to tolerate these conditions. At high productivity on fertile, mesic soils, competitive dominants take over and exclude most other species (see Chapter 10). This has been called a hump-shaped distribution, because the graph of species richness along a gradient of productivity is predicted to be low at both ends of the gradient with a "hump" (a peak) in the middle. Other ecologists have reported that they have noted a positive relationship, that is, that species diversity increases with productivity, with the highest diversity at the highest end of a productivity gradient. The problem is that the evidence has been confusing and unclear when the entire body of research on this question has been considered.

The effects of productivity on patterns of diversity have been studied extensively by many different researchers in many different systems and, critically, at many different grains and extents. In a review of many published studies, Gary Mittelbach and colleagues (2001) looked at the effects of productivity on diversity across a range of geographic extents (the total spatial area examined) and ecological extents (within vs. across communities; Figure 19.16). They found that at all spatial extents, studies reported a variety of patterns, although the most common pattern was hump-shaped (that is, higher in the middle and lower at both ends). At the largest spatial extent, hump-shaped and positive patterns were both common. Len Gillman and Shane Wright (2006) extended the analysis and concluded that positive relationships were common at regional to continental extents, while a variety of patterns were found at smaller extents. One reason for the differences in the conclusions is disagreement among researchers about which studies should be included in the analysis (Gillman and Wright 2010; Hillebrand and Cardinale 2010; Mittelbach 2010; Whittaker 2010). Other more technical issues about how studies are combined and compared may also be responsible for disparate results (e.g., Gurevitch and Mengersen

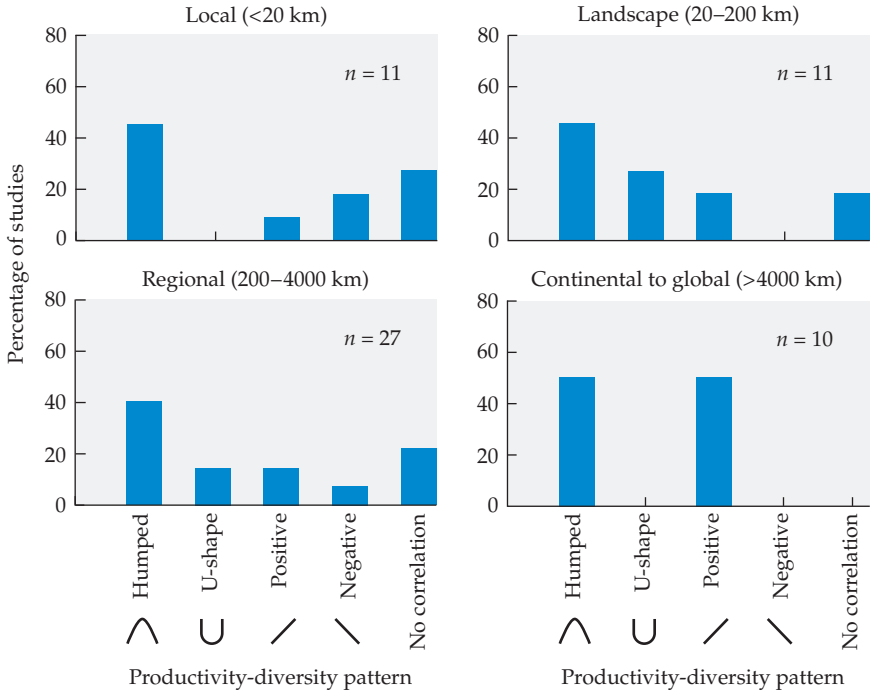



**Figure 19.16** The variation in relationships (hump-shaped, U-shaped, negative, positive) between productivity and vascular plant species richness for studies done at different geographic scales. Geographic scale is based on the farthest distance between sites in a data set. Sample sizes (*n*) refer to the number of data sets. (After G. G. Mittelbach et al. 2001. *Ecology* 82: 2381–2396.)

## BOX 19B Explaining Diversity along Ecological Gradients

A central question in ecology is explaining why there are more species in some places than others. Sometimes we can arrange those places along a line—a gradient. If that gradient is a line in space, such as the tropics to temperate regions to polar regions, or up the side of a mountain, we speak of geographic gradients. If the line is defined by variation in an environmental factor, such as disturbance, salinity, temperature, or succession, we speak of ecological gradients. There are many models (explanations) for patterns of diversity along ecological gradients. Here we describe an overarching theory of ecological gradients that combines those explanations into a broader conceptual framework.

What ecological gradients contribute to variation in species richness? Ecologists agree that gradients in energy availability are important, but beyond that there are myriad processes operating at a variety of extents. Many seemingly contradictory models have been proposed to explain the many patterns of diversity along ecological gradients (e.g., Grime 1973; Huston 1979; Tilman 1982; Rosenzweig 1992; Huston and DeAngelis 1994; Grace 1999). All of the models consider environmental gradients in species richness; the gradient can extend over very short spatial distances or be global, or it can extend over short or very long periods of time (this is the domain of these theories). Like any theory or hypothesis, all of these make assumptions. Assumptions differ among models but some of them are that (1) systems are at equilibrium at some spatial or temporal scale; (2) the species are roughly equivalent in their resource requirements, dispersal abilities, and extinction probabilities; (3) each species restricts itself more than it restricts other species; (4) local assemblages tend to be in persistent states (certain models); and (5) the regional species pool contains only species that can coexist with one another (certain models).

Samuel Scheiner and Michael Willig (2005; see also Fox et al. 2011) showed that the 24 proposed models all can be derived from the many different combinations of only four different rules:

1. There is variation in some environmental factor, which causes variation in the number of individuals, creating a gradient in the density of individuals. The environmental factor could be a resource or some condition such as stress or disturbance. If the gradient exists with respect to a single limiting resource, then each species restricts itself more than it restricts other species, thereby minimizing competitive exclusion and creating a richness gradient. Otherwise, species richness would be the same everywhere. If there are multiple resources associated with the gradient, the minimum (limiting) resource dominates (Liebig's law of the minimum, van der Ploeg et al. 1999), but the identity of that resource may be species-specific and may change along the gradient.
2. More individuals lead to more species (Fisher et al. 1943; Preston 1962a, b; Srivastava and Lawton 1998), as explained in the main text.
3. As the average value of a resource (e.g., water) or condition (e.g., soil texture) increases, its variability within a site also increases. If the resource is zero or very low, there cannot be much variation in that resource by definition. Increasing the average value allows for the possible increase in the variability, unless an upper bound also exists (e.g., water saturation of soil). Increases in environmental variation lead to increases in the number of species because species may specialize on distinct environmental conditions.
4. A trade-off in performance or species characteristics along the environmental gradient can lead to species richness being higher or lower at the ends of an environmental gradient than in the middle.

All diversity gradient models include the first two rules. Not all of these rules hold under all circumstances, nor are they mutually exclusive. Many models propose trade-offs, but they differ in specifics. For example, a commonly invoked trade-off is one between competitive ability for two different resources. Other trade-offs include competitive ability versus stress tolerance and competitive ability versus colonizing ability. The process of unifying the many disparate theories into a single conceptual framework demonstrated how seemingly contradictory ideas were closely related once they were reduced to their essentials. This effort has led to a similar process being applied to other areas of ecology (Scheiner and Willig 2011).

2010). Synthesizing the results of different studies to determine general patterns in ecology and account for factors contributing to differences in results among studies is a challenging undertaking (see Box 10A).

Peter Adler and a large group of colleagues (2011) used a different approach, an extensive manipulative field experiment, to resolving the question of the relationship between diversity and productivity. They designed a group study that used a single protocol across 48 sites and five continents in herbaceous-dominated communities. At each site researchers set up ten 1 $m^2$ plots where they counted all of the plant species present and measured peak-growing-season biomass as a measure of net primary productivity. They tested the influence of spatial scale in three ways: by examining how richness changed with productivity across individual plots within sites; by averaging results across sites within regions; and at the global extent by comparing richness and productivity across all

48 sites and five continents (Figure 19.17). The patterns of variation in species diversity as productivity changed within sites were not consistent either among sites, or with averaging across sites at the regional scale (although there were hints of a possible hump-shaped relationship among a subset of central North American grassland sites). At the global scale there was a hump-shaped distribution only when sites with a human-dominated land use history were included; when those sites were excluded, there was a positive relationship, with the highest richness at the highest productivity, but this pattern was not consistent at other scales. The authors concluded that the patterns were dependent on the type of vegetation and on spatial scale.

Even for a single geographic area and set of data, the relationship between productivity and diversity can vary with changes in the extent and grain of the analysis. This effect was demonstrated by Samuel Scheiner and Sharon Jones (2002) using data on the vegetation of Wisconsin originally collected by John T. Curtis and his students in the 1950s (Curtis 1959). At the largest extent—the state as a whole—the relationship between productivity and the species richness of terrestrial vascular plants was U-shaped, with the lowest species richness at intermediate productivities—a rather unusual pattern. At the smallest extent (7.5 × 7.5 km$^2$) the relationship between productivity and species richness was hump-shaped—that is, species richness was highest at intermediate levels of productivity (Figure 19.18A). (You will notice that this is a much larger "small" area than that of Adler and colleagues.) As the extent got larger, the hump became broader, flattening out at 120 × 120 km$^2$, and then becoming U-shaped at 240 × 240 km$^2$. When the grain of the data was manipulated by aggregation of sites, for single sites the relationship was U-shaped, at intermediate grain sizes the relationship was negative, while at the largest grain size the relationship was hump-shaped (Figure 19.18B). Michael Jennings and

John T. Curtis

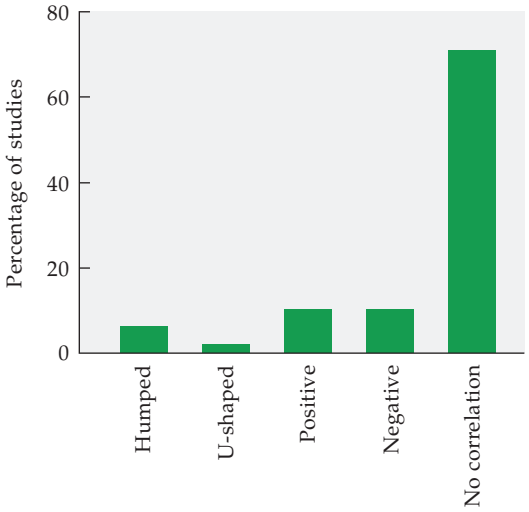



**Figure 19.17** Percentage of studies showing various relationships (hump-shaped, U-shaped, negative, positive) between productivity and vascular plant species richness for 48 herbaceous-dominated sites across five continents. (Data from P. B. Adler et al. 2011. *Science* 333: 1750–1753.)

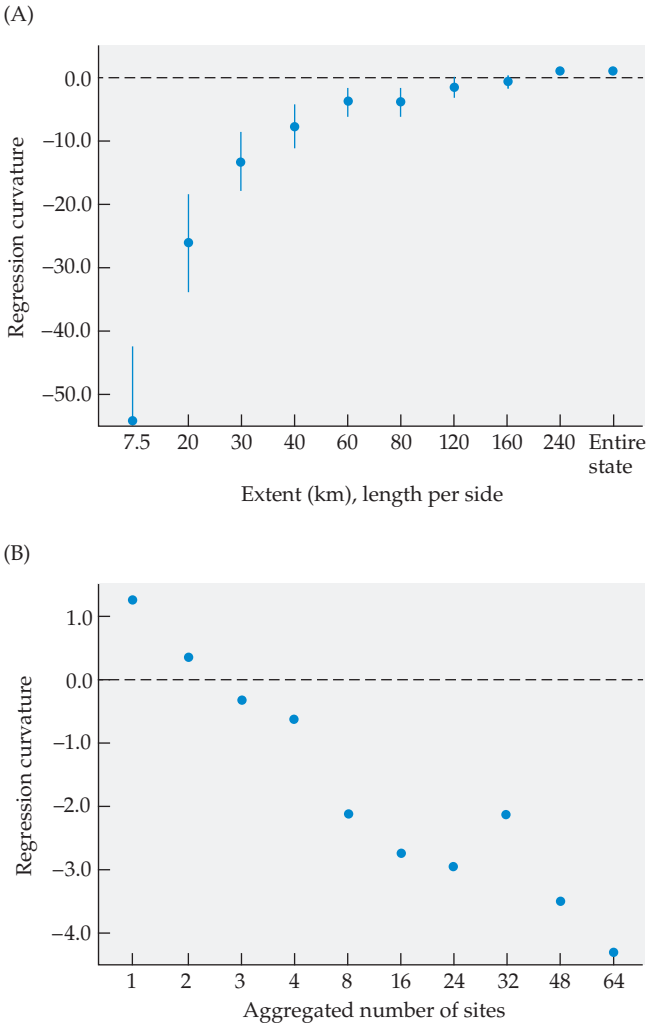



**Figure 19.18** Effects of changing (A) the extent and (B) the grain of the data on the relationship between plant species richness and productivity across the state of Wisconsin. (A) Average amount of regression curvature for 1000 randomly chosen samples at each extent; values are the length of the side of a square in km. Negative curvature indicates a hump-shaped relationship; positive curvature indicates a U-shaped relationship. Error bars show the 95% confidence interval. (B) Grain size was manipulated by combining data from adjacent sites, beginning with two sites, then three sites, and so forth. The graph shows the relationship across the entire state for each grain size. (Data from S. M. Scheiner and S. Jones. 2002. *Evol Ecol Res* 4: 1097–1117.)

associates (2005) performed a similar analysis in plant communities in the northwestern United States. They found that the overall relationship was hump-shaped and, as in Wisconsin, this relationship changed as the extent changed.

All of these analyses taken together show that no unique relationship exists between productivity and species richness. The relationship depends on the scale of the analysis, with different components of scale having different effects. One reason for the change in the relationship with scale is that the total productivity (energy captured) increases in a linear fashion with area, but species richness does not (see Figure 12.10). Another reason is that many processes can play a role in determining how diversity changes along ecological gradients (see Box 19B). More work is needed before ecologists can understand the processes responsible for species-area curves and how they relate to other factors, such as productivity.

## 19.6 Biodiversity Is Rapidly Being Lost Globally

### What is being lost?

About 80% of the plants that have become extinct in recent centuries are from the tropics and subtropics (Humphreys et al. 2019). Most are woody perennials; herbaceous plants are much less likely to be reported as extinct, although the reasons for this are unclear. **Extinction debt** exists when changing conditions doom species to eventual extinction in a community or region, but there is a lag before they actually disappear. Species facing extinction might be saved if immediate steps are taken, but planners and conservationists must recognize which species are in the greatest danger (Kuussaari et al. 2009). Long-lived plants, habitat specialists, those with small populations, those vulnerable to climate change, and those whose habitats have been degraded or are subject to fragmentation are most likely to be subject to extinction debt. For example, long-lived trees may appear to be surviving, but if they are unable to reproduce due to changed conditions, they are ultimately going to become extinct (at least in that area). Extinction can occur over a very short period of time, when a small, clustered population is wiped out due to habitat destruction or environmental change. On the other hand, specialist species, those in small habitat remnants, and those with small population sizes and experiencing changing environments may gradually decline to extinction. For a community with a large extinction debt, the number of species threatened with extinction might be underestimated, and the consequences of habitat damage and climate change to biodiversity might not be fully recognized.

Taylor Ricketts and colleagues (2005) evaluated which species globally are most threatened with extinction and the sites where those species are found. Many sites and species are in imminent danger of extinction. The situation is particularly acute in Central America, South America, and Southeast Asia and on islands worldwide. A large proportion of these sites lack any legal protection, and even when there is legal protection, it may not be enforced or enforceable. Not surprisingly, many of these areas are in biodiversity hotspots, but those are not the only areas with species in imminent danger of extinction. Tropical forests are the best-known threatened biome. About half of the species on Earth live in tropical forests, even though these forests now encompass only about 7% of terrestrial land cover. Globally, oceanic islands and archipelagos are only about 5% of the terrestrial land, but they contribute a disproportionately large number of endemic species (see Table 19.2). Thousands of island endemic plant species are threatened or in danger of extinction globally (Bramwell and Caujapé-Castells 2011).

Biodiversity losses may be subtle and require evaluation at multiple scales and use of multiple measures of biodiversity. One way to determine conservation priorities is to focus on **ecoregions**, major ecosystems defined by distinctive geography, solar radiation, and water status. An analysis by Dinerstein and colleagues (2017) found that for 88% of the world's ecoregions, less than half of the area was protected; more alarming still, 24% of the world's ecoregions are in great peril because on average only 4% of their area remains (Figure 19.19). For example, the ecoregion encompassing the grasslands in the southern part of South America, extending across Argentina, Uruguay, and the southern part of Brazil, is among the largest extent of temperate grasslands globally. They are not well protected, and about half have been converted to agricultural fields and tree plantations. They are very high in biodiversity, with estimates of 3000 native grassland plant species and a high level of endemism. In an extensive survey, Ingmar Staude and colleagues (2018) found that diversity losses in these grasslands were severe. Local extinctions were greatest in landscapes with the smallest remaining habitat, and phylogenetic diversity was most eroded in smaller habitat fragments. Although overall β-diversity was not changed by the size of habitat fragments, the two components of β-diversity—species turnover and nestedness—did change. Species turnover decreased as habitat fragments became smaller, but nestedness increased. The plant communities are becoming more homogenized, with the disturbance-tolerant, widespread, common species increasing and the narrowly distributed, rare species disappearing. The spread of anthropogenically adapted species (often non-native or invasive; see Chapter 14) and the disappearance of endemic species is the

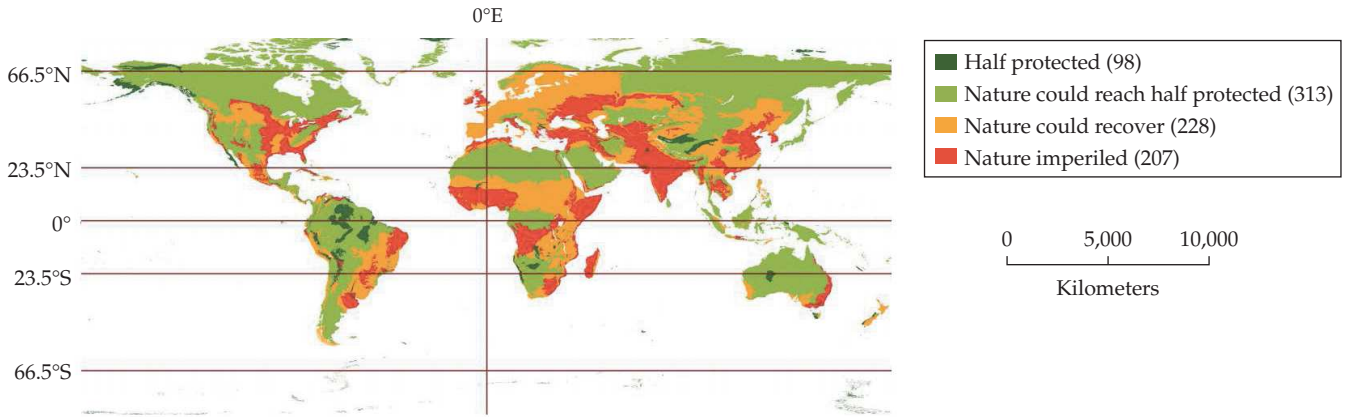



**Figure 19.19** The protection statuses of the ecoregions of the world. The ecoregion protection categories are defined as the following: Half Protected, more than 50% of the area is protected; Nature Could Reach Half Protected, less than 50% of the total ecoregion area is protected, but the sum of the total ecoregion protected and unprotected natural habitat remaining is more than 50%; Nature Could Recover, the sum of the amount of natural habitat remaining and the amount of the total ecoregion that is protected is less than 50% but more than 20%; Nature Imperiled, the sum of the amount of natural habitat remaining and the amount of the total ecoregion that is protected is less than or equal to 20%. Although enough habitat remains for nearly half of the ecoregions to exceed 50% protected, many areas lack adequate legal and conservation plans. (After E. Dinerstein et al. 2017. *Bioscience* 67: 534–545.)

likely cause of the increasing nestedness and decreasing species turnover. At regional and global scales, it means a large loss of biodiversity.

Another ecoregion with high biodiversity that is threatened with high levels of species loss is the Cerrado in South America (Figure 19.20; see Chapter 18). This region, which includes savanna and shrublands, has approximately 11,000 plant species, of which over 5000 are endemics. Since the 1950s, over half of the vegetation of the Cerrado has been destroyed and converted to soy and beef production, and this conversion has accelerated in the twenty-first century. The acceleration of extinctions in the Cerrado has received little attention globally.

We have to be careful about how we interpret the loss of plant biodiversity. Species richness has increased in many areas as new, non-native species enter regions, landscapes, and communities. This increase in species numbers has typically been greatest on islands. For example, the Hawaiian Islands and New Zealand both now have more non-native than native species. While the introduction of large numbers of non-native plant species has commonly led to a decrease in the biomass and abundance of native species, the native species have generally persisted. At local extents, the results are mixed (Vellend et al. 2017). Sites that had large disturbances, especially major landuse transitions, have lost species as more natives have disappeared than non-natives have colonized. Plant species richness is about a third lower in areas primarily given over to cultivation than relatively undisturbed areas. Other local effects depend on the specific situation. For instance, urbanization can increase species diversity due to the planting and maintaining of diverse, non-native plants in gardens and parks. Other effects of urbanization, such as increases in paved surfaces and decreases in green spaces, can lead to a loss of native plants.

## Biodiversity is threatened by human activity

Species have always become extinct, and species composition has always changed across every spatial scale, from local to global, over the history of life on Earth. Natural rates of plant extinction and colonization following climate warming after the last glaciation were documented by Bradley Butterfield and his colleagues (2019) using data from packrat middens (see Chapter 18). They

Courtesy of J. Gurevitch

**Figure 19.20** The Cerrado in Brazil is a highly threatened ecological region with high plant biodiversity.

showed that there is a persistent extinction debt: species not adapted to changing conditions can hang on for a while, but eventually extinction catches up with them. Even after the rate of warming slowed, extinction and colonization rates were higher for thousands of years.

However, there are several reasons that extinctions are different now, all having to do with humans. Human activities have vastly speeded up the decline and extinction of other species, including plants. What generally took place earlier in Earth's history over long periods of time is now occurring in "the blink of an eye" geologically. Human land and resource use has monopolized an increasing proportion of Earth's space and materials. Organisms and communities often cannot disperse or migrate to other areas because the routes are blocked or other areas are already dominated by people. We have changed fire regimes, altered river courses, drained wetlands, and made other profound changes in natural habitats. Humans have changed conditions for other species in multiple interacting ways, so now they are faced not only with individual threats, but with many threats simultaneously.

While the specific threats to biodiversity differ among biomes (Sala et al. 2000), the most important factor is changes in land use, which result in the destruction, degradation, and fragmentation of habitats worldwide (Table 19.4). Besides being entirely destroyed, habitats can be degraded by many factors, including the elimination of critical species, damage to the soil (such as by erosion or nitrogen deposition), overgrazing, removal of water or other resources for human consumption, and changes in disturbance regimes caused by humans and their domesticated animals. **Deforestation**—the complete removal of forests and conversion to other systems—is occurring in tropical forests worldwide. The rainforests of tropical Southeast Asia are particularly threatened. One of the current driving forces is the establishment of massive palm oil plantations, particularly in Indonesia. Palm oil is used in an enormous number of products—including shampoos and other beauty and cosmetic products, snack foods, pastries, and peanut butter. Zengzhong Zeng and associates (2018) used satellite data to document the loss of 239,000 km$^2$ of forests in Southeast Asia between 2000 and 2014 and their conversion to agriculture, with devastating effects for erosion, climate change, and plant and animal biodiversity (Figure 19.21; Figure 19.22).

Other forests, including highly diverse tropical deciduous forests, are also severely threatened. Climate effects sometimes interact with other factors to intensify damage. For instance, the intense 1997–1998 El Niño event caused severe and damaging fires in the rainforests of the island of Borneo, exacerbating the damage already done by extensive logging and land clearing. Complete clearing is done for lumber and to clear land for crops, while selective logging focuses on valuable (and sometimes rare) species for lumber. Even partial logging can affect many other plants and many other components of the forest. Gregory Asner and colleagues (2009) estimated that in South America, 156,000 km$^2$ (1.8%) of the remaining

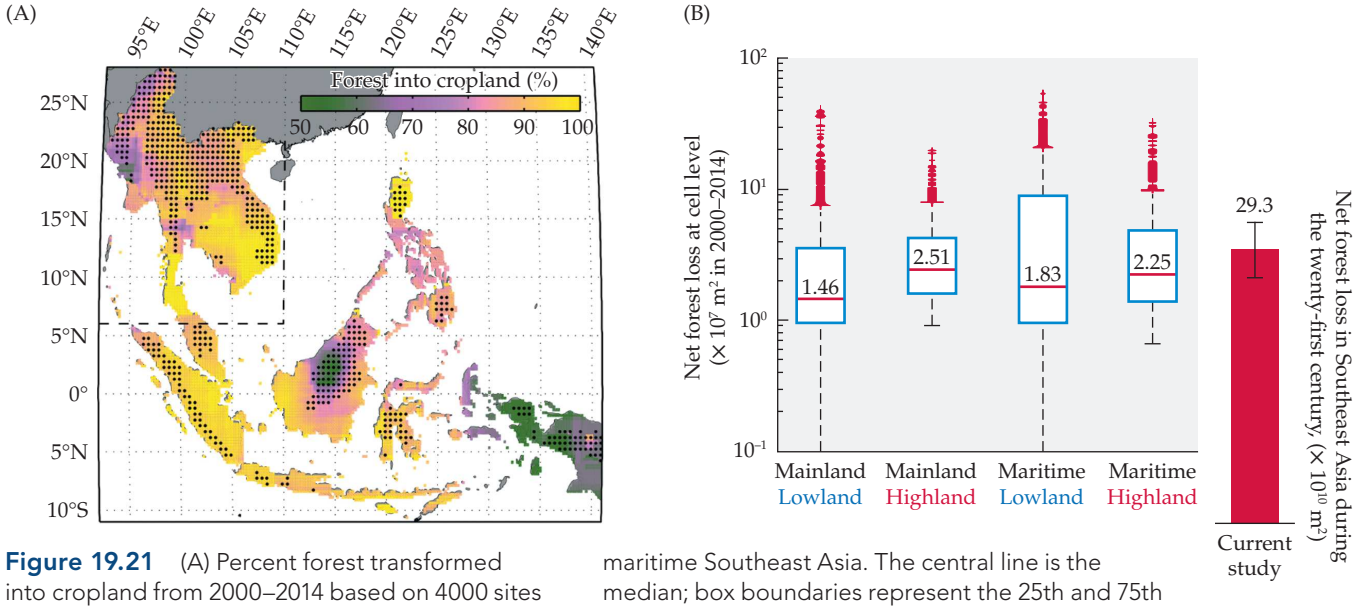



**Figure 19.21** (A) Percent forest transformed into cropland from 2000–2014 based on 4000 sites in Southeast Asia. Stippled areas are highlands. Dashed line separates mainland from maritime areas. (B) Net forest loss at level of 0.25° cells (× $10^7$ m$^2$) in the lowlands and highlands of mainland and maritime Southeast Asia. The central line is the median; box boundaries represent the 25th and 75th percentiles; whiskers represent 99.3% coverage of the data; the red bar shows total magnitude of net forest loss × $10^{10}$ m$^2$ during 2001–2014. (After Z. Zeng et al. 2018. *Nat Geosci* 11: 556–562.)

tropical rainforest were totally cleared from 2000 to 2005, and another 1,600,000 km² (over 18%) were partially logged. In Africa, 15,000 km² (0.5%) were cleared during this time period, and 560,000 km² (19.2%) were partially logged. Globally, 275,000 km² were deforested and 4,000,000 km² were partially logged.

A major factor exacerbating forest loss is road building, which is accelerating greatly. The Chinese Belt and Road project, begun in 2013, is the most ambitious infrastructure development project in history, involving extensive road building in 152 countries in Asia, Europe, Africa, the Middle East, and the Americas. Considerable concern about the enormous effects on biodiversity and ecosystem sustainability have been raised about the plans for this project. Deforestation is far greater along roads. Rivers can also act as conduits for human activities. In the Amazon, over 90% of deforestation has been carried out within 5.5 km of roads or within 1 km of rivers (Barber et al. 2014).

Biodiversity losses due to habitat destruction are often combined with and exacerbated by other threats. In many parts of the world, grazing by cattle, sheep, and goats has severely degraded grassland ecosystems. In many grasslands, climate change, increases in atmospheric $CO_2$, grazing, alteration of fire regimes, and other factors have led to woody plant encroachment, particularly of shrubs.

**TABLE 19.4** Current and likely future factors causing direct and indirect changes to global, regional, and local biodiversity

| |
|---|
| Land and land use transformation, including soil erosion |
| Harvesting and overexploitation of natural populations |
| Freshwater appropriation for human use |
| Climate change |
| Species invasions |
| Direct effects of elevated and artificial atmospheric and marine $CO_2$, tropospheric ozone, chlorofluorocarbons and other chemicals that destroy stratospheric ozone, $SO_2$, and other acid precipitation components |
| Pollution from fertilizers, pesticides, plastic, etc. |
| Accumulation of trash and discarded materials in landfills, oceans, etc. |

*Note*: There are overlaps between categories; for example, landfills generate methane, which contributes to climate change.

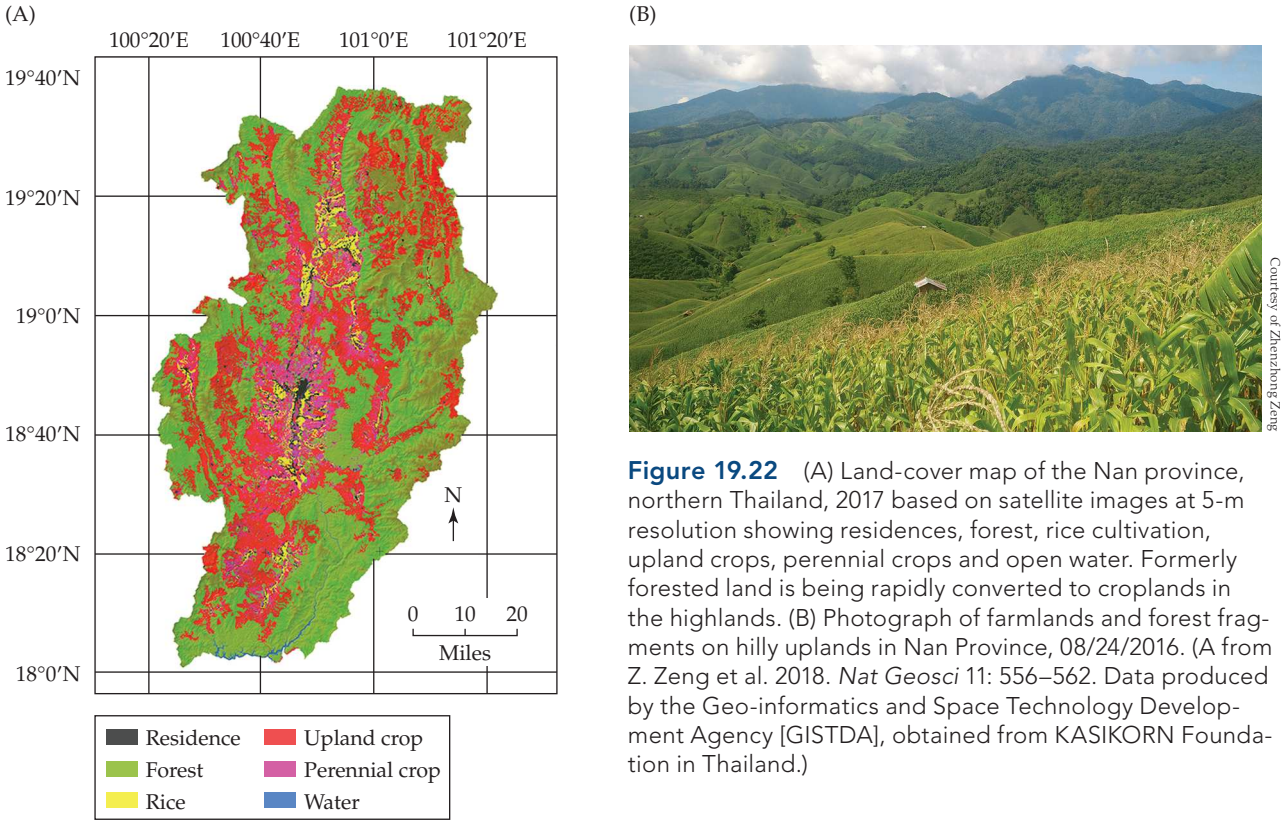



**Figure 19.22** (A) Land-cover map of the Nan province, northern Thailand, 2017 based on satellite images at 5-m resolution showing residences, forest, rice cultivation, upland crops, perennial crops and open water. Formerly forested land is being rapidly converted to croplands in the highlands. (B) Photograph of farmlands and forest fragments on hilly uplands in Nan Province, 08/24/2016. (A from Z. Zeng et al. 2018. *Nat Geosci* 11: 556–562. Data produced by the Geo-informatics and Space Technology Development Agency [GISTDA], obtained from KASIKORN Foundation in Thailand.)

### Does human domination require a new definition of the biomes?

Over three-quarters of terrestrial environments have been altered, sometimes totally transformed, by humans. Erle Ellis and Navin Ramankutty (2008) suggested a redefinition of Earth's biomes. Instead of basing their classification on potential vegetation and climate (see Chapter 17), they based it on an analysis of their actual current status as they have been restructured by humans. Using data on human populations (urban or nonurban), land use (percent area of pasture, crops, irrigation, rice, urban land), and land cover (percent area of trees and bare earth), they identified 18 anthropogenic biomes. By recognizing human transformation of the environment, this classification offers valuable insight into ecological, biogeographic, and social factors determining biodiversity. However, the classification can vary with the scale at which the data are analyzed (Alessa and Chapin 2008). It may be that these classifications are more accurate descriptions of something other than biomes, which are large, relatively consistent areas. Do urban metropolitan areas and megacities constitute a new biome? Perhaps so. A good case could be made, as they share many characteristics in common, whether they are in Asia, Africa, North America, or elsewhere, and they dominate substantial regions. Other authors have delimited larger human-dominated systems as "social-ecological regions" (Jones et al. 2019). Nested within the large biome-like social-ecological regions (e.g., the American West) there are more specific social-ecological systems (e.g., Seattle-Tacoma metropolitan region in Washington; the (recently designated but possibly former) Grand Staircase-Escalante National Monument in Utah), similar to the hierarchical classification of plant communities (see Table 12.1). Other classification systems are possible, such as by ecoregions (see Figure 19.19). In any case, human domination of the biosphere is reshaping our thinking of the natural world.

### Both rare and common species face threats in a range of communities

Not only high-diversity communities and rare species are threatened by human actions. Less diverse communities, from cold deserts to marshes to pine barrens, in many areas of the world are also subject to damage from a wide range of direct and indirect anthropogenic effects. These lower-diversity communities or regions may be valued by people and be of high ecological importance, even though they have lower species numbers. They may harbor phylogenetically divergent endemic species, provide ecosystem services like water purification and storm protection (coastal marshes), protect water supplies through aquifer regeneration (pine barrens), or be valued for their beauty and potential to inspire awe and emotional and spiritual regeneration.

Even originally very common species and ecosystems may be endangered. In North America *Castanea dentata* (American chestnut, Fagaceae; see Figure 11.24) was decimated by an invasive pathogen transported by humans in the early to mid-twentieth century (see Chapter 11). More recently, in Southeast Asia various species of *Tectona* (teak, Verbenaceae) have become endangered due to logging for their valuable wood and by habitat destruction. Once-widespread communities from pine barrens to many kinds of wetlands have been greatly reduced in extent or have disappeared altogether. Climate change and human land uses have created and dramatically increased the extent of newly created, depauperate deserts at the expense of fragile grasslands and shrublands. In the Sahel region, south of the Sahara, climate change and overgrazing have converted vast areas of once productive grasslands and shrublands into deserts of shifting sands. Similarly, in Lagan in central China near the headwaters of the Yellow River, extensive grasslands once supported people and their grazing animals. They are now barren wastelands supporting little vegetation, with recurrent sandstorms and vanishing rivers, lakes, and pastures, apparently due to a combination of intensive land use (particularly overgrazing) and climate change.

Overexploitation by collection and removal has reduced populations of many plant species. Carlos Peres and coworkers (2003) showed that populations of *Bertholletia excelsa* (Brazil nuts, Lecythidaceae)—the most important food crop collected entirely from wild trees—are declining in areas of the Amazon rainforest where the nuts are harvested. As a result of the harvesting, populations in these areas tend to have few or no young trees, and if the harvesting persists at present levels, they are threatened with extinction. Several plant taxa, particularly orchids and cacti, are threatened by amateur collectors (Orlean 2000). While collecting and selling plants taken from natural populations is illegal in many areas, steep prices and minimal fines conspire to maximize these depredations.

Many plant species are threatened indirectly by overexploitation of other species. The killing of major predators, for example, can lead to rises in herbivore populations, resulting in overgrazing or overbrowsing. Extensive losses of understory vegetation in forests all over the northeastern United States in recent years are due, in part, to a tremendous rise in populations of white-tailed deer. Increased deer herbivory is implicated as a major contributor to the decline of *Trillium grandiflorum* (great white trillium, Melanthiaceae; see Figure 8.4) populations in Pennsylvania (Knight 2003, 2004).

### Human population growth and land use contribute to biodiversity loss

The number of people in the world has grown dramatically since the Industrial Revolution, and it is continuing

to grow rapidly. The consequences of these global increases in the human population have been a cause for alarm and controversy for many years—as exemplified by the wide range of reactions to the much-disputed 1968 book by Paul Ehrlich, *The Population Bomb*. Some of the predictions made about the consequences of human population growth were flawed and alarmist, but attempts at outright dismissal of these concerns have also been flawed. Both ecologists and social scientists have increasingly focused on the complex issues involving environmental effects of ever increasing human populations. Work on documenting the "global human footprint" has highlighted the geographic areas that have been totally or partially transformed by human populations and activities (Sanderson et al. 2002, Venter et al. 2016). Elements contributing to the human footprint include the built environment, agriculture, population density, transportation, and night-time lighting. The most heavily impacted areas are eastern and portions of midwestern North America, much of western Europe, eastern Brazil, most of India and large parts of China and southeast Asia. Parts that are relatively less impacted include parts of the Amazon Basin, the far northern parts of Asia and North America, the largest deserts, the Tibetan plateau and surrounding Himalayas, and smaller scattered areas that are protected or inaccessible.

Recent growth in human populations has been much greater in developing than industrialized regions and tends to be greatest in the tropics, where a large proportion of the people are living at or below subsistence levels. Population growth has slowed in many developed, industrialized countries. Rapid human population growth can be associated with large environmental consequences—such as increased resource use, habitat destruction, and species extinctions. But how much of the deforestation and other damage to natural communities can we ascribe to population growth itself? In many cases, the resource extraction is either carried out by or used to benefit people in industrialized countries far from the source of the resources and the areas where population growth is highest. Human societies are complex systems, and it is not surprising that human populations affect the environment in indirect and multifactorial ways. Both local and international economic and political forces interact with population size to influence changes in land use (Lambin et al. 2001).

Social scientists who seek to characterize and quantify the interactions between humans and their environments disagree about the extent to which different factors, including population size, are responsible for environmental problems such as deforestation. On the one hand, there is solid evidence that in many developing countries, population growth results in greater deforestation (Ehrhardt-Martinez 1998). The negative relationship between human population growth and forest area is complex and changing, however (Mather and Needle 2000). In studies focused at the scale of particular countries or regions, there have been contradictory results. In two studies of the effects of human population density on deforestation in the Amazon basin of Brazil, its effects were at best indirect. Alexander Pfaff (1999) found that in the Brazilian Amazon, population density did not directly determine tropical forest clearing. Rather, many factors were implicated, from land characteristics and distance to markets, to government development projects. However, Pfaff also found that while a static measure of population size was not a good predictor of deforestation, the first settlers in a region had a far greater effect on forest destruction than later immigrants. He concluded that the timing and distribution of population growth are more important than its absolute magnitude in determining its effects. Govindan Parayil and Florence Tong (1998) found that population size and growth were not directly responsible for deforestation in the Amazon basin. Instead, multiple agents were involved, operating at different scales. Major agents responsible for deforestation included large-scale clearing of forests to create cattle ranches and logging of forests for lumber (both generally carried out by large multinational corporations). A third major agent of forest destruction was the very small scale clearing of forests for farming (slash-and-burn agriculture) by long-term residents as well as by impoverished immigrants to the region from other parts of the country. Thus, Parayil and Tong found that factors other than population growth were the major direct contributors to plant community destruction at a regional scale.

William Laurance showed in detail how patterns of tropical forest loss and fragmentation are also changing (Laurance 2000; Laurance et al. 2001, 2018; see Chapter 15). In past decades, many of the largest-scale forest losses in the Amazon basin occurred at its southern and eastern edges. The development of major new highways has led to great increases in forest loss. Katrina Brown and François Ekoko (2001) found the opening of new roads in forested areas to be the most important factor in deforestation in the rainforest of Cameroon, in central Africa. These researchers showed that multiple agents (from the government to corporate loggers to local village interests and needs) and causes of deforestation were acting together to the detriment of natural systems. China is now a leading force for development in tropical areas, resulting in vastly increased road networks that will undoubtedly contribute to the erosion of tropical forests and biodiversity, although some efforts are being made to counter these effects.

Courtesy of Professor William F. Laurance, James Cook University

William Laurance

Patterns that are broadly similar in their complexity, but different in important details, have been found in human-environment interactions in the United States. In the past half century, two distinct patterns of changing residential land use have had large effects on the reduction and fragmentation of natural communities. The rise of suburbanization—the establishment of vast low-density residential areas surrounding metropolitan centers, particularly in North America—has expanded human affects over thousands of square kilometers. This shift has changed land that was once agricultural or occupied by natural communities into areas dominated by widely spaced homes on large plots of land, with commercial development in other areas. Another major land use change has been the rise of the urban-wildland interface: the incursion of residential, vacation, and retirement homes into wilderness, particularly in coastal, montane, and desert areas (Bartlett et al. 2000). While this use does not consume as much land as suburban development, it can cause fragmentation of natural areas, changes in fire regimes, erosion, and incursions of invasive plant and pest species.

Local planning and policy decisions that create suburban development patterns can also lead to other unintended consequences, such as increases in per capita fossil fuel consumption by commuters. In suburban areas, home heating and cooling use much greater amounts of energy—and produce much more greenhouse gases—on a per-home-unit basis than in dense, urban housing. Ironically, urban living can be greener in some ways than exurban or suburban areas.

Human domination of land use poses another problem. If species are forced to migrate in response to climate change, this may be prevented by large urban areas. Kathryn Miller and Brian McGill (2018) found that along the east coast of the U.S., predicted dispersal of tree species in response to climate change would be slowed by up to 40%, and the trees with the greatest predicted shifts were the ones most affected by urbanization barriers.

## 19.7 Ecosystem Services Are One Way of Quantifying the Benefits of Natural Systems to Humans

The natural world provides many different kinds of benefits to people's quality of life, livelihoods, and even survival. In an effort to define concrete benefits of nature to humans (and sometimes estimate their monetary worth), the concept of ecosystem services was introduced in 1970 and became popularized by the turn of the twenty-first century (Daily et al. 2000; Brown et al. 2007). An **ecosystem service** is a direct or indirect benefit that ecosystems provide to people. Four categories of ecosystem services were defined in the Millennium Ecosystem Assessment (2005)—supporting, provisioning, cultural, and regulating. Examples of **supporting** services include pollination (crops, forest species) and nutrient cycling. **Provisioning** includes food, lumber, medicines, and genetic resources. **Regulating** includes flood control by marsh plants and mangrove ecosystems, decomposition by microorganisms, and regulation of rainfall and the atmosphere by forests. **Cultural** services include flowers (weddings, funerals), spiritual values (awe), and psychological refreshment and regeneration in nature. In a paper synthesizing the research on ecosystem services, Bradley Cardinale and collaborators (2012) concluded that biodiversity is strongly related to some ecosystem services, particularly provisioning and regulating services. For example, crop yields are strongly and positively related to genetic diversity of the crops, and plant species richness enhances wood production, carbon sequestration, and soil nutrients and organic matter.

### Why should anyone care about plant biodiversity?

While there is no question that humans benefit from natural systems and depend on them, the question of whether biodiversity itself is beneficial to humans has been much more controversial.

Plants are the foundation of almost all terrestrial food webs. They are the habitat for other organisms. They are essential components of the water, carbon, and nitrogen cycles regionally and globally (see Chapter 5). But does it matter if biodiversity is diminished? There is considerable evidence that the functional diversity and species richness of a community can exert strong influences on ecosystem properties and services, including productivity and its variability, at least under certain conditions (see Chapter 14). Species extinctions can cause the degradation of ecosystem services, although not all species have large effects on ecosystem properties or ecosystem services (Hooper et al. 2005). Ecosystem services may depend on complex interactions among species and functional groups in ecological systems. These complex interactions may, therefore, both directly and indirectly affect biodiversity (Harrison et al. 2014). Species richness and ecological complexity may buffer ecosystems against climate change.

Should we be concerned about loss of plant (or other) biodiversity for practical, aesthetic, financial, spiritual, economic, or other reasons? Is human well-being the only priority, and is preservation of the natural world a luxury that is really in conflict with human well-being? Should jobs and human priorities be "balanced" with preservation of the natural world, or is this argument just an excuse for ignoring the protection of nature? It

has been argued that the most vulnerable people suffer the most from ecological degradation and enironmental disasters. Ultimately these are not scientific questions; they are questions people need to answer for themselves, based on their own values and ideally after becoming aware of the issues and evidence.

There are two fundamental arguments for preservation of natural areas, species, and communities. The first argument is that these organisms and natural systems provide critical benefits to people, and their loss has costs to human well-being. Documenting and quantifying these ecosystem services can help make a case for the value of preserving the biodiversity that is their basis. However, the benefits and costs of such preservation are not shared equally by all people. If intact mangrove ecosystems prevent the loss of homes and farms to coastal storms and support fish populations for villages dependent on fishing, that is a local or regional benefit. If destroying them to build a power plant benefits the investors in the power plant and the urban dwellers who get cheap electricity, the villagers suffer but others may benefit, particularly over the short term.

The second reason is more abstract and is an issue of values rather than something scientific, or a thing that can be monetized. Is everything to be valued only because of its benefits to us? We do not value our friends only because they might provide us with job contacts, although they might do so. We do not care for family members only because doing so benefits us financially, although they often provide such benefits. So why should we value natural systems only because of what they do for us personally? Do other organisms have a right to exist just because they exist? What if an ecosystem service only benefits someone else? Should we consider it valuable if we gain no benefit from it ourselves? These are things you'll have to ponder and decide for yourself.

## Conservation and restoration of biodiversity: a ray of hope?

Humans live in and depend on the natural world in myriad ways. We are part of the natural world, although we play a unique role and can be enormously destructive. As a species with intelligence, strength, memory, foresight, and the ability to work together, we are also capable of protecting and restoring natural systems.

Is it all gloom and doom? Is the world going to hell in a handbasket? There are clearly reasons to think so. However, there are also people working mightily to change the statement "All of this will happen if the current rates of [fill in your favorite bad thing] remain unchanged." Perhaps some of these things can be changed. Human behavior is not immutable. If current practices are altered, the consequences of our actions can also be changed. Major international and national governmental and private efforts have been undertaken in recent years to quantify, address, and remediate the environmental and biodiversity issues outlined here; particularly notable among them are the Intergovernmental Panel on Climate Change and the Millennium Ecosystem Assessment, both under the banner of the United Nations, and the International Union for Conservation of Nature (IUCN), publisher of the Red List of Threatened Species.

Restoration ecologists work to restore biodiversity in forests, wetlands, and prairie remnants. They re-create seminatural reserves in the middle of urban landscapes, islands of habitat surrounded by a contrasting urban matrix. The best practices for restoration are still being determined. For example, we do not know how many species to start with. If invasive species are removed, they might return or just be replaced by other invasive species. What will ensure that mutualists like pollinators and mycorrhizal symbionts arrive? Natural communities might not survive human overuse and non-native plant intrusion (Handel 2017). These are very challenging issues, but there have been some notable successes. Local efforts at ecological restoration can yield remarkable results, sometimes in the most unlikely places. In New Delhi, India, one of the densest, most polluted cities in the world, what was once a wreckage of garbage and polluted water is now a restored wetland refuge named Yamuna Biodiversity Park. It supports a high diversity of wetland and forest plants and animals, particularly birds, and is a welcome oasis for park visitors, including youth groups. On the other side of the world in the U.S., high diversity prairies once encompassed vast areas but are now a tiny remnant of their former extent. Experiments with promising results are showing how to best restore and maintain prairie biodiversity, including in formerly cultivated areas (Baer et al. 2020). In Kansas, Konza Prairie Biological Station, a Long Term Ecological Research Network site (see Figure 1.2 and Box 12D), is a 3500 ha restored and protected native tallgrass prairie supporting high plant and animal biodiversity. Research there has made substantial progress in our understanding of the roles that fire, climate, topography, and grazing by bison and cattle play in biodiversity and ecosystem function in these grasslands.

People are working in many countries to reduce carbon emissions, stop deforestation and habitat fragmentation and deterioration, and restore damaged or destroyed ecosystems. We have the option of supporting efforts to protect species and habitats locally and around the world, and of changing our own habits and practices to reduce our personal impact on the environment. We also have the option of acting to alter current patterns of environmental destruction. The consequences of doing nothing should be obvious.

## Summary

- One of the best-known biodiversity patterns is the latitudinal diversity gradient, with greatest species richness in the tropics, intermediate richness in temperate zones, and lowest richness in polar regions.
- Some places have very high species richness, and some areas, such as islands, have unusually high numbers of endemic species,.
- There are many hypotheses for large-scale patterns of plant species richness.
- The relationship between species richness as a function of productivity varies as a function of grain and extent.
- No one process accounts for all of the observed patterns in biodiversity, with different processes important in different circumstances and at different scales.
- Differences in species composition among sites, β-diversity, vary globally and can be greater in transitions between botanical regions.
- Ecological and evolutionary histories of particular regions and interactions between local and regional processes are important in affecting biodiversity.
- Regional and local diversities can influence each other.
- Habitat destruction, global climate change, and other factors have increased the number of species threatened with extinction.
- Human populations have complex effects on and interactions with biodiversity loss.
- Protecting plant biodiversity can maintain ecosystem services.
- Conservation and restoration efforts can help sustain biodiversity.

# Glossary

Numbers in brackets refer to the chapter(s) where the term is introduced.

## A

**abiotic** Nonliving. [1, 15]

**absolute competition index (ACI)** A measure of the intensity of competition. [10]

**abundance curve** A graph describing a hierarchy of relative abundance among the species in a community. [14]

**acclimation** A potentially reversible adjustment to environmental conditions. [2, 3]

***A-C*$_i$ curve** The relationship between $CO_2$ availability and photosynthetic rate. [3]

**acid precipitation** Precipitation that is more acidic than normal. [5]

**activity limit** The range of temperatures at which a plant can function effectively. [3]

**actual evapotranspiration (AET)** The amount of water that enters a system in precipitation, minus the amount that is lost in runoff and in percolation to groundwater. [5]

**adaptive trait** A trait that increases the probability that an organism will survive, reproduce, or leave descendants. [9]

**additive design** An experimental design for testing the effects of competition by altering the total density of neighbors while keeping the density of the target species constant. [10]

**additive series** An experimental design for testing the effects of competition by altering both the total density and the frequency of neighbors. [10]

**adiabatic lapse rate** The rate of cooling of rising air. [16]

**adsorb** To attract and hold to a surface. [4]

**adventitious roots** Roots that originate from aboveground tissue. [6]

**aerenchyma** Aerated plant tissues. [3]

**afforestation** The planting of trees to expand forested land. [16]

**agamospermy** The production of seeds without fertilization or meiosis. [6]

**age-based** Of or referring to methods for the study of populations that require only information on the population's age structure. [8]

**age-structured** Of or referring to a population described by the relative frequencies of individuals by age class. [8]

**alkaloid** A small compound that contains an aromatic ring and nitrogen. [11]

**allelochemical** A chemical released by one organism that inhibits the growth of another organism. [10]

**allelopathy** The chemical inhibition of the growth of one organism by another. [10]

**allopatric speciation** Speciation that occurs when a geographically isolated population becomes a new species. [9]

**allopolyploidy** Polyploidy resulting from the union of gametes from individuals of different species. [9]

**alpha (α) diversity** Inventory diversity within a single community. [12]

**alternation of generations** A reproductive cycle in which a haploid multicellular organism gives rise sexually to a diploid organism, while the meiotic products of the diploid organism grow directly by meiosis into a haploid organism. [6]

**ammonification** The release of nitrogen from organic compounds in the form of ammonium ions ($NH_4^+$). [5]

**amphistomatous** Having stomata on both sides of the leaf. [3]

**androdioecy** A system of gender expression in which some plants in a population have only staminate flowers, while others have either perfect flowers or a mixture of both staminate and pistillate flowers. [6]

**andromonoecy** A system of gender expression in which both staminate and perfect flowers occur on the same individuals. [6]

**aneuploidy** Changes in the genome of a species due to the gain and loss of individual chromosomes. [9]

**angle of incidence** The angle that a ray of sunlight makes with a line perpendicular to the surface of the Earth. [16]

**annual** A plant in which the life cycle is completed in one year or less. [7]

**anoxic** Without oxygen; anaerobic. [17]

**anther** The terminal portion of a stamen containing pollen in pollen sacs. [6]

**anthropogenic erosion** Accelerated soil erosion due directly or indirectly to human activities; sometimes just called erosion. [4]

**aphelion** Earth's farthest point from the sun. [16]

**apical dominance** The prevention of growth of other meristems by the apical meristem. [6]

**apical meristem** A group of undifferentiated plant cells at the tip of a growing stem or branch. [6]

**apomixis** Reproduction by maturing seeds in the absence of fertilization. [6]

**apparent competition** Density-dependent negative interactions between species that appear to be due to competition for resources but are actually due to a shared predator or herbivore. [11]

**aquatic plant** A plant that grows submerged in water. [3]

**arbuscular mycorrhizae (AM)** A group of fungi that grow inside the cortex cells and intercellular spaces of roots. [11]

**arbuscule** The delicately branched structure of an arbuscular mycorrhizal fungus inside the cortical cells of the roots and in the intercellular spaces between root cells of a plant. [11]

**archaeophytes** Ancient introductions of exotic plant species. [14]

**archipelago** A string of islands. [15]

**architecture** The arrangement of regenerating parts of a plant. [6]

**assemblage** A group of phylogenetically related organisms living in the same place. [12]

**assimilation rate** The rate at which $CO_2$ is taken up by a leaf. [2]

**association** A particular community type, found in many places and with a certain physiognomy and species composition. [12]

**assortative mating** The mating of individuals with similar phenotypes at a greater frequency than would occur by chance alone. [6]

**asymmetric competition** Competition in which one individual has very large negative effects on its neighbors, but not the reverse. [10]

**autecology** The ecological study of individuals. [1]

**autopolyploidy** Polyploidy resulting from the union of gametes from individuals of the same species. [9]

**axillary meristem** A group of undifferentiated plant cells at a leaf or stem axil. [6]

## B

**basal area** The area occupied by the base of a plant such as a tussock grass or a tree with a definable base. *See also* basal cover. [12]

**basal cover** The fraction of area occupied by the base of a plant such as a tussock grass or a tree with a definable base. *See also* basal area. [12]

**beneficial mineral element** Minerals that either are essential for only some plant species, or are not essential but stimulate growth. [4]

**beta (β) diversity** Differences in community composition along an environmental gradient or among communities in a landscape. [12]

**biennial** A semelparous plant that flowers after 2 or more years. [7]

**biodiversity** All of the populations, species, and communities in a defined area. [12]

**biodiversity hotspot** A threatened area with very high species diversity. [19]

**biogeochemical cycle** The major fluxes and pools of a substance in a system. [5]

**biological control** The deliberate use of one species to control another undesirable species. [11]

**biological species concept** The concept that a species is a group of actually or potentially interbreeding organisms that are completely or nearly completely genetically isolated from other such groups such that each group is an evolutionarily independent lineage. [9]

**biomass** Mass of living organisms. [12]

**biome** A major biogeographic region that differs from other such regions in the structure of its vegetation and its dominant plant species. [18]

**biosphere** The living organisms on Earth and their physical environment. [5]

**biotic** Living. [1, 15]

**bottom-up control** Limit of herbivore populations by their food supply. [11]

**boundary layer** A blanket of relatively still fluid that surrounds all objects in air or water. [3]

**broad-sense heritability** Heritability that is due to total genetic variation. [9]

**browser** An animal that eats leaves from trees or shrubs. [11]

**bulb** An underground rosette stem that stores nutrients. [6]

**bulbil** A tiny bulblike organ vegetatively produced in an inflorescence or leaf axil. [6]

**buzz pollination** A type of pollination in which pollen can be effectively extracted from flowers only by vibrating the anthers at an appropriate frequency. [6]

## C

**$C_3$ photosynthesis** A form of photosynthesis in which $CO_2$ is captured by RuBP carboxylase/oxygenase and the first stable product is a three-carbon compound. [2]

**$C_4$ photosynthesis** A form of photosynthesis in which $CO_2$ is captured by PEP carboxylase and the first stable product is a four-carbon compound. [2]

**calcic horizon** A cement-hard layer of soil, formed by a concentration of calcium, that is found in many deserts. [4]

**Calvin-Benson cycle** The set of biochemical reactions by which carbon from $CO_2$ is incorporated into organic compounds. [2]

**CAM photosynthesis** *See* crassulacean acid metabolism. [2]

**cambium** A zone of meristematic cells between primary xylem and phloem. [6]

**canonical correspondence analysis (CCA)** An ordination method based on the correlation of a matrix of community similarities with a matrix of environmental relationships in which the axes are restricted to be linear combinations of explanatory variables. [12]

**canopy cover** The amount of ground covered by the highest level of vegetation of a community. [12]

**capillary action** The pulling of water through a narrow tube or a small space by the attraction of the water molecules to the surfaces of charged particles or to one another. [4]

**carbon reduction** The part of photosynthesis in which carbon is taken up from the atmosphere. [2]

**carboxylation** A chemical reaction in which a $CO_2$ molecule is added to another molecule. [2]

**cation** A positively charged ion. [4]

**cation exchange capacity (CEC)** A measure of the total ability of the colloids in a soil to adsorb cations. [4]

**cavitation** Breakage of the water column in xylem capillaries. [3]

**cellular respiration** A process by which organic compounds are broken down to release energy through the use of oxygen and the release of $CO_2$. [2]

**chamaephytes** Shrubs with buds less than 25 cm above the ground. [12]

**characteristic dimension** The effective width of a leaf with respect to fluxes of energy and mass. [3]

**characteristic value** An eigenvalue. [8]

**characteristic vector** An eigenvector. [8]

**chasmogamous** Of or referring to flowers that are open and potentially cross-pollinated. [6]

**chronic herbivory** Herbivory that occurs over long time periods. [11]

**chronosequence** A set of communities of different ages since disturbance, assumed to represent a single community over time. [13]

**cladogram** A phylogenetic tree. [12]

**clayey soil** Soil with at least 35% clay particles. [4]

**clear-cutting** The removal of all trees. [13]

**cleistogamous** Of or referring to flowers that are closed and self-pollinated. [6]

**climate** The long-term statistical distribution of the weather in a given area. [16]

**climax** The hypothetical equilibrium end point of a successional sequence. [13]

**cline** A gradient in allele frequencies or other population characteristics. [9]

**clonal** Consisting of potentially independent ramets. [6, 9]

**clonal fragmentation** A form of reproduction in which pieces of a plant are capable of breaking off and rooting to form new independent plants. [6]

**clone** A genetic duplicate. [6]

**cohort** A group of individuals in a population that germinated, reached a particular size, or entered a study at about the same time. [8]

**cold air drainage** Sinking of cold air masses to low-lying spots due to the greater density of colder air. [3]

**collenchyma** Support tissue made of living elongated cells with irregularly thickened walls. [11]

**colloid** A substance in which one or more materials in a finely divided state are suspended or dispersed throughout a second material. [4]

**commensal** Of or referring to a species interaction in which one species benefits and the other neither benefits nor is harmed. [4, 11]

**common garden** A garden in which plants from diverse locations are grown together to determine whether differences among populations are due to genetic differences. [8]

**common mycorrhizal network (CMN)** A network of extensive mycorrhizal connections that link plants of many different species. [10]

**community** A group of populations that coexist in space and time and interact with one another directly or indirectly. [12]

**community type** A group of communities that share the same dominant species. [12]

**compensation** A hypothesized response to herbivory resulting in no net difference between grazed and ungrazed individuals. [11]

**competition** A reduction in fitness due to shared use of a resource in limited supply. [10]

**competitive effect** The ability of an individual to affect neighbors through resource use. [10]

**competitive response** The response of neighbors to resource use by an individual. [10]

**compositional similarity** The extent to which adjacent patches share a similar set of species at similar frequencies. [15]

**conduction** The direct transfer of heat energy from the molecules of a warmer object to the molecules of a cooler object. [3]

**connectivity** The degree to which a landscape facilitates or impedes the movement of organisms. [15]

**constant final yield** A consistent total final dry weight (or "yield") in a given area for plants sown over a wide range of densities. [10]

**constitutive defense** A defense that is present in a plant regardless of herbivore damage. [11]

**continental** A climate found in the interiors of large continents at mid-latitudes and characterized by seasonal changes in precipitation and temperature. [16]

**continentality** Influence of land masses on weather. [17]

**controlled burn** A fire that is planned and purposely set; also called a prescribed burn. [13]

**convection** Heat transport by a packet of air or water moving as a unit. [3]

**convergent evolution** The independent evolution of similar features in unrelated taxa. [18]

**Coriolis effect** The tendency of an object moving toward or away from the equator to be deflected, relative to the direction in which it is moving, toward the right in the Northern Hemisphere and toward the left in the Southern Hemisphere. [16]

**correlational selection** A type of selection that occurs when the pattern of selection on one trait depends on the value of another trait. [9]

**cosexual** Of or referring to plants that function as males and females more or less simultaneously. [6]

**cover** The percentage of the ground covered by the canopy of a given species. [12]

**crassulacean acid metabolism (CAM)** A form of photosynthesis in which $CO_2$ is captured by PEP carboxylase and the capture of light energy is temporally separated from the uptake of $CO_2$. [2]

**Croll-Milankovitch effects** Changes in the Earth's orbit around the sun over very long time scales that affect climate. [16]

**crown fire** A fire that burns the tree canopy. [13]

**cryptic dioecy** A system of gender expression in which sporophytes produce only female or only male gametophytes even when apparently capable of producing both types. [6]

**cryptic species** Species that look identical and yet are reproductively isolated. [9]

**cryptogamic crust** A thin crust made of mosses, fungi, algae, lichens, and cyanobacteria. [13]

**cryptophytes** Perennial herbs with perennating organs below the ground surface. [12]

**cultural** The type pf ecosystem service that includes provision of flowers (weddings, funerals), spiritual values (awe), and psychological refreshment and regeneration in nature. [19]

**cumulative survivorship probability** The chance of surviving from an initial time to a new census. [8]

**cuticle** A waxy, nonliving covering over the exposed epidermal cells that prevents desiccation by the air. [3]

**cycle** The major fluxes and pools of a material in a system. [5]

## D

**deciduous** The loss of leaves on a seasonal basis. [7]

**deforestation** The complete removal of a forest. [16, 19]

**degree-days** A measure of the sum of temperatures experienced over a specified time period. [7]

**demographic stochasticity** Variation in vital rates because of random variation in individuals' survival and reproduction. [8]

**dendrochronology** The study of tree rings. [13]

**denitrification** The reduction of $NO_2^-$ and $NO_3^-$ to gaseous nitrogen. [5]

**density** The number of individuals of a species per unit area. [8, 12]

**density-dependent** Of or referring to population processes that vary as a function of population density. [8]

**density-independent** Of or referring to population processes that vary as a function of extrinsic factors unrelated to population density. [8]

**desert** A region where potential evapotranspiration exceeds actual evapotranspiration. [18]

**deterministic** Having a fixed outcome. [13]

**detritus** Debris. Organic detritus is the dead bodies and waste products of living and once-living organisms. [5]

**dew point** The point at which water molecules in the air have lost the kinetic energy to remain a gas and they condense to a liquid state. [16]

**diameter at breast height (DBH)** The diameter of a tree trunk measured at 1.3 m above the ground in some countries and 1.4 in others. [8, 12]

**diaspore** The dispersal unit of a plant, including the seed and any additional related structures. [6]

**differentiation diversity** The diversity of a sample based on the way the species are grouped into subunits. *See also* beta diversity. [12]

**diffuse** The characteristic of competition of many species competing simultaneously, so the competition between any particular pair of species is fairly weak. [10]

**dimension reduction** The process of collapsing highly multivariate data into a small number of dimensions. [12]

**dioecy** A system of gender expression in which at least some individuals in a plant population have only pistillate or only staminate flowers. [6]

**direct gradient analysis** Analysis of species distributions or community similarities along a single environmental gradient. [12]

**directional selection** A type of selection that occurs when individuals with the most extreme trait values have the highest fitness in a population. [9]

**disassortative mating** The mating of individuals with dissimilar phenotypes at a greater frequency than would occur by chance alone. [6]

**dispersal kernel** A model giving the probability of dispersing various distances. [6]

**dispersion** The difference in functional trait values of a particular species from all other species in a set. [12]

**disruptive selection** A type of selection that occurs when individuals with both high and low values of a trait have higher fitness than those with intermediate values of that trait. [9]

**dissected** In reference to leaves, having a blade divided into small but connected parts. [9]

**disturbance** An event that is relatively discrete in time and causes abrupt change in ecosystem, community, or population structure and changes resources, substrate availability, or the physical environment. [13]

**disturbance regime** The characteristics of disturbances occurring in a given ecosystem, generally described by intensity, size, and frequency. [13]

**dominance** Of or referring to the expression of an allele that depends on the properties of the other allele at that locus. [9]

**dominant eigenvalue** The largest eigenvalue. [8]

**dominant species** The most conspicuous and numerically abundant species in a community. [12]

**double fertilization** The process common in angiosperms in which one sperm cell fertilizes the egg cell and a second sperm cell fuses with a specialized cell in the embryo sac and contributes to endosperm. [6]

**drought avoidance** A strategy of growth only during periods of sufficient rainfall, used by plants living in areas with periods of prolonged drought. [3]

**drought-deciduous** Referring to plants that shed leaves in response to low water availability. [3, 18]

**drought-tolerant** Having the ability to live and grow even when experiencing very low water potential. [3]

**dry deposition** The deposition of molecules on a surface as particulate matter. [5]

**dry weight** The weight of a sample after it has lost all its moisture and is at a constant weight. [12]

**dynamic stability** The tendency of a system to return to its original state after a small perturbation. [14]

## E

**eccentricity** The change in the shape of Earth's orbit around the sun, from nearly circular to a more elliptical trajectory. [16]

**ecology** The study of the relationships between living organisms and their environments. [1]

**ecoregions** Major ecosystems defined by distinctive geography, solar radiation, and water status. [19]

**ecosystem** All of the organisms in an area and all of the abiotic materials and energy with which they interact. [5]

**ecosystem concept** The idea that the biological and physical components of the environment form a single interactive system. [5]

**ecosystem service** A direct or indirect benefit that ecosystems provide to people. [19]

**ecotype** One of several populations of a species from different habitats or locations that possess a similar set of genetically based adaptations. [9]

**ectendomycorrhizae** Mycorrhizae with a rudimentary hyphal mantle on the outside of roots, a partial hartig net,

and intracellular clusters of thick-walled hyphal cells. [11]

**ectomycorrhizae (ECM)** Mycorrhizae that form a sheath around the root, and whose hyphae penetrate the intercellular spaces of the root cortex, but not the root cells. [11]

**eddy covariance** A method for measuring net ecosystem exchange. [5]

**edge effect** A systematic difference between the edge of a habitat patch and its interior. [15]

**eigenvalue** A number that, when used to multiply a population vector, gives the same result as multiplication by the transition matrix. Given a matrix $\mathbf{A}$, if $\mathbf{A}\,\mathbf{x} = \lambda\,\mathbf{x}$, then $\lambda$ is an eigenvalue of $\mathbf{A}$, and $\mathbf{x}$ is the associated (right) eigenvector. [8]

**eigenvector** A nonzero vector that, when multiplied by an associated eigenvalue, gives the same result as multiplying by the transition matrix. The eigenvector associated with the dominant eigenvalue gives the stable age or stage distribution of the population. [8]

**El Niño Southern Oscillation** A 3-to-7-year cycle of weather patterns in the Pacific Ocean that drives dramatic global fluctuations in weather. [16]

**elaiosome** A lipid body attached to a seed that serves as a reward to a seed disperser. [6]

**elasticity** The extent to which a proportional change in a transition rate causes a proportional change in the long-term population growth rate. [8]

**eluviation** The removal from soil layers of soil material in suspension or in solution. [4]

**embryo sac** A megagametophyte; a cluster of haploid cells in the ovule of a flowering plant in which fertilization of the egg and development of the embryo occur. [6]

**emergent property** A property that occurs at one level of organization as a result of properties and processes that are unique to that level of organization and is not merely an aggregate of properties and processes at a lower level. [12]

**emergent tree** A tree with a canopy that extends above the general canopy of the forest. [18]

**emissivity** The efficiency of a body's emission of energy. [3]

**endemic species** A species found only in a single, limited area. [14, 19]

**endomycorrhizae** Mycorrhizae whose hyphae penetrate the cells of the root cortex, but do not form a sheath around the root. [11]

**endophytes** Intracellular, symbiotic microorganisms. [11]

**endosperm** A triploid tissue that provides nutrients to the embryo in most angiosperm seeds and that arises through double fertilization. [6]

**energy balance** The amount of energy absorbed, emitted, and stored by an object. [3]

**ensemble** A group of phylogenetically related organisms living in the same place that use biotic or abiotic resources in a similar way. [12]

**environmental stochasticity** Variation in vital rates due to environmental factors that affect all the individuals in a class or population in roughly the same way. [8]

**epigenetic effects** Changes that do not involve alterations of the DNA base-pair sequence. [9]

**epiphyte** A plant that grows on another plant and is not rooted in the ground. [2, 18]

**epistasis** Of or referring to the expression of an allele that depends on the properties of alleles at other loci. [9]

**equability** The extent to which species in a set are equally different from each other in functional trait values. [12]

**equalizing mechanisms** The differences in average fitness between a species and its competitors. [10]

**ericoid mycorrhiza** A symbiosis between a fungus and the roots of a plant in the family Ericaceae. [11]

**errors of development** Small, random differences in when and how genes are expressed during development. [9]

**essential mineral nutrients** Those nutrients that plants need to live and grow. [4]

**evapotranspiration** The transfer of water to the atmosphere by evaporation plus transpiration. [5]

**evenness** The extent to which the species in a community are equally abundant. [12]

**exchangeable base** A cation that contributes to making the soil more alkaline. [4]

**exotic** A species that is outside its native range. [14]

**experiment** A test of a hypothesis. [1]

**extent** A component of scale indicating the total range of a set of data. [15]

**extinction debt** The state in which changing conditions doom species to eventual extinction in a community or region, but there is a lag before they actually disappear. [19]

**extinction probability** The percentage of replicate populations expected to go extinct in a given time period. [8]

## F

**facilitation** Positive interaction among plants; a process by which early successional species increase the survival of later successional species. [10, 13]

**facultative** Existing only under some conditions. [4, 11]

**fecundity** The potential reproductive capacity of an organism measured as the number of gametes or propagules it produces. [8, 9]

**female fitness** Fitness through maturing seeds. [7]

**fertility** The number of viable offspring produced per unit time; the birth rate. [8, 9]

**fibrous root system** A root system that forms a dense network. [3]

**field capacity** The point at which a soil is completely saturated with water. [4]

**floral bract** A specialized leaf below a flower. [6]

**flux** Rate of flow. [5]

**flux equation** A model of flow rates of the following general form: flux = (conductance) × (driving force). [2]

**focus** A component of scale indicating the sampling level or area represented by each data point; also called resolution. [15]

**forb** A broad-leaved herbaceous plant. [2]

**forest** An area dominated by trees in which the canopies touch one another. [18]

**formation** A community defined by its physiognomy. [12]

**fractal dimension** A measure of the complexity or degree of convolution of a line or surface. [15]

**frequency** (of a disturbance) How often, on average, a disturbance occurs in a particular place. *See also* return interval. [13]

**frequency** (of a species in a community) The percentage of samples in which a species appears. [12]

**frequency-dependent selection** A type of selection that occurs when the fitness of a genotype depends on whether it is common or rare in a population. [6]

**frugivory** Fruit eating. [6, 11]

**fruit** The matured ovary or ovaries of one or more flowers and associated structures. [6]

**fuel load** The amount of combustible plant material in an area. [13]

**functional ecology** The study of how the biochemistry and physiology of an individual determines its response to its environment. [2]

**functional diversity** Variation in species' trait values. [12]

**functional group** A group of organisms that contain a set of similar traits that allows them to use biotic or abiotic resources in a similar way. *See also* guild. [12]

**fundamental niche** The range of environmental values in which a species is physiologically capable of growing. [9, 14]

## G

**gametophyte** The haploid generation of a plant that produces gametes. [6]

**gametophytic SI** Self-incompatibility due to the pollen grain's haploid genotype. [6]

**gamma (γ) diversity** Inventory diversity across a landscape or a set of samples including more than one community. [12]

**gap** An opening in a forest canopy. [13]

**gene-environment covariation** A nonrandom relationship between genes and the environment. [9]

**genet** A genetic individual; the product of a single seed. *See also* ramet. [9]

**genetic drift** Change in gene frequencies due to random sampling effects. [9]

**genetic response** That part of natural selection consisting of change in a population across generations. [9]

**genome** An individual's entire DNA sequence. [9]

**genotype** The information contained in an individual's genome. [9]

**genotype-environment interaction** Variation in genetic expression as a function of the environment. [9]

**geoengineering** The deliberate large-scale manipulation of the planetary environment to counteract anthropogenic climate change. [16]

**geographic information systems (GIS)** A method of depicting the relationships between different kinds of geographic information for the same region. [15]

**geometric mean** The *n*th root of the product of *n* numbers. [8]

**grain** (in landscape ecology) A component of scale indicating the size of the primary unit in a set of data. [15]

**grain** (in reproduction) The fruit of a member of the grass family (Poaceae). *See also* pollen. [6]

**grana stack** In chloroplasts, a group of disc-shaped, flattened vesicles bearing photosynthetic pigments. [2]

**granivory** Seed eating. [6, 7, 11, 18]

**gravitational potential** Water potential due to the pull of gravity. [3]

**grazer** An herbivore that eats grasses and other ground-growing plants. [11]

**greenhouse effect** The reabsorption of longwave radiation emitted by the Earth's surface by greenhouse gases. [16]

**greenhouse gas** An atmospheric gas that contributes to the greenhouse effect. [16]

**gross photosynthesis** The total amount of $CO_2$ captured by an individual. [2]

**gross photosynthetic production (GPP)** The total energy (or carbon) fixed by primary producers; also called gross primary production. [5]

**gross primary production (GPP)** The total energy (or carbon) fixed by primary producers; also called gross photosynthetic production. [5]

**ground-truthing** Testing of the reliability of remote sensing data by ground-based measurements. [12, 15]

**groundwater** Underground water that has percolated to a zone that is saturated. [4]

**guard cell** The two cells on either side of a stoma that determine the size of the stomatal opening. [2]

**guild** A group of organisms that use biotic or abiotic resources in similar ways. [12]

**gynodioecy** A system of gender expression in which some plants in a population have only pistillate flowers, while others have either perfect flowers or a mixture of both staminate and pistillate flowers. [6]

**gynomonoecy** A system of gender expression in which both pistillate and perfect flowers occur on the same individuals. [6]

**gyre** A large circular pattern of oceanic circulation. [16]

## H

**habitat** The kind of environment a population or species generally inhabits; an area distinguished by its community. [1, 15]

**habitat construction** The process by which an organism alters its own (or another species') environment. [9]

**Hadley cell** A large pattern of air circulation in which air rises, traverses a distance, falls, and then moves back again along the Earth's surface. [16]

**Hartig net** A complex of mycelia that grows in between the root cortical cells of plants and enmeshes them. [11]

**halophyte** A plant that lives in saline soil. [3]

**heat** A measure of the total kinetic energy of the molecules in a substance. [16]

**heath** A vegetation type dominated by low-growing shrubs in the Ericaceae. [18]

**hemicryptophytes** Perennial herbs with buds at the ground surface. [12]

**hemiparasite** A plant that facultatively taps into the roots of other plants to gain resources. [11]

**herbivory** The consumption of all or part of a living plant. [11]

**heritability ($h^2$)** The resemblance among relatives due to the sharing of genes; the amount of phenotypic variation in a population that is due to genetic variation. [9]

**heritable** Having a genetic basis. [9]

**hermaphrodite** An individual that produces male and female gametes simultaneously. [6]

**heterostyly** A mating system in which different individuals in a population have styles and stigmas of different lengths. [6]

**horizon** A characteristic layer of a soil. [4]

**hybrid swarm** A group of species that hybridize frequently. [9]

**hybridization** Interbreeding between individuals of different species. [9]

**hydraulic lift** A process by which plants take up water in deep soil horizons and then lose it to the soil at upper horizons. [3, 4]

**hydraulic redistribution** Rapid water redistribution in the soil in any direction by roots. [3, 4]

**hygrophyte** A plant that lives in permanently moist soil. [3]

**hypothesis** A possible explanation for a particular observation. [1]

**hysteresis** When a change in a system property lags behind the effect causing the change. [13]

## I

**igneous** Of or referring to rock of volcanic origin. [4]

**illuviation** A process by which soil material is precipitated (from solution) or deposited (from suspension) into a soil layer from an overlying layer. [4]

**imbibition** A process by which early successional species prevent colonization of later successional species by monopolizing resources. [8]

**immobilization** The incorporation of nutrients into soil microorganisms, making them less available for uptake by plants. [5]

**importance value (IV)** The sum of the relative cover, relative density, and relative frequency of a species in a community. [12]

**inbreeding** Mating among relatives. [6]

**inbreeding depression** A decrease in fitness due to inbreeding. [6]

**inclusive fitness** The total fitness of all copies of an allele in a group, usually of siblings. [7]

**indicator species** A species used to place a community in a predetermined group. [12]

**indirect gradient analysis** A form of community analysis in which the stands are ordered by their similarity in species composition, and the environmental factors responsible for the resulting patterns are inferred. [12]

**individual-based model** A model consisting of simulations in which the properties of individual organisms are explicitly modeled. [15]

**induced defense** A response to herbivory that protects the plant, whether or not it harms the herbivore. [11]

**induced resistance** A response to herbivory that harms the herbivore, whether or not it protects the plant. [11]

**induced response** A response that is elicited by an attack by an herbivore. [11]

**inflorescence** An aggregation of flowers. [6]

**infrared gas analyzer (IRGA)** A device that uses infrared light to measure the uptake or respiration of $CO_2$. [3]

**inhibition** A process by which early successional species decrease the survival of later successional species. [13]

**initial floristic composition** A theory of succession that emphasized the process of colonization and differences in the life-spans of species. [13]

**intensity (of a disturbance)** The amount of change caused by a disturbance. [13]

**intercalary meristem** A group of undifferentiated plant cells located between nodes. [6]

**intermediate disturbance hypothesis** A hypothesis that species diversity is highest at intermediate levels of disturbance. [14]

**internode** The stem segment between nodes. [6]

**intertropical convergence zone (ITCZ)** The latitude at which the Hadley cells meet on Earth's surface and the trade winds converge. [18]

**intracellular conductance** For a given concentration gradient, the rate at which a gas moves through leaf cells and cell walls. [2]

**invasive species** A species that is rapidly expanding outside of its native range. [14]

**inventory diversity** The diversity of a sample based on species number or evenness. *See also* alpha diversity, gamma diversity. [12]

**island** An area that is suitable for the survival and reproduction of a species, surrounded by unsuitable habitat. [15]

**island biogeography** *See* theory of island biogeography. [15]

**isobilateral** In reference to leaves, having a distinctive, symmetrical internal architecture with palisade tissue on both the upper and lower sides of the leaf. [3]

**iteroparous** Reproducing in more than one bout in a lifetime. [7]

## J

**jet stream** A river of air in the upper atmosphere that moves at high speeds. [16]

## K

**krummholz** Stunted forest or scrub in windswept alpine regions close to the tree line. [18]

## L

**lag phase** A period during the early population growth of an invasive species when its numbers are low. [14]

**landscape demography** The study of the demographic properties of multiple populations and their drivers across spatial scales. [15]

**landscape ecology** The study of the spatial distribution of individuals, populations, and communities and its causes and consequences. [15]

**latent heat exchange** The transfer of energy in the process of converting water between the liquid and vapor states. [3]

**latent heat loss** The loss of energy through the process of converting water between the liquid and vapor states. [3]

**latex** Milky sap. [11]

**leaching** Loss of a material from a surface soil through water drainage. [4]

**leading eigenvalue** The largest eigenvalue. [8]

**leaf area index (LAI)** The area of photosynthetic surface per unit of ground area. [5]

**leaf axil** The place where a leaf and stem join. [6]

**leaf conductance** For a given concentration gradient, the rate at which a gas moves into or out of a leaf. [2]

**left eigenvector** A vector $\mathbf{y}$ that, when it premultiplies a matrix $\mathbf{A}$, gives the same result as multiplication by an associated eigenvalue; $\mathbf{y}\,\mathbf{A} = \lambda\,\mathbf{y}$. The left eigenvector associated with the dominant eigenvalue gives the reproductive values. [8]

**lethal limit** The highest or lowest temperature at which a plant can remain alive. [3]

**liana** A woody vine rooted in the ground. [18]

**life cycle graph** A drawing showing the size, age, or stage classes of a population that uses arrows to indicate transitions among classes. [8]

**life history** The schedule of birth, mortality, and growth of individuals in a population. [7]

**life table** A list of the estimated mortality, cumulative survival, and related rates of a cohort. [8]

**life table response experiment (LTRE)** A method to examine and analyze matrix demographic data to ask how environmental differences contribute to changes in $\lambda$. [8]

**light compensation point** The light level at which photosynthetic gain exactly matches respiratory loss. [2]

**light reactions** The part of photosynthesis in which light energy is captured. [2]

**light response curve** The rate of fixation of $CO_2$ as a function of light intensity. [3]

**light saturation ($A_{sat}$)** The light level at which the maximum photosynthetic rate is reached. [2]

**lignin** A phenolic compound that impregnates secondary cell walls, providing strength. [11]

**litter** Recently fallen plant material that is only partially decomposed. [4]

**LMA** *See* specific leaf mass. [2]

**loamy soil** Soil with approximately equal proportions of sand, silt, and clay particles. [4]

**local guild** A group of organisms living in the same place that use the same resources. [12]

**loess** Fine unconsolidated windblown sediment. [4]

**log response ratio (LRR)** A measure of the intensity of competition. *See also* absolute competition index, relative competition index. [10]

**longwave radiation** Radiation at wavelengths greater than 700 nm. [3]

## M

**macrofossil** A large fossil, such as a leaf, flower, stem, or seed. [17]

**macronutrient** An element that is needed by plants in large quantities. [4]

**manipulative experiment** A test of a hypothesis that involves making a deliberate change in the physical world. [1]

**mantle** A dense network of hyphae that partially or fully ensheathes the outside of a plant root. [11]

**maritime** A climate often found at the eastern edges of continents, characterized by uniform precipitation throughout the year. [16]

**mass balance** An approach to constructing nutrient budgets, in accounting for the magnitude of pools and fluxes in an ecosystem, that states that inputs – outputs = change in storage. [5]

**masting** Large and erratic variation among years in the size of the seed crop that is synchronized among most of the plants in a population. [7, 11]

**mating system** The biological factors that govern who can mate with whom. [6]

**matric potential** Water potential due to the cohesive force that binds water to physical objects. [3]

**mean field approach** A method of analysis that focuses on describing average parameter values. [12]

**mean field models** Metapopulation models in which all individuals in the population are assumed to have the same properties, so only mean population parameters are simulated. [15]

**mean residence time** The length of time a unit of a nutrient is retained in a leaf. Also called turnover time. [5]

**mediterranean** A climate often found at the western edges of continents, characterized by rainy winters and dry summers. [16]

**meristem** A group of undifferentiated plant cells. [6]

**mesic** Moist. [3]

**mesophyll** The photosynthetic tissue between the upper and lower epidermis of a leaf. [2]

**mesophyll conductance** For a given concentration gradient, the rate at which a gas moves through leaf mesophyll cells and cell walls. [2]

**mesophyte** A plant that lives in moderately moist soils. [3]

**meta-analysis** A statistical synthesis of the results of independent experiments. [1, 10]

**metabolic energy** Energy used by plants for maintenance, growth, and reproduction. [3]

**metacommunity** A group of communities linked by migration and extinction. [15]

**metamorphic** Of or referring to rocks changed by the action of great pressures and temperatures deep underground. [4]

**metapopulation** A group of populations linked by migration and extinction. [15]

**microfossil** A small fossil, such as a pollen grain or single-celled organism. [17]

**microhabitat** The conditions in the immediate surroundings of an individual organism. [1, 15]

**micronutrient** An element that is needed by plants in very small quantities. [4]

**midden** An accumulation of large amounts of plant material in a packrat den. [17]

**migration** The movement of organisms or propagules. [9]

**Milankovitch cycles** Three kinds of changes in the Earth's orbit around the sun that determine the amount of sunlight reaching Earth's surface, seasonal differences at different latitudes, and the advance and retreat of Earth's ice ages. [16]

**mineralization** The set of processes by which microorganisms decompose decaying biomass, releasing carbon as $CO_2$ and nutrients in inorganic form. [5]

**minimum viable population (MVP)** The smallest size necessary to give a population a probability $x$ of surviving $N$ years. [8]

**model** An abstraction or simplification that expresses structures or relationships. [1]

**modular** Having a structure made up of repeated units (in plants, nodes, lateral organs, and internodes). [6]

**monocarpic** Reproducing in a single bout in a lifetime; a botanical term for semelparous. [7]

**monoculture** A population or planting of a single species. [10]

**monoecy** A system of gender expression in which both staminate flowers and pistillate flowers occur on the same individual. [6]

**monothetic-divisive analysis** A classification process based on dividing one group into two groups depending on the presence or absence of a single key species. [12]

**multiple regression** The regression of a single dependent variable simultaneously on multiple independent variables. [12]

**multivariate** Consisting of more than one variable. [12]

**muskeg** A grassy bog habitat with scattered, stunted conifers. [17]

**mutation** A change in a DNA sequence. [9]

**mutualism** A symbiosis in which both organisms benefit. [4, 11]

**mycorrhiza (pl. mycorrhizae)** A symbiosis between a fungus and the roots of a terrestrial plant. [11]

## N

**narrow-sense heritability** Heritability that is due only to additive genetic variation. [9]

**natural experiment** A test of a hypothesis based on a change in the physical world caused by some natural occurrence. [1]

**natural selection** The process by which individuals with different phenotypic characteristics leave different numbers of offspring. *See also* sexual selection. [9]

**nearest neighbor probability** The probability that two types of habitat will be adjacent. [15]

**necrosis** Localized tissue death. [11]

**negative assortative mating** Mating of individuals with dissimilar phenotypes at a greater frequency than would occur by chance alone. [6]

**neophytes** Recent introductions of exotic plant species. [14]

**nestedness** The tendency of communities to be subsets of other communities. [15]

**net ecosystem carbon balance (NECB)** The rate of carbon accumulation or loss in an ecosystem. [5]

**net ecosystem exchange (NEE)** The net transfer of $CO_2$ into or out of an ecosystem. [5]

**net ecosystem production (NEP)** The net accumulation of carbon by an ecosystem over the course of succession. [5]

**net primary production (NPP)** The total energy (or carbon) fixed by primary producers minus respiratory losses. [5]

**net reproductive rate ($R_0$)** The expected number of offspring that an individual will have over its entire life. [8]

**neutral theory of biodiversity and biogeography** A theory that predicts the patterns of local and regional species richness and abundance by using simple assumptions of rates of speciation, dispersal, and random extinction. [19]

**niche** The range of ecological conditions in which a plant species will grow. [9, 14]

**nitrification** The oxidization of ammonia to nitrate. [5]

**nitrifying bacteria** Specialized bacteria that carry out nitrification. [5]

**nitrogen fixation** The process by which nitrogen enters the biosphere from the atmosphere. [4]

**nitrogen saturation** Nitrogen deposition that exceeds the ability of a system to use nitrogen. [5]

**nitrogen use efficiency** The maximum photosynthetic rate per gram of nitrogen in a leaf. [2]

**node** A point of proliferation for the development of leaves or flowers. [6]

**normalized difference vegetation index (NDVI)** The relative reflectance of vegetation in the near-infrared and visible wavelengths. [5]

**null model** A model that describes how data would look if only a specified set of random processes were operating. [19]

**nurse plant** A plant that increases the survival of another plant, especially in very harsh environments such as a desert. [10]

## O

**obligate** Necessary. [4, 11]

**obliquity** The degree of tilt of the Earth's axis. [16]

**observational experiment** A test of a hypothesis based on the systematic study of natural variation. [1]

**old-field succession** Succession in abandoned agricultural fields. [13]

**orchidaceous mycorrhiza** A symbiosis between a fungus and the roots of a plant in the family Orchidaceae. [11]

**ordination** The process of taking points arrayed in an *n*-dimensional space and reducing them to fewer dimensions. [12]

**organic matter** The decaying and decomposed material in soil that comes from living things. [4]

**osmoregulation** The regulation by an organism of its water potential. [3]

**osmotic potential** Water potential due to solutes dissolved in water. [3]

**outbreeding depression** Reduced fitness due to mating with unrelated plants. [6]

**outcrossing** The avoidance of mating among relatives. [6]

**ovary** The ovule-bearing region of a pistil. [6]

**overcompensation** A hypothesized response to herbivory resulting in more growth than in ungrazed plants. [11]

**ovule** In seed plants, the sporophytic structure containing the female gametophyte. [6]

## P

**paleoecology** The study of historical ecology. [17]

**palynology** The study of spores and pollen. [17]

**Pangaea** A supercontinent consisting of all of the continents that existed at the beginning of the Mesozoic era. [17]

**parapatric speciation** Speciation that occurs when a new species arises in a population adjacent to other populations of the ancestral species. [9]

**parasitism** An intimate relationship between two organisms in which one (the parasite) lives on, off, or at the expense of the other (host). [4, 11]

**parent material** The upper layers of the heterogeneous mass that results from the action of weathering and other forces on rock. [4]

**patch** A specified area, either one that is defined arbitrarily or, often, one that is in some way relatively homogeneous or internally consistent. [15]

**pathogenicity** The ability of a microorganism to cause disease in a host. [11]

**pattern** A set of relationships between pieces or entities of the natural world. [1]

**PEP carboxylase** The enzyme responsible for the initial capture of $CO_2$ in $C_4$ photosynthesis. [2]

**percentage base saturation** The proportion of the cation exchange capacity that is occupied by bases. [4]

**perennial** A plant in which the life cycle is completed in more than one year. [7]

**perfect blackbody radiator** An object that absorbs all the radiation that falls on it and emits all of the energy possible for an object at its temperature. [3]

**perfect flower** A flower containing both functional stamens and functional stigmas. [6]

**perianth** The combination of the petals and sepals of a flower. [6]

**perihelion** Earth's nearest point to the sun. [16]

**permafrost** A permanently frozen layer of soil. [18]

**permanent wilting point** The soil water potential at which plants can no longer can extract water from the soil and exist in a permanently wilted condition. [4]

**pH** The negative logarithm of the concentration of $H^+$ ions. [4]

**phanerophytes** Trees or tall shrubs with buds more than 25 cm above the ground. [12]

**phenolic** A chemical compound consisting of an aromatic ring with an attached hydroxyl group. [11]

**phenology** The timing of life history events, such as initiation of growth, flowering, and dormancy. [2, 7]

**phenotype** All of the physical attributes of an organism. [9]

**phenotypic plasticity** Variation in the phenotype of a single genotype caused by environmental influences. [2, 9]

**phenotypic selection** That part of natural selection consisting of the relationship between phenotypic variation and fitness differences. [9]

**phloem plugging** A defense against infection in which the phloem clogs up, preventing the spread of the infectious agent. [11]

**photon** A discrete packet of energy that travels at the speed of light. [3]

**photoperiod** The number of hours of light and darkness in a 24-hour cycle. [2]

**photorespiration** A process in plants that consumes $O_2$ and releases $CO_2$, which depends on light. [2]

**photosynthesis** The biochemical processes by which an organism captures energy from sunlight and fixes carbon from the atmosphere. [2]

**photosynthetic induction** The start-up time required for plants to reach maximum photosynthesis after exposure to bright light. [2]

**photosynthetic photon flux density (PPFD)** The amount of photosynthetically active radiation impinging upon a leaf per unit time. [2]

**photosynthetically active radiation (PAR)** The wavelengths of light that can be used in photosynthesis. [2]

**photosystem** An array of chlorophyll and accessory pigment molecules that captures light energy. There are two photosystems: I and II. [2]

**phreatophyte** A plant with roots that extend to the water table. [3]

**phylogenetic diversity** The extent to which species within a set (such as in a community) are closely or distantly related. [12]

**phylogeny** A pattern of relationships among species (or higher taxa) based on evolutionary ancestry. [2, 12]

**physiognomy** The form, structure, or appearance of a plant community. [12]

**physiological ecology** The study of the physiological mechanisms that underlie whole-individual responses to the environment. [2]

**phytoalexin** A secondary compound that acts as a defense against pathogens. [11]

**pistillate flower** A flower with functional pistils but not functional stamens. [6]

**plantlet** A small plant created vegetatively. [6]

**plant litter** A variety of materials including dead roots, dead tree trunks, and shed leaves, needles, and bark. [5]

**plasmodesmata** Thin strands of living tissue that form connections between plant cells. [2]

**pollen** A microgametophyte; a microspore containing a mature or immature male gametophyte. A pollen grain is a single pollen individual. [6]

**pollination syndrome** A particular combination of floral traits that is associated with a particular type of pollinator. [6]

**polycarpic** Reproducing in more than one bout in a lifetime. [7]

**polyploidy** A change in chromosome number due to the duplication of the entire set of chromosomes. [9]

**polythetic-agglomerative analysis** A classification process based on linking groups into a hierarchy based on the similarity of objects or groups of objects. [12]

**pool** The stored quantity of a nutrient or element. [5]

**population dynamics** Changes in the number of individuals, genetic composition, structure, or spatial dispersion of a population. [8]

**population structure** The relative frequency of individuals in each size class, age class, or stage class. [8]

**population vector** A list of the numbers of individuals in each stage class at a given time. [8]

**porosity** In soil, the total volume of pores and their size, shape, and arrangement between and within aggregates. [4]

**potential evapotranspiration (PET)** The maximum amount of water that would be lost by evapotranspiration in a particular place if water were freely available in the soil and plant cover were 100%. [5]

**precession** The wobble of the Earth's axis. [16]

**precipitation** The amount and pattern of rainfall and snowfall. [16]

**prescribed burn** A fire that is planned and purposely set; also called a controlled burn. [13]

**pressure potential** Water potential due to the hydrostatic or pneumatic pressure in a system. [3]

**primary growth** The process of increasing length that takes place in shoots or roots. [6]

**primary metabolite** A compound necessary for the basic functioning of a plant. [11]

**primary productivity** The amount of carbon transformed from $CO_2$ to organic carbon per unit area in a specified time period. [5]

**primary research** The gathering of information or finding out of facts not known before. [1]

**primary succession** Succession in which plants colonize ground that was not previously vegetated. [13]

**productivity** The amount of carbon or energy transferred from one trophic level to the next per unit area in a specified time period. [5]

**propagule** A seed, vegetative reproductive unit, or other dispersal structure. [10]

**propagule rain** In metapopulation models, occupancy-independent immigration. [15]

**proteoid roots** Dense clusters of short, fine lateral rootlets that produce organic acids or other chelating agents. [4]

**provisioning** The type of ecosystem service that includes food, lumber, medicines, and genetic resources. [19]

**pubescence** A mat of short hairs on a leaf surface. [3]

**pyrogenic** Fire-promoting. [13]

## Q

**quadrat** A sample plot. [12]

**quantum yield** (also quantum efficiency) The number of moles of $CO_2$ fixed per mole of photons absorbed. [2]

## R

***R**** The average amount or concentration of a resource remaining in the environment after a population of a single species grown alone has used all that it can take up. [10]

**radiant energy** Energy transferred from one object to another by photons. [3]

**radicle** The structure that will develop into the root structure. [6]

**rain shadow** An area on a mountain's leeward side that tends to have low precipitation. [16]

**ramet** A potentially physiologically independent unit of a genet. [9]

**rarefaction** A method for estimating the total number of species in an area based on a sample of individuals. [12]

**realized fertility** The chance of surviving to maturity times the fecundity of survivors. [8]

**realized niche** The environmental range within which a species is found. [9]

**recruitment** The establishment of seedlings. [8]

**regeneration niche** The set of environmental requirements necessary for germination and establishment of a plant species. [10]

**regulating** The type of ecosystem service that includes flood control by marsh plants and mangrove ecosystems, decomposition by microorganisms, and regulation of rainfall and the atmosphere by forests. [19]

**relative competition index (RCI)** A measure of the intensity of competition. *See also* absolute competition index, log response ratio. [10]

**relative growth rate (RGR)** A measure of growth in terms of increase in weight per unit time with respect to initial weight. [10]

**relative humidity** The water vapor in the air as a percentage of the amount in fully saturated air at the same temperature. [16]

**relay floristics** A theory of succession that emphasized the facilitation of late successional species by earlier ones and the replacement of one integrated community by another. [13]

**relevé** A single, large sample plot. [12]

**remote sensing** The process of collecting data about an entity of interest without being in direct contact with that entity. [5]

**reproductive effort** The fraction of resources allocated to reproduction at a given age or stage. [7]

**reproductive isolation** A condition in which two populations are unable to share genes due to barriers to mating or inviability of offspring. [9]

**reproductive value (RV)** The contribution in offspring that an average individual now in stage $x$ will make to the next generation before it dies. [8]

**rescue effect** In metapopulation models, the dependence of population extinction on the frequency of occupied patches. [15]

**resistance** The ability of a potential host to avoid herbivory or infection by a pathogen. [11]

**resolution** A component of scale indicating the sampling level or area represented by each data point; also called focus. [15]

**response surface experiment** An experimental design for testing the effects of competition by altering both the total density and the frequency of neighbors. [10]

**retention time** The average length of time a material resides in an ecosystem component. [5]

**return interval** The average time between disturbances in a particular place. [13]

**rhizome** An underground horizontal stem growing near the soil surface. [6, 9]

**rhizosphere** The microbe-rich environment immediately surrounding plant roots. [4]

**richness** *See* species richness.

**riparian** Adjacent to a stream or river. [13]

**root-to-shoot ratio** The ratio of root mass to aboveground tissue mass. [3]

**root exudates** Chemicals released from plant roots that attract, sustain, and interact in complex ways with the many different species of bacteria, fungi, nematodes, and other organisms living in that environment. [4]

**rosette plant** A plant with short or unnoticeable internodes. [6]

**rubisco** RuBP carboxylase/oxygenase; the enzyme responsible for the initial capture of $CO_2$ in $C_3$ photosynthesis or the Calvin-Benson cycle. [2]

**ruderal** Weedy. A ruderal species is one that grows well in temporary habitats and areas subject to frequent disturbance. [7, 14]

**rumen** A specialized stomach compartment of some mammalian grazers. [11]

## S

**safe site** A favorable spot for seed germination and establishment. [13]

**sampling error** Variation in the estimate of a parameter because of differences between a sample and the entire population from which it comes. [8]

**sandy soil** Soil with more than 50% sand particles. [4]

**saturated** The condition of soil when all of its pores are filled with water. [4]

**saprophyte** A plant or fungus that acquires carbon from organic detritus. [5]

**scale** The spatial or ecological context of a set of data, measured as grain, extent, and focus. [15]

**scarification** The abrasion of seeds, required by some seeds for germination. [6]

**scientific method** A process for obtaining knowledge of the natural world that consists of observations, descriptions, quantifications, posing of hypotheses, tests of those hypotheses using experiments, and verification, rejection, or revision of the hypotheses, followed by retesting of the new or modified hypotheses. [1]

**sclereid** A sclerenchyma cell with thick, lignified, much-pitted walls. [11]

**sclerenchyma** A plant tissue containing elongated fiber cells. [11]

**sclerophyllous** Having leaves that are tough, evergreen, relatively small, roundish, and thick. [3, 18]

**secondary chemical** A compound that serves one of a wide variety of functions outside a plant's basic metabolism, including defense and attraction of pollinators. [11]

**secondary growth** The process of increasing in girth by producing woody tissues. [6]

**secondary metabolite** *See* secondary chemical. [11]

**secondary research** Gathering of data or confirmation of facts that are already known. [1]

**secondary succession** Succession in which plants colonize ground previously occupied by a living community. [13]

**sedge** A grasslike plant in the Cyperaceae. [18]

**sedimentary** Of or referring to rock formed by the deposition and recementation of material derived from other rock. [4]

**seed** An embryonic sporophyte embedded in a female gametophyte and covered with one or more integuments derived from the maternal sporophyte. [6]

**seed bank** The seeds buried in the soil. The term can refer to either a single species or an entire community. [6]

**seed fern** An extinct group of vascular plants that superficially resembled ferns but reproduced through seeds rather than spores. [17]

**seed predator** An herbivore that consumes seeds or grains; a granivore. [11]

**seed trap** A container designed to capture windblown seeds. [6]

**selective logging** The removal of only some trees. [13]

**self-fertilizing** The mating of an individual with itself. [6]

**self-incompatibility** The prevention of self-fertilization or matings between certain individuals by a genetically based recognition system. [6]

**self-pruning** The shedding of branches by a shrub or tree. [6]

**self-thinning** The process of density-dependent mortality. [10]

**semelparous** Reproducing in a single bout in a lifetime. [7]

**semelparous perennial** An individual that lives more than 1 year but reproduces in a single bout in a lifetime. [7]

**sensible heat** Heat that can be sensed; conduction + convection. [3]

**sensible heat exchange** Energy exchange that results in a change in temperature. [3]

**sensitivity** The extent to which an absolute change in a transition rate causes a change in the long-term population growth rate. [8]

**sequential hermaphroditism** A life history in which an individual begins life as one gender and then switches to the other. [6]

**seral** Of or referring to a plant community that is part of a successional sequence. [13]

**serotiny** The retention by cones or fruits of seeds that are released only after exposure to heat, normally by a fire. [6, 13]

**sexual selection** The process by which individuals with different phenotypic characteristics leave different numbers of offspring due to differences in access to mates and fertility. [6]

**shoot** The stem and leaves of a plant. [6]

**shortwave radiation** Radiation at wavelengths less than 700 nm. [3]

**shrubland** A plant community dominated by low-growing woody plants with multiple stems. [18]

**silty soil** Soil with more than 50% silt particles. [4]

**similarity measure** A measure of the extent to which two communities share the same species, based on presence or abundance. [12]

**sink population** A population that is maintained by immigration. [15]

**size (of a disturbance)** Spatial extent. The amount of area affected by a disturbance. [13]

**size hierarchy** A highly unequal size distribution among individuals in a population or community. [10]

**size-structured** A demographic model in which individuals are characterized based on their size. [8]

**soil erosion** The removal of the terrestrial substrate by liquid water, ice, or wind, in the presence of gravity. [4]

**soil order** The broadest category in a system of soil classification. [4]

**soil profile** The sequence of horizons that characterizes a soil. [4]

**soil series** A designation based on the texture of the surface layer, topography, parent material, and the vegetation under which the soils were formed and defined at the regional level. [4]

**soil solution** The water in the soil and its associated dissolved minerals. [4]

**soil structure** The physical arrangement of soil particles into larger clusters. [4]

**soil texture** The relative proportions of sand, silt, and clay particles in a soil. [4]

**soil type** A designation based on the texture of the surface layer, topography, parent material, and the vegetation under which the soils were formed and defined at the landscape level. [4]

**solar tracking** The tracking of the sun across the sky during the day by a leaf. [2]

**somatic mutation** A mutation in a non-gamete-producing cell. [9]

**source population** A population that supplies migrants to other populations. [15]

**spatial autocorrelation** Resemblance as a function of distance. [9]

**spatially explicit approach** A method of analysis that depends on determining the particular spatial arrangement of species and landscape elements. [12]

**speciation** The process of origination of new species. [9]

**species** *See* biological species concept, taxonomic species. [9]

**species-area curve** A mathematical function showing how the number of species increases as the area sampled increases. [12]

**species density** The number of species per unit area. [12]

**species pool** The species in a region available to colonize a site. [13]

**species richness** The number of species. [12]

**species richness relationships** Any relationship that describes how species richness changes as a function of the area and/or time over which those species are sampled. [12]

**species sorting** An ecological process by which the same or closely related species are found in a given biome because that form is well suited to that environment. [18]

**species-area curve** A graph describing the increase in the number of species found as the area sampled increases. [12]

**specific heat** The amount of heat energy required to warm 1 g by 1°C. [16]

**specific leaf area** Leaf area per gram of dry weight. [2]

**specific leaf mass** Leaf dry weight per unit leaf area; also called LMA (leaf mass per leaf area) or specific leaf weight. [2]

**specific leaf weight** *See* specific leaf mass. [2]

**sporophyte** The diploid generation of a plant that produces spores. [6]

**sporophytic SI** Self-incompatibility due to the pollen parent's diploid genotype. [6]

**spring annual** A plant that germinates in early spring, grows quickly, and flowers by late spring or early summer. [7]

**spring ephemeral** A plant that is active above ground only for a short time in spring. [2]

**stabilizing mechanisms** The differences between the niches of competing species. [10]

**stabilizing selection** A type of selection that occurs when individuals with intermediate trait values have the highest fitness in a population. [9]

**stable growth rate** The rate of population growth achieved when the proportion of individuals in each age or stage class ceases to change. [8]

**stable stage distribution** The population structure at which the proportion of individuals in each stage class stays constant each generation; also called a stable structure. [8]

**stage-structured** Of or referring to a population in which individuals are classified by some combination of age, size, developmental stage, or other indicators of status. [8]

**stable structure** *See* stable stage distribution. [8]

**stamen** The pollen-bearing organ or microsporophyll of a flowering plant. [6]

**staminate flower** A flower with functional stamens but not functional pistils. [6]

**stand** All of the plants in an area; a group of trees growing together. [12]

**standing biomass** The sum of all organic matter in living vegetation, including the nonliving structural material in wood. [5]

**standing crop** The total of living and dead biomass in an area that is not lying on the ground. [14]

**state-structured** A population in which individuals are classified by their life history state. [8]

**Stefan-Boltzmann equation** An equation describing the radiant energy emitted by the surface of an object. [3]

**stipule** An appendage or bract situated at either side of a leaf axil. [18]

**stochastic** Random. [8]

**stoichiometry** The relative proportions of different elements. [4]

**stolon** A stem that runs along the ground. [6]

**storage effect** A mechanism of species coexistence in which recruitment fluctuations occur, with each species responding differently to the environment and thus recruiting in different years. [10]

**strategy** A set of coordinated adaptive traits. [3]

**stroma** The watery matrix that fills the chloroplast. [2]

**stroma lamellae** A membrane in a chloroplast that connects grana stacks. [2]

**structural coloration** Color caused by the way light bounces off a physical surface with particular optical properties. [2]

**substitutive design** An experimental design for testing the relative strength of intraspecific versus interspecific competition by altering the frequencies of two hypo-thesized competitors while keeping the total density constant. [10]

**subtropical** Of or referring to an area that rarely experiences temperatures below freezing. [18]

**succession** Directional change in community composition and structure through time. [13]

**succulent** A plant that stores large amounts of water in leaves, stems, or other tissues. [3]

**sucker** A shoot that arises from a bud on the near-surface roots of a woody plant. [6]

**summer annual** A plant that completes its aboveground life cycle in the warm months. [7]

**superorganism** A view of communities according to a theory proposing that they are analogous to individuals and are born, develop, grow, and eventually die. [12]

**supervised classification** A classification process that uses remote sensing data and requires that groups be defined explicitly by the user. [12]

**supporting** The type of ecosystem service that includes pollination (crops, forest species) and nutrient cycling. [19]

**surface fire** A fire that runs along the soil surface or ground vegetation. [13]

**survivorship** The chance of surviving. [9]

**symbiosis** A relationship between two intimately interacting organisms. [11]

**sympatric speciation** Formation of two species from a single population in one location. [9]

**synecology** The ecological study of whole communities. [1]

## T

**tannin** A phenolic compound that reduces the digestibility of plant tissues. [11]

**taproot** A root that extends deep into the soil and may also be thickened and capable of storing food. [3]

**taxonomic guilds** Groups of related organisms that use the same class of resources in a similar way. [12]

**taxonomic species** A monophyletic group that shares a set of unique traits. [9]

**temperature** The average random kinetic energy of the molecules in a substance. [3, 16]

**temperature lag** Delay in the effects of warming and cooling due to the storage of energy in the atmosphere, ground, and ocean. [16]

**temporal niche** The range of time in which a plant species lives. [10]

**terpene** A compound composed of multiple units of isoprene ($C_5H_8$). [11]

**tetraploid** Having twice the usual number of chromosomes ($4n$). [9]

**theory** A broad, comprehensive explanation of a large body of information. [1]

**theory of island biogeography** A theory that predicts the patterns of species' presence and absence on islands or in island-like habitats. [15]

**thermohaline circulation** Oceanic circulation driven by differences in the density of water. [16]

**–3/2 thinning law** A hypothesis that density-dependent mortality results in a consistent relationship between population number and mass of –3/2 on a log-log scale. [10]

**threshold effect** An abrupt change in the properties of a system when that system reaches a certain point. [15]

**throughflow** Rainwater that passes through the canopy and travels down the surface of tree trunks, carrying dissolved ions leached from the plant surfaces. [5]

**thylakoid membrane** A double membrane in a chloroplast that bears photosynthetic pigments, consisting of grana stacks alternating with stroma thylakoids. [2]

**tiller** A node on a rhizome that can produce new leaves and roots. [9]

**tolerance** (in herbivory) The ability of a plant to maintain its fitness when consumed by an herbivore or infected by a pathogen. [11]

**tolerance** (in succession) A process by which early successional species neither increase nor decrease the survival of later successional species. [13]

**top-down population control** Limit of herbivore populations by their predators. [11]

**tracheid** A type of elongated cell that forms the water-conducting tissue in primitive vascular plants. [3]

**trade winds** Winds moving from 30°N and 30°S toward the equator, from northeast to southwest in the Northern Hemisphere and from southeast to northwest in the Southern Hemisphere. [16]

**trade-off** A situation in which increasing one feature of an organism necessitates decreasing another feature. [9]

**transect** A line along which samples are taken. [9, 12]

**transgenerational plasticity** The indirect effect by the environment on the phenotype of individuals in subsequent generations. [9]

**transient analysis** Analysis of fluctuations in population growth rates. [8]

**transition matrix model** A matrix of the survival and reproductive rate coefficients of a population. [8]

**transition zone** An area of contact between two communities, regions, or biomes. [19]

**tree line** An often abrupt demarcation at high altitude or high latitude where trees are replaced by low-growing vegetation. [18]

**trichome** A single-celled plant hair. [11]

**trophic cascade** A process by which predators reduce the population size of their prey, setting off still more changes at the next trophic level. [11]

**troposphere** The lowest layer of the atmosphere, where living organisms reside. [16]

**tundra** A biome at high altitude or high latitude dominated by herbaceous plants because the temperature is too cold for tree growth. [18]

**turnover** (of species along a gradient) The average number of species that appear and disappear as you pass from one community to the next along a gradient. [12]

**turnover time** A measure of how rapidly materials move through a system; the total mass of a component material divided by its flux in or out of the system. Also called mean residence time. [4, 5]

## U

**unified theory** A theory that consists of a few, general propositions that characterize a wide domain of phenomena and from which can be derived an array of models. [1]

**univariate** Concerning a single variable. [12]

**unsupervised classification** A classification process that uses remote sensing data and separates pixels into a user-defined number of classes based on statistical similarities between the pixels. [12]

## V

**vapor pressure deficit** The difference between water vapor in the leaf and the air. [3]

**vascular cambium** A large, tubular meristem just inside the stem, responsible for secondary growth of xylem and phloem. [6]

**vegetation lag** A situation in which change in the vegetation happens more slowly than climate change. [17]

**vegetative reproduction** Reproduction by vegetative growth of a new ramet. [6]

**vernalization** Acceleration of germination or flowering by exposure to cold temperatures. [7]

**virulence** The ability of an infectious agent to produce disease; a measure of the degree of damage inflicted by a pathogen on the host organism. [11]

**vital rates** The reproductive rates and chances of survival of individuals in a population. [8]

**volatilization** To cause to pass off in vapor. [5]

## W

**water potential** The difference in potential energy between pure water and the water in some system. [3]

**water use efficiency** The grams of carbon fixed in photosynthesis per gram of water lost in transpiration. [2, 3]

**watershed** An area that includes a stream or river and all of the land that drains into it. [1]

**weather** The immediate or short-term climatic conditions in a given area. [16]

**weedy** Of or referring to a species growing where it is not wanted, for example, an uncultivated species that proliferates in agricultural settings. Also called ruderal. [14]

**westerlies** Winds moving from west to east from 30°N and 30°S toward the poles. [16]

**wet deposition** The deposition on a surface of molecules dissolved in water. [5]

**windthrow** Blowdown of a branch, part of a tree, a whole tree, or a group of trees. [13]

**winter annual** A plant that germinates in autumn, grows through the winter, reproduces in spring, and then dies. [7]

**Wisconsin glaciation** The most recent glacial advance in North America, named for one of the locations of its southernmost extent. [17]

**wood** Secondary xylem. [6]

**woodland** An area dominated by trees that are spaced apart, varying from closely spaced to sparse. [18]

## X

**xeric** Dry. [3]

**xerophyte** A plant that lives in regions with frequent or extended droughts. [3]

**xylem** The tissue that carries water throughout a vascular plant. [3]

**xylem vessel** A vessel that transports water in a plant. [3]

**xylem vessel element** A type of cell found in some angiosperms that forms xylem vessels. [3]

for more ebook/ testbank/ solution manuals requests: email 960126734@qq.com

# Index

Page numbers in *italic* indicate material in an illustration.

## A


A horizons, 94, 95
Abert's buckwheat (*Eriogonum abertianum*), 195, 196, *197*
*Abies,* 532
*Abies amabilis* (Pacific silver fir), 127, 387
*Abies balsamea* (balsam fir), *509,* 532
*Abies balsamifera. See* Abies balsamea
*Abies concolor* (white fir), 430, *431, 511*
*Abies procera* (noble fir), 387
*Abies sibirica* (Siberian fir), 399
Abiotic factors
  concept of, 340
  habitats and, 13, 424
  heterogeneity of ecological phenomena and, 13
Abramsky, Zvika, 282
Abscisic acid, 67
Absolute competition index (ACI), 271, 286
Absorption spectra, *25*
Abundance
  abundance-weighted measures of species richness, 343, 344
  biodiversity and, 342
  geographic range and, 400
  inventory diversity and, 343–344, *345*
  measures of species abundance, 356
  using to describe species distribution across landscapes, 365–366
  *See also* Local abundance
Abundance curves, 411–412
Abundance magnitude, 343
Abundance variability (evenness), 343–344
*Acacia,* 107, 316, 527
*Acacia cyclops,* 548
*Acacia saligna,* 548
Acacia savanna, 536
Acanthaceae, *40*
Acaulosporaceae, *325*
Accessory pigments, 22, 24
Acclimate/Acclimation, 48, 54, 233
Acclimatization, 233
*Acer. See* Maples
*Acer rubrum* (red maple), *173,* 281, 326, *388, 504*
*Acer saccharum. See* Sugar maple
*Acer spicatum* (mountain maple or moosewood), *339*
Acetylcholine, 315
*Achillea,* 241, 251–253
*Achillea lanulosa, 251, 253*
*Achillea millefolium,* 398
*Achnatherum hymenoides* (Indian ricegrass), 302
ACI. *See* Absolute competition index
*A*-$C_i$ curves, 56
Acid precipitation, 132, 133, 136
Acidic soils, 93, 94
Acorales, *40*
Acorns, 303
*Acrobeloides nanus, 124*
*Actinotus helianthi* (flannelflower), 228–229
Action spectra, *25*
Activity limit, 80
Actual evapotranspiration (AET), 127, 139
Adaptations
  adaptive plasticity to light conditions, 243–245
  environmental effects over generations, 245
  heavy-metal tolerance, 241–243, 329, 521
  natural selection and patterns of, 240–246
  through hybridization, 256–257
Adaptive speculations, 231–232
Adaptive traits, 232–235
Adder's tongue (*Botrychium australe*), 308
Additive designs, 272
Additive series designs, 272
Additive variance, 239
*Adelges tsugae* (hemlock woolly adelgid), 309
Adenosine triphosphate (ATP), photosynthesis and, *23,* 24–25, *35,* 37
*Adenostoma fasciculatum* (chamise), 387, *534*
Adhesion, *93*
Adiabatic lapse rate, 459
Adler, Peter, 203, 294, 354, 564
Adventitious roots, 145, *146*
*Aegiceras corniculatum* (khalsi), 489
*Aegopodium podagraria* (Bishop's goatweed), *157*
*Aeonium, 39*
Aerenchyma, 72
Aerts, Rein, 135, 282
*Aesculus,* 529
*Aesculus octandra* (yellow buckeye), *339*
AET. *See* Actual evapotranspiration
Afforestation, 480–481, 492
Africa, Mesozoic era, 499, *500*
African beli (*Paraberlinia bifoliolata*), 281
African buffalo, 382

African daisy (*Osteospermum pinnatum*), *62*
Agamospermy, 150, 255–256
*Agathosma ciliata, 534*
Agavaceae, 42
*Agave,* 150, 183–184
*Agave palmeri* (blue or Palmer agave), *43*
*Agave parryi* (Parry's agave), *64*
Agaves, 42, 64–65
Age-based methods, 203
Age-structured populations, 203, 215–216
Agrawal, Anurag, 315
Agricultural land
  primary productivity, *524*
  succession and, 372, 377, 394
Agriculture
  carbon emissions and, *479,* 481
  crop mating systems and gene flow, 169–170
  impact on temperate grassland, 534–535
  as a source of greenhouse gases, 482
  *See also* Cattle
*Agrostis capillaris,* 308
*Agrostis stolonifera* (creeping bentgrass), 148, 425–426
*Agrostis tenuis,* 241, *243*
Agulhas Current, *462*
Airborne laser-guided imaging spectroscopy, 428–429
Akçakaya, Resit, 434
Alanine, *35*
Alberta wild rose (*Rosa acicularis*), *61*
Alders, 107, 386, 532
Alexander, Helen, 209
Alfisols, *87, 88*
Alismatales, *40*
Alkaloids, 312, 313, 314–315, 318
Allele frequencies, heritability of a trait and, 236–237
Allelopathy, 262, 273, 276–279
Allen, Emily, 556–557
*Alliaria petiolata* (garlic mustard), 310
*Allium,* 150
*Allium tricoccum* (wild leek), 28–29, 33
Allopatric speciation, 254
Allophane, 90
Allopolyploidy, 255
Almendro tree (*Dipteryx panamensis*), *381*
*Alnus,* 107, 386, 532
*Alnus rubra* (red alder), 387
*Alnus sinuata,* 386, 392
*Aloe,* 145, *146*
*Aloe arborenscens* (candelabra aloe), *43*
*Aloe dichotoma* (quiver tree), *333*
Alofs, Karen, 429
*Aloidendrum dichotomum, 146*
Alpha (α) diversity, 344, 345, 347, 550
Alpine forget-me-not (*Myosotis alpestris*), 399–400
Alpine grassland, *515, 516,* 539, *540*
Alpine plants
  convergent evolution, 521
  effects of global climate change on, 491–492
  genetic variation in responses to day length, 46–47
Alpine shrubland, *515, 516,* 539
Alpine sorrel (*Oxyria digyna*), 46–47, 540, 541
Alpine tree line, 518, *519*
Alpine tundra, *514,* 540–541
Alps, 482, 484, 491, 551
*Alternanthera,* 41
Alternation of generations, 150–153, 498
Altitudinal gradient in species frequencies, 337–338, *339*
Aluminum, *91, 93,* 94
Aluminum oxide, 135
AM. *See* Arbuscular mycorrhizae
*Amanita muscaria,* 327, *328*
Amaranthaceae, 39, *40,* 325, *509*
*Amaranthus caudatus* (love lies bleeding), *40*
Amazon Basin
  deterministic succession in tropical rainforests, 394
  effects of plants on precipitation, 138
  habitat fragmentation and the Biological Dynamics of Forest Fragments Project, 441–443, 444
  human population growth and biodiversity loss, 571
  hydraulic redistribution, 58–59
  land use changes and anthropogenic carbon emissions, 482
  Pleistocene glaciation and, 507
  species abundance and rarity, 400, *401*
  tree species richness and secondary succession, 386
  *See also* Tropical rainforest
*Ambrosia deltoidea,* 392
*Ambrosia dumosa* (burro bush), *61*
*Ambrosia psilostachya,* 148
American beech (*Fagus grandifolia*), *339,* 375, *384,* 509, *510, 529*
American bison, *297,* 300, 382
American chestnut (*Castanea dentata*), 322–323, 384, 508, 509, 529, 570
American elm (*Ulmus americana*), 324, *338*
American hornbeam (*Carpinus caroliniana*), *339*
Amino acids, 105, 129
Amino groups, 130
Ammonia, 105, 129
Ammonification, 129, 131
Ammonium
  Haber process, 134
  nitrogen cycle and, 129–130, 131
  nitrogen deposition and, 134
  plant nutrition and, 105
  volatilization by cattle, 133
  wetland soils, 130, 131
Ammonium nitrate, 134
*Ammophila arenaria* (marram grass), 324
*Ammophila breviligulata* (beach grass), 375
Amoebas, 125
*Amorphophallus titanum* (voodoo lily), 75, 159
*Amphicarpum,* 168
Amphistomatous plants, 69
Amur honeysuckle (*Lonicera maackii*), 405
*Anabaena,* 106, 107
*Anabaena azollae, 106*
Anaerobic stress proteins, 67
Analysis of variance (ANOVA), 360
Andaman, *559*
Anderson, Kristina, 386–387
Andes, 464, 484
Andrewartha, Herbert, 300
Androdioecy, 163
Andromonoecy, 163
*Andropogon gerardii* (big bluestem), *40,* 534
*Andropogon semiberbis,* 228
*Andropogon virginicus* (broomsedge), 372
Aneuploidy, 256
Angiosperms
  adaptations to terrestrial environments, 54–55
  dominance in the Cenozoic era, 503–504
  dominance in the Cretaceous period, 501–503
  evolutionary history and the geological time scale, *496*
  evolutionary relationships of land plants, *497*
  forms of gender expression, 163, 164
  mycorrhizae and, 325 (*see also* Mycorrhizae)
  number of known species, 544
  phylogeny of $C_4$ photosynthesis, *40*
  Pleistocene extinctions, *552*
  sexual life cycle, 151–153
  xylem, 73, *74*
Angle of incidence, 451–452, *453*
Animals
  disturbance and, 382, *384*
  evolutionary consequences of plant-herbivore interactions, 316–317
  perception of color, 155
  plant reproductive phenologies and, 196–197

seed dispersal and, 170–173
*See also* Pollinators
Ankerwycke Yew, *122*
Annosus root and butt rot, 320
Annual growth rings, 376
Annuals
adaptations to arid environments, 62–63
demographic life history theory, 184
environmental stochasticity and population growth in desert annuals, 223–224, *225*
in hot deserts, 538
life history variations, 182–183
Raunkiaer system of community physiognomy, *357*
seed banks and bet-hedging life histories, 190, 191
ANOVA. *See* Analysis of variance
Anoxic conditions, 504
Antarctic, global climate change and, 485
Anther smut, *408*
Anthers, 151, *152, 154*
Anthocyanidin, *312*
Anthocyanins, 50, 313, 319
*Anthoxanthum odoratum,* 241
Anthropogenic erosion, 96
Anthropogenic global climate change
basic scientific facts concerning, 471
biodiversity loss and, 570
Camp Fire and, 380
challenges for modern plant communities, 511
direct effects on plants of increasing levels of atmospheric carbon dioxide, 474
global carbon cycle and, 471, 472–474, 477–482
human efforts to address, 492–493
impact on global biodiversity, *439*
impact on species and ecological systems, 489–492
impact on the fynbos, 548
long-term community studies, 357
predicted changes from, 486–489
real and probable effects of, 471–472, 483–485
significance for small populations and conservation, 250
tundra and, 541
*Anticharis, 40*
Ants
fynbos, 547–548
seed dispersal and, 173
Apatite, 108, 135
Aphelion, *455,* 456
Aphids, 298, *299,* 324, *408*
Apiales, *40*
Apical dominance, 144
Apical meristems, 144, 145
*Apis mellifera. See* Honeybees
Apocynaceae, 156, 159, 313
*Apocynum,* 313
Apomixis, 149–150
Appalachian forests, 395
Appalachian Mountains, 465
Apparency hypothesis, 314
Apparent competition, 308
*Aptenia cordifolia* (baby sun rose), *60*
Aquatic plants
defined, 61
phenotypic plasticity, 241, *242*
pollination, 161–162
Aquifoliales, *40*
Arabian Desert, 459, 537
*Arabidopsis thaliana,* 182, 245
*Arabis,* 321
Araceae, 156, 159
Arachnids, 125
*Araucaria,* 500
*Araucaria araucana* (monkey puzzle tree), 500
*Araucaria excelsa* (Norfolk Island pine), 500
Arbuscular mycorrhizae (AM), *325,* 326–327, 328, 329, 330
Arbuscules, *326,* 327
Arbutoid ectendomycorrhizae, *325,* 329
*Arbutus,* 323
*Arbutus unedo, 534*
Archaeophytes, 403
Archipelagos
biodiversity loss and, 566
defined, 431
geographic isolation and endemics, 557–559
island biogeography theory, 431–433
metapopulation theory and, 434
Arctic
effects of bird herbivory on salt marshes, 307
ozone layer and, 30–31
permafrost and global climate change, 485
plant genetic variation in responses to day length, 46–47
sea ice and global climate change, 484, *485,* 488
Arctic tundra
characteristics of, 540–541
effects of global climate change on, 491
global distribution, *514*
β-diversity, 550
*Arctostaphylos, 290*
*Arctostaphylos glauca, 534*
*Arctostaphylos manzanita* (manzanita), 323
Area hypothesis of species richness, 546
Arecaceae, 311, 556
Arecales, *40*
*Arenaria serpyllifolia* (sandwort), 387
Argentine ant (*Iridomyrmex humilis*), 548
*Argyroxiphium,* 177
*Argyroxiphium sandwicense* (Mauna Kea silversword), 220, *221*
Aridisols, *87, 88*
*Arisaema triphyllum* (jack-in-the-pulpit), 159
*Aristida longiseta,* 273
Aristotle, 14
Arrhenius, Svante, 477
*Artemisia, 509*
*Artemisia californica* (California sagebrush), 277
*Artemisia tridentata* (sagebrush), *511, 538*
*Artemisia tridentata vaseyana* (mountain big sagebrush), *61*
Ascension island, *559*
*Asclepias,* 313, *314,* 316–317
*Asclepias meadii* (Mead's milkweed), 209
*Asclepias syriaca* (common milkweed), 316–317
Ascomycota, *325*
Asexual reproduction, 149–150, 201
Ashman, Tia-Lynn, 158, 164
Asia
carbon emissions, 478, 479
continental patterns in plant biodiversity, 551–552
Asian soybean rust, 320
*Asimina, 552*
Asner, Gregory, 428–429, 568–569
Asparagales, *40*
Aspartate, *35*
Aspens, 387, 390–391, 532
Assemblages, 335
Assimilation rate, 32
Associations, 334
Assortative mating, 166–167
*Aster ericoides* (white aster), 372
Asteraceae
agamospermy, 150
convergent evolution and, 521
distribution of $C_4$ photosynthesis, *40*
fynbos, 547
number of species, 544
seed size and, 180
Asterales, *40*
*Asterogyne martiana,* 29–30
Asteroids, 503
*Astragalus,* 309
*Astragalus scaphoides* (bitterroot milkvetch), 192–193
*Astrocaryum mexicanum* (chocho), 228
Asymmetric competition, 268–269
Asymmetrical similarity, *362*
Atacama Desert, 459, 464, 465, 537, 538
Atlantic Ocean, 463
Atmosphere
carbon pool, 115–116
effects of global air movements on ocean currents, 462–464
forms of nitrogen found in, 132–133

global patterns of air movement, 457–464
greenhouse effect, 450–451 (*see also* Greenhouse effect)
oxygen levels in the Carboniferous period, 499
plant nutrients and, 115–116
radiant energy balance of the Earth and, 449
Rayleigh scattering, 77
soil–plant–atmosphere continuum of water movement, 57–59
Atmospheric carbon dioxide
anthropogenic fluxes, 473–474
Cenozoic era, 503
Cretaceous period, 501
direct effects on plants of increasing levels of, 474
effects of deforestation and land use change on, 481–482
effects on global biodiversity, *439*
evolution of photosynthesis and, 39, 41
fossil fuel combustion and, 477–481
global carbon cycle, 472–474
global climate change and (*see* Anthropogenic global climate change)
greenhouse effect and, 451 (*see also* Greenhouse effect)
historic increases in, *473*
levels over the past 400 million years, *501*
Mesozoic era, 499
Milankovitch cycles and, 456
sequestration and, 493
*See also* Greenhouse gases
Atmospheric deposition
of nitrogen, 132–135, 137
of sulfur, 132, 136
ATP synthase, *23*
*Atriplex,* 255, 256
*Atriplex canescens* (fourwing saltbush or shadscale), *40, 70,* 255, *511*
*Atriplex polycarpa* (desert saltbush), 62
Auge, Gabriela, 182
Augspurger, Carol, 171, *172,* 194, 303
Augustine, David, 306
Austin, Michael, 364
Australia
*Cactoblastis cactorum* moth and prickly pears, 303, *304*
early human impact on plant communities, 508
human migration to, 508
wildfires and global climate change, 483
*Austrocedrus chilensis* (Chilean cedar), 74
Autecology, 17–18
Autopolyploidy, 255
*Avena, 36*
*Avena fatua* (wild oats), 323–324
Avirulence genes, 321, 322
Axillary flowers, *157*
Axillary meristems, 144–145
*Azolla,* 107
*Azorella compacta* (yareta), *539*


## B


B horizons, 94, 95
Baby sun rose (*Aptenia cordifolia*), *60*
Bacon, Francis, 6
Bacteria
biological immobilization of nitrogen, 132
cellulose-digesting, 316
endophytes, 324
mineralization of nitrogen, 129–131
nitrogen fixing symbioses, 105–108
pathogenic, 318, 320
soil organic matter and, 125
Bacteroids, *106*
Baker, Herbert, 181, 404
Bald cypress (*Taxodium distichum*), *66*
Baldwin, Ian, 314
Balsam fir (*Abies balsamea*), *509,* 532
Balsam poplar (*Populus balsamifera*), 532
Balzter, Heiko, 290
Banff National Park, 365
Bangladesh, 489
*Banksia,* 175
*Banksia ericifolia* (heath-leaved banksia), *502*
Barberis, Ignacio, 273
Barbour, Michael, 62
Bark beetles, 301–302, 315–316, 324, 382
Barro Colorado Island, *172*
Bartholomew, Bruce, 277, 280
Basal angiosperms, *40*
Basal cover (basal area), 356
Bases, exchangeable, 93–94
Basidiomycota, 320–321, *325*
Bassham, James, 26
Basswood (*Tilia americana*), 375, 391, 529
Bat pollination, 158
Bay of Bengal, 489
Bazzaz, Fakhri, 188
BDFFP. *See* Biological Dynamics of Forest Fragments Project
Beach grass (*Ammophila breviligulata*), 375
*Beagle* (HMS), 14, *15*
Beal, William J., 175
Beck's water marigold (*Megalodonta beckii*), *242*
Bee pollination, 156, 169
Beeches, 507, 528, *529*
Bees, 161, 169
Beetles, 158, 316
*Begonia pavonina* (peacock begonia), *49*
Begonias, 49
Belsky, Joy, 300
Belt transects, 355
Beneficial mineral elements, *103,* 104
Benguela Current, *462*
Benítez-Malvido, Julieta, 443
Benson, Andrew, 26
Berberiodopsidales, *40*
*Berberis thunbergii* (Japanese barberry), 310
Berendse, Frank, 135
Beringia, 507, 552
*Bertholletia excelsa* (Brazil nut), 570
Bertness, Mark, 282
Beta (β) diversity
analyzing phylogenetic diversity with respect to, 347
analyzing variation among sites and, 426
biodiversity loss in grasslands, 566
definition and concept of, 345, *346*
neutral theory of biodiversity and biogeography, 562
regional and global patterns in, 550
Wog Wog Experiment and, 445
*Beta vulgaris maritima* (sea beet), 182
Beta-carotene, 24, *25*
Bet-hedging life histories, 190–191
*Betula,* 502, *509,* 532, 539
*Betula allegheniensis* (yellow birch), *339*
*Betula lenta* (black birch), 529
*Betula papyrifera* (paper birch), 532
Betulaceae, *325,* 327, 502
Bever, James, 330
Bicarbonate ions, 94, 130
*Bidens, 283*
Biennials, 183
Big bluestem (*Andropogon gerardii*), *40,* 534
Bigger, David, 310
Bignoniaceae, 165
Bigpod ceanothus (*Ceanothus megacarpus*), *121*
Big-toothed aspen (*Populus grandidentata*), 387, *388,* 390
Billings, Dwight, 540
Bimodal patch occupancy, 435–436, *437*
Bindweed, 249–250, *504*
Biocrusts, 96
Biodiversity
benefits to humans, 572–573
biodiversity metrics, 342, 349
concept of, 341–342
conservation and restoration efforts, 573
differentiation diversity, 345, *346*
ecosystem services and, 572–573
effects of invasive species on, 410–411
functional diversity, 347–349
fynbos, 547–548
geographic or environmental gradients in, 561–566
habitat fragmentation and, 439–443
inventory diversity, 342–345
loss of (*see* Biodiversity loss)

major threats to, *439*
phylogenetic diversity, 345–347
variability across the Earth, 544–546
worldwide leaf economic spectrum, 275, 276
*See also* Global biodiversity; Latitudinal diversity gradient; Species diversity
Biodiversity hotspots, 560–561
Biodiversity loss
biodiversity hotspots and, 560–561
current and likely future factors causing, *569*
ecoregion surveys, 566–567
extinction debt, 566
from human activity, 567–569, 570–572
human population growth and, 570–572
introduction, 543
threats to rare and common species, 570
Biogeochemical cycles
calcium cycle, 136–137
carbon cycle, 116–127
defined, 114
nitrogen cycle, 127–132
overview and characteristics of, 114–116
phosphorus cycle, 135–136
quantifying pools and fluxes, 114
sulfur cycle, 136
water cycle, 137–139
Biogeochemistry, 18
Biogeography, 18
Biological control, 303–305
Biological Dynamics of Forest Fragments Project (BDFFP), 219, 441–443, 444
Biological legacies, 387
Biological species concept, 254–255
Biomass
of major biome types, *524*
measuring by dry weight, 356
productivity and standing crop, 412
Biomes
abiotic factors determining, 514
alpine and arctic vegetation, 539–541
categorization of vegetation by structure and function, 514–520 (*see also* Physiognomy)
convergent evolution and, 520–523
definition and overview, 513
distribution by temperature and precipitation, *515*
dominant growth forms and general climatic conditions, *516*
grasslands, 534–536, *537*
impact of human domination, 570
major world biomes, *514*
moist tropical forests, 523–526
seasonal tropical forests and woodlands, 526–528
taiga, 532–533
temperate deciduous forest, 528–529
temperate evergreen forest, 530–531
temperate rainforest, 529–530
temperate shrubland, 533, *534*
temperate woodland, 531, *532*
Alexander von Humboldt and, 522–523
Biosphere, 113
Biotic exchange, *439*
Biotic factors
concept of, 340
habitats and, 13, 424
heterogeneity of ecological phenomena and, 13
Birch, Charles, 300
Birches, 502, *509*, 532, 539
Birds, herbivory and, 307
Bishop's goatweed (*Aegopodium podagraria*), *157*
Bison, *297*, 300
Bitcoin, 480
Bittercress, 29, 314
Bitterroot milkvetch (*Astragalus scaphoides*), 192–193
Black birch (*Betula lenta*), 529
Black cottonwood (*Populus trichocarpa*), *73*, 392
Black oak (*Quercus velutina*), 391
Black sally (*Eucalyptus stellulata*), 302
Black spruce (*Picea maritima*), 518, *519*, 532
Black willow (*Salix nigra*), *338*
Black-beard sugarbush (*Protea lepicarpodendron*), *548*
Blackberries, 150, 372
Blackbrush (*Coleogyne ramosissima*), *511*
Black-eyed Susan (*Rudbeckia hirta*), *155*
Bladder campion (*Silene latifolia*), 321, 408
Blair, Amy, 278
*Blepharis*, *40*
Blocks, 6
Blue agave (*Agave palmeri*), *43*
Blue iridescence, 49, 50
Blue light, 47
Blue wildrye (*Elymus glaucus*), 323–324
Blueberries, 93
Bluebonnet (*Lupinus texensis*), 166
Blue-eyed Mary (*Collinsia verna*), 208, 210, 222
Bluegrass (*Poa compressa*), 398
Blue-stain disease, 324
Bogs, 130–131, 134
Bolting, 144
Bond, William, 474
Bonneville Salt Flats, 511
Bonpland, Aime, 544
Booth, Rosemary, 364
Boraginaceae, *40*, 311
Boreal forest
description of, 532–533
effects of global climate change on, 491
fire, 380
physiognomy, 516
transition zones, 553–554
*See also* Taiga
Bormann, Frank, 376
Borneo, 568
Boron, *103*
Boryaceae, 66
*Botrychium australe* (adder's tongue), 308
Bottlebrush, 547–548
Bottom-up population control, 300–301
Boundary layer, 75, 78, 79
Boundary layer conductance, 60
Bounded ranges model, 562
*Bouteloua rigidiseta*, 273
Bowers, Michael, 324
Bowman, William, 286
Box, George, 205
Boyden, Suzanne, 279, *280*
Braak, Cajo ter, 364
*Brabejum stellatifolium*, *534*
Bradshaw, A. D., 241
Bradshaw, H. D. ("Toby"), 169
*Bradyrhizobium*, 106–107
*Brassica campestris*, 236
*Brassica rapa* (field mustard), 160
*Brassica septiceps* (turnip), *30*
Brassicaceae, *40*, 313–314, 325
Brassicales, *40*
Braun, E. Lucy, 515
Braun-Blanquet, Josias, 336, 338, 430
Braun-Blanquet method, *356*
Bray, J. Roger, 363
Brazil Current, *462*
Brazil nut (*Bertholletia excelsa*), 570
Brazilian Amazon. *See* Amazon Basin
British Ecological Society, 15
Brittle willow (*Salix fragilis*), 59
Brittlebush (*Encelia farinosa*), *38*, 79–80
Broadleaf lupine (*Lupinus latifolius*), 387
Broad-leaved deciduous forests, 502, 516, 517
Broad-leaved evergreen forests, 516–517
Broad-sense heritability, 239
Brokaw, Nicholas, 415
Broken tulips, 319
Bromeliads, *44*, 156, 183–184, *523*
*Bromus tectorum* (cheatgrass), 379, 408, 538
Bronstein, Judith, 279
Brooks, J. Renée, 335
Broom (*Sarothamnus scoparius*), 302
Broomrapes, 317
Broomsedge (*Andropogon virginicus*), 372
Brown, James, 306, 402, 403

Brown, Katrina, 571
Brown, Kerry, 394, 525
Brown, Valerie, 306
Brown food web, 122–125
Browsers, 298
  *See also* Herbivory
Bruchid beetles, 303
*Bruguiera gymnorrhiza* (kankra), 489
Bruna, Emilio, 218–219, 442
*Bryophyllum daigremontianum* x *B. delagoense,* 405
Bryophytes, 54
Bucci, Sandra, 63
Buckeyes, 529
Buckley, Yvonne, 219–220
Bulbils, 150
Bulbs, 150
Bulliform cells, *36*
Bumblebees, 161
Bundle sheath cells, 34, 35–36, 41
Burgman, Mark, 402–403
Burney, David, 508
Burro bush (*Ambrosia dumosa*), *61*
Bush beardgrass, 408
Busing, Richard, 415
Butt rots, 320
Buttercups, 150
Butterfield, Bradley, 567–568
Buxales, *40*
Buzz pollination, 158

## C

C horizons, 94, 95
*C* selection, 186
$C_2$ photosynthesis, 34, 41
$C_3$ photosynthesis
  CAM photosynthesis and, 38, 39
  carbon isotopes and, 37
  description of, 33–35
  evolution of, 39, 41
  evolution of $C_4$ and CAM pathways, 39, 41–42
  response to temperature, *36*
$C_3$ plants
  behavior of stomata in, 67
  carbon reduction in, 25
  Cenozoic era, 503–504
  competition among grasses along an environmental gradient, 284–286
  direct effects of increasing levels of atmospheric carbon dioxide on, 474
  early human impact on in Australia, 508
  geographic distributions, 44–46
  growth forms and habitats, 42, 43
  leaf anatomy, *30, 36, 59*
  leaf light environment and photosynthesis, 29
  phenology, 43–44
  water use efficiency, 62
$C_4$ photosynthesis
  CAM photosynthesis and, 38, 39
  carbon isotopes and, 37
  discovery and elucidation of, 26–27
  overview and description of, 33, 35–37, *38*
  photorespiration and, 34, 35
  phylogeny and evolution of, 39, *40,* 41–42
  response to temperature, *36*
$C_4$ plants
  behavior of stomata in, *67*
  competition among grasses along an environmental gradient, 284–286
  direct effects of increasing levels of atmospheric carbon dioxide on, 474
  early human impact on in Australia, 508
  evolution in the Cenozoic era, 503–504
  geographic distributions, 44–46
  growth forms and habitats, 42, 43
  phenology, 43–44
  water use efficiency, 62
Cacti (Cactaceae)
  adaptations to reduced water availability, 64–65
  CAM photosynthesis, 42
  convergent evolution and, 521–522
  effects of temperature extremes on species distributions, 81
  factors determining whole-plant temperature, 79–80
*Cactoblastis cactorum,* 303–304
Cadenasso, Mary, 440
Caffeine, 313
*Cakile edentula* (sea rocket), *234*
*Calathea ovadensis,* 158
Calcareous grasslands, 135
Calcareous soils, 92, 108
Calcic horizon, 95–97
Calcium
  global calcium cycle, 136–137
  plant nutrition, *102*
  soil organic matter and, 97
  in soils, 92, 93–94
  sources of, 116
Calcium carbonate, 95–97, 136–137
Calcium depletion, 137
Caldwell, Martyn, 31
Caliche, 95–97
California chaparral. *See* Chaparral
California Current, 462, 465
California poppy (*Eschscholzia californica*), 191
California sagebrush (*Artemisia californica*), 277
Callaway, Inderjit, 278
Callaway, Ragan, 278, 282, 287, 288
*Callistemon, 231,* 547–548
*Calluna, 290*
*Calluna vulgaris,* 134
*Calochortus,* 227
*Calochortus albus, 227*
*Calochortus gunnisonii, 227*
*Calochortus obispoensis, 227*
*Calochortus pulchellus, 227*
*Calochortus tiburonensis, 227*
Calotropagenin, *314*
Calvin, Melvin, 26
Calvin-Benson cycle
  $C_2$ photosynthesis, 34, 41
  $C_3$ photosynthesis, 33–34
  $C_4$ photosynthesis, 34, 35, 37
  CAM photosynthesis, *38,* 39
  description of, 25, *27*
  discovery and elucidation of, 26–27
  evolution of $C_4$ photosynthesis, 41–42
Calyx spurs, *154*
CAM. *See* Crassulacean acid metabolism
CAM plants
  behavior of stomata in, 67
  geographic distribution, 45, 46
  habitats and growth forms, 42–43, *44*
  succulents, 64–65
  water use efficiency, 62
  *See also* Crassulacean acid metabolism
*Cameraria, 311*
Camp Fire, 380
*Campanula americana* (tall bellflower), 220
Campanulaceae, 521
Canada, plant biodiversity, 550–551
Canada goldenrod (*Solidgao canadensis*), *61*
Canada thistle (*Cirsium arvense*), 425–426
Canadian pondweed (*Elodea canadensis*), 150
Canadian Rockies, 365, *367, 368*
Canary Current, *462*
Candelabra aloe (*Aloe arborenscens*), *43*
Cankers, 321, 322–323
*Cannabis sativa* (marijuana), *68*
Canonical correspondence analysis (CCA), 365, *366*
Canopy cover, 356
Cape Floristic Region, 547–548
Cape reeds, 547
Cape region, 559–560
  *See also* Fynbos
Cape sugarbird (*Promerops cafer*), *548*
Capillary action, 58, 99
Capitulums, 156, *157*
Carbon
  C:N ratios, 132
  factors affecting plant carbon uptake, 31–32, *33*
  global carbon cycle (*see* Carbon cycle)
  gross photosynthesis and, 28
  light compensation point and, 28–29

plant nutrition, 101, *102*
reduction (*see* Carbon reduction)
soil organic carbon, *8,* 97
*See also* Carbon isotopes
Carbon cycle
anthropogenic global climate change and, 471, 473–474
bacteria and soil organic matter, 125
carbon pools, 115–116, 119–122
effects of deforestation and land use change on, 481–482
fossil fuel combustion and, 477–481
interconnection to other biogeochemical cycles, 115
models of, 118–119
overview and description of, 472–473
productivity, 116–118, 125–127 (*see also* Productivity)
soil food webs and decomposition, 122–125
taiga and, 533
Carbon dioxide
atmospheric carbonic acid and, 132
carbon reduction in photosynthesis, 22, 24–25, 26–27
cellular respiration and, 28
eddy covariance method of quantifying, 126, 127
factors effecting plant carbon uptake, 31–32, *33*
global carbon cycle, 472–474
leaf conductance, 32
primary productivity and, 116
rubisco and, 34
soil pH and, 94
water loss through transpiration and, 60
*See also* Atmospheric carbon dioxide
Carbon fixation rate, 56
Carbon isotopes
analysis of human impact on $C_3$ and $C_4$ plants in Australia, 508
elucidation of carbon reduction and, 26
measurement of water use efficiency, 57
photosynthesis and, 37
Carbon monoxide, *118,* 476
Carbon pools, 115–116
Carbon reduction
$C_3$ photosynthesis, 33–35
$C_4$ photosynthesis, 33, 35–37, *38*
crassulacean acid metabolism, 33, 38–39
defined, 22
discovery and elucidation of, 26–27
overview of the main pathways in, 33
in photosynthesis, 24–25
Carbonate, 116
Carbonic acid, 93, 94, 132
Carboniferous period, *496,* 499, *500*
Carboxylation, 34
*Cardamine cordifolia,* 314
*Cardamine leucantha,* 29
Cardenolides, *314,* 316–317
Cardiac glycosides, 313
Cardinal flower (*Lobelia cardinalis*), *155*
Cardinale, Bradley, 572
Cardón, *222,* 279
*Carduus nutans* (musk thistle), 304, 305
*Carex, 283*
*Carex nudata,* 425–426
Carlson, Bradley, 491–492
*Carnegiea gigantea* (saguaro), 45, *64,* 81, 392, 557
Carnivorous plants, 130, *131*
Carotenes, 22
Carotenoids, 22, 24
Carousel model, 290
*Carpinus caroliniana* (American hornbeam), *339*
Carrying capacity, 185
*Carya,* 327, 372, 507, *509,* 529, *552*
Caryophyllaceae, 320–321, *325*
Caryophyllales, *40*
*Castanea,* 528
*Castanea dentata* (American chestnut), 322–323, 384, 508, 509, 529, 570
*Castilleja indivisa* (Indian paintbrush), 317, 318
Caswell, Hal, 200, 218
Catbrier (*Smilax glauca*), 316
Catechin, *312*
Categorical view of nature, 429, 430, 431
Cathedral termite (*Nasutitermes triodiae*), *537*
Cation exchange capacity (CEC), 94
Catkins, 156
Catskill Mountains, 134
Cattails, 247, 282, 283
Cattle
effects of grazing (*see* Livestock grazing)
grassland restoration and, 395
impact on tropical rainforests, 525
volatilization of ammonium, 133
Cavitation, 74
Cayuela Delgado, Luis, 428
CCA. *See* Canonical correspondence analysis
*Ceanothus,* 107
*Ceanothus greggii, 534*
*Ceanothus leucodermis, 534*
*Ceanothus megacarpus* (bigpod ceanothus), *121*
CEC. *See* Cation exchange capacity
Cecil soils, *97*
Cedar Creek Natural History Area, 376
Cedars, 507
Celastrales, *40*
Cellular automata models, 290
Cellular respiration, 28
Cement production, 479, 480
Cenophytic flora, *496*
Cenozoic era, 503–504
Censored data points, 208
Censuses, 206
*Centaurea diffusa* (diffuse knapweed), 278
*Centaurea stoebe* (spotted knapweed), 278
Central America, forest fragmentation, *441*
Central Valley (California), 535
Centrifugal theory of community organization, 282–283, 402
CENTURY model, 119
Century plant, 150, 183–184
*Cercis canadensis* (eastern redbud), 184
Cermák, Jan, 59
Cerrado, 536, 567
CFCs. *See* Chlorofluorocarbons
Chalk grasslands, 308–309
*Chamaecyparis, 552*
*Chamaenerion angustifolium* (fireweed), 387, *389*
Chamaephytes, *147,* 357
Chamise (*Adenostoma fasciculatum*), 387, *534*
Chamonix Valley, 491–492
Chance, Gleasonian ecology and, 337
Chaparral
allelopathy, 277
convergent evolution and, *534*
description of, 533
fire, 380
nutrient cycling in, *121*
Chapman Peak, *485*
Characteristic values, 213–214
Characteristic vectors, 213–214
Charnov, Eric, 164, 184
Charnov-Schaffer model of perennial life histories, 188
Charophyta, 497
Chase, Jonathan, 246
Chasmogamy, 168–169
Chazdon, Robin, 29–30, 49–50
Cheatgrass (*Bromus tectorum*), 379, 408, 538
Che-Castaldo, Judy, 229
Chemical fertilizers, 133, 134
Chemical plant defenses
constitutive and induced, 314–316
evolutionary consequences of plant-herbivore interactions, 316–317
to pathogens, 321
secondary chemicals, 312–314
Chenopodiaceae, 39, *40*
*Chenopodium album* (lamb's quarters), *504*
Cherry picking, 11–12
Chesson, Peter, 292, 294
Chestnut blight, 322–323, 384
Chestnuts, 528
*See also* American chestnut

Chihuahuan Desert
characteristics of hot deserts, 537–538
effects of herbivory on community composition, 306
global patterns of air movement and, 459
physiognomy of hot deserts and, 520
reproductive phenology of *Eriogonum abertianum, 197*
Chile, earthquakes, 382, 384
Chilean cedar (*Austrocedrus chilensis*), 74
China, carbon emissions, 478, 479, *480*
*Chionochloa pallens* (snow tussock), 193
*Chlorella pyrenoidosa, 26*
Chlorine, *103*
Chlorofluorocarbons (CFCs), 475, 476
Chlorophyll
photosynthesis, 22, 24
remote sensing and estimates of net primary production, 126
Chlorophyta, 107
Chloroplasts
$C_4$ photosynthesis, *35*
CAM photosynthesis, *38*
carbon reduction in, 24–25, *27*
in guard cells, *60*
iridoplasts, 49
light availability and photosynthesis, 29–30
light reactions of photosynthesis, 22, *23,* 24
in non-leaf structures, *31*
structure of, *23*
Cholla cacti (*Opuntia*), 201
*Chorizanthe rigida,* 174
Christensen, Norman, 221, *376*
Chromosomes
aneuploidy, 256
epigenetic effects, 245
polyploidy, 255–256
Chronic herbivory, 302
Chronosequence, 374–376
Chrysomelid beetles, 316
Chu, Chengjin, 203
*Cibotium glaucum* (hapu'u), *386*
Ciliates, 125
Cinquefoils, 150
Circular quadrats, 355
*Cirsium,* 206, *207,* 304–305
*Cirsium arvense* (Canada thistle), 425–426
*Cirsium canescens* (Platte thistle), *183,* 302, 304–305
*Cirsium centaureae,* 304
*Cirsium undulatum,* 304
*Cistus salvifolius, 534*
CITES. *See* Convention on International Trade in Endangered Species of Wild Fauna and Flora
*Citrus,* 150, 320
Citrus canker, 320
*Citrus paradisi* (grapefruit), 255–256
*Cladium, 283*
Cladograms, 346–347
*Cladonia, 290*
Clark, James, 294
*Clarkia, 256*
*Clarkia concinna* (red ribbons), 430
Clark's nutcracker, 172–173
Classification
of communities in a landscape, 366–368
of landscape ecology data, 430, 431
Clausen, Jens, 251–253
Clauss, Maria, 192
Clay, Keith, 324
Clay loam, *90*
Clay particles
adsorption of nutrient cations, 92
characteristics of, 90, *91,* 92
soil pH and, 93
soil porosity and, *92*
Clayey soils, 89, 90, 92, *101*
Clear-cutting, 384–385
Clegg, Michael, 249–250
*Cleistes bifaria* (rosebud orchid), 209
Cleistogamy, 168–169
Clements, Frederic, 274, 336–337, 338, 339, 372–373, 391, 395
Clementsian ecology, 336–337, 338–339
*Cleome, 40*
Climate
biomes and, 514, *515, 516*
Cenozoic era, 503
climate variation and the latitudinal diversity gradient, 549–550
continental-scale movement of air and water, 464–465, *466*
defined, 448
distinguished from weather, 448
El Niño Southern Oscillation, 468–470
factors affecting the magnitude of daily variation in temperature, 454–455
global climate change (*see* Anthropogenic global climate change)
global climate models, 486–487
global patterns of air movement, 457–464
Little Ice Age and modern vegetation, 510–511
long-term cycles in the Earth's orbit affecting radiant energy input, 455–457
major components of, 448
net primary production and, 117–118
during the Pleistocene glacial maximum, 507
radiant energy balance of the Earth, 448–451
seasonal variation in precipitation, 465–468
significance of temperature and precipitation predictability for plant ecology and evolution, 470–471
solar angle of incidence and, 451–452, *453*
temperature lag and, 452, 454–455
variation in, 448
Climate change
evolution of $C_4$ photosynthesis and, 41
impact on plant responses to day length, 46
Milankovitch cycles and, 456, 457
*See also* Anthropogenic global climate change
Climate diagrams, *467*
Climax, 373, 395
Clines, 241, 254
*See also* Ecotypes
Clonal fragmentation, 150
Clonal plants
agamospermy and, 150
considerations for population dynamics, 201
genets and ramets, 246–247
spatial dispersion and resource uptake, 147–149
Cloud forest, 526
*See also* Tropical montane forest
Clouds, 457
Club mosses, *497,* 498
Cluster analysis, 367, *368*
CMNs. *See* Common mycorrhizal networks
C:N ratios, 132
CO protein, 47
Coal, 477, 479, 499
*See also* Fossil fuels
Coarse scale, 423
Coast Range, *466*
Cobalt, *103*
Coca (*Erythroxylum coca*), *313*
Cocaine, 313
Cock, Matthew, 305
Cocklebur (*Xanthium strumarium*), *171*
Coco de mer (*Lodoicea maldivica*), 179
Coconut (*Cocos nucifera*), 174
Cocos island, *559*
*Cocos nucifera* (coconut), 174
Cody, Martin, 181
Coevolution
in plant-herbivore interactions, 316–317
in plant-pathogen interactions, 321–322
pollination syndromes and, 159
Coexistence theory, 290–291
*See also* Species coexistence
Coffee (*Coffea arabica*), 169, *313*
Cohesion, *93*
Cohorts, 207
Cold air drainage, 80–81
Cold deserts
characteristics of, 538
dominant growth forms and general climatic conditions, *516*

global distribution, *514*
temperature and precipitation, *515*
Cold grassland, *514*
Cold shrubland, *514*
Cold temperature
adaptations to, 80
effects on plant communities, 80–81
extremes for leaf damage, *81*
global distribution of subfreezing temperatures, *517*
Cole, Kenneth, 511
Cole, Lamont, 184
*Coleogyne ramosissima* (blackbrush), *511*
Coleoptera, 306
Cole's paradox, 184
Collemboa, *124,* 125
Collenchyma, 311
Colletxó (*Moricandia arvensis*), 41, *42*
*Colliguaya odorifera, 534*
Collins, Cathy, 445
Collins, Scott, 435–436
*Collinsia verna* (blue-eyed Mary), 208, 210, 222
Colloids, 90
Colombian bluestem, 379
Colonization
community restoration and, 395
dispersal capacity of species and, 385–387
effects of disturbance size on, 377–379
initial floristic composition theory and, 390, 393
propagule pool and, 387–389
random events affecting, 393
species interactions during succession and, 391–392
species pool and, 375
*See also* Succession
Color
animal perception of, 155
flower color and animal pollination, 154–156
Columnar cacti, 81
Colwell, Robert, 562
*Commelina communis* (dayflower), *33*
Commelinales, *40*
Commensalisms, 105, 525
Common lilac (*Syringa vulgaris*), 409
Common milkweed (*Asclepias syriaca*), 316–317
Common mullein (*Verbascum thapsus*), 175
Common mycorrhizal networks (CMNs), 281
Common plantain (*Plantago major*), 425–426
Commonness
scheme for describing, *399*
species characteristics associated with, 403
in time or space, 399, 400–401
variability with spatial extent, 417–418
Communities
abundance curves, 411–412
biodiversity and, 341–349 (*see also* Biodiversity)
causes and effects of habitat fragmentation, 439–443, 444–445
centrifugal theory of community organization, 282–283
Clements-Gleason debate, 336–338
competitive hierarchies, 275–276
Cretaceous period, 502–503
debate on community assembly rules, 433–434
debate on the reality of, 340–341
definition and concept of, 334, 335, 336
dynamic equilibrium at the landscape scale, 395–396
effects of global climate change on, 489
effects of habitat fragmentation on community assembly, 444–445
effects of herbivory on, 305–310
effects of increased diversity on productivity and stability, 416–418
effects of invasive species on, 403, 410–411
effects of nitrogen deposition on community composition, 134–135
effects of plant pathogens on, 322–324
effects of productivity on species diversity, 412–414
effects of temperature extremes on, 80–81
emergence of modern plant communities, 508–510
emergent properties and, 338–339, 340
example classification of a North American plant community, *334*
factors creating vulnerability to invasive species, 405–409
interactions of regional diversity with local diversity, 556–561
introduction to, 333–334
issues of patterns and process, 338–340
Little Ice Age and modern vegetation, 510–511
long-term studies, 357–358
measures of species richness, 349–354
metacommunity theory, 437–439
metapopulation theory, 434–436, *437*
methods of comparing, 358–361, *362*
nestedness, 444, 445
patterns of distribution across landscapes, 362–368
physiognomy, 349, 357
plant pathogen interactions with other organisms and, 324
during the Pleistocene glaciation, 507–508
sampling methods, 354–357
specialization and species diversity, 414–415
superorganism concept, 391
terminology, 334, 335
threats to rare and common species in, 570
worldwide leaf economic spectrum, 275, 276
Community assembly rules, 433–434
Community disassembly, 444
Community dynamics
Clements-Gleason dichotomy, 372–374
disturbance and, 378–379
hierarchical framework for a theory of, *372*
history of perspectives on, 373–374
Community restoration, successional processes and, 394–395
Community types, 334
COMPADRE database, 189, 219, 229
Competition
allelopathy, 276–279
among plants for pollinators, 165
among pollen grains, 165
asymmetric, 268–269
competition models and understanding species coexistence, 289–293
definition and overview, 261–262
density dependence, 200
diffuse or focused, 292
effects of mycorrhizae on, 330–331
effects of productivity gradients on, 282–288
experimental approaches for studying, 270–273
Grime's triangular model, 186, *187*
herbivory and apparent competition, 308
between individuals for limiting resources, 262–270
natural selection and, 261
between seedlings, 266–270
species coexistence and, 273
during succession, 392–393
Competition indexes, 270–271
Competitive effect, 266, 274
Competitive exclusion, 281–282
Competitive hierarchies, 275–276
Competitive response, 266, 274
Competitive strategies, 274–275
Compositional similarity, 427
Compound umbels, *157*
Concrete, 94
Condit, Richard, 550
Conduction, 75, 77, *78*
Cones, serotiny, 175, 387, 389

Conifers
  bark beetles and, 301–302, 315–316
  coniferous forests during the Pleistocene glaciation, 507
  Cretaceous period, 502
  decomposition of needles, 124
  defenses to herbivory, 315–316
  evolutionary history and the geological time scale, *496*
  evolutionary relationships of land plants, *497*
  Mesozoic era, 501
  needle-leaved evergreen forests, 517
Connectivity, 443–444
Connell, Joseph, 322, 391
Connolly, John, 276
Conservation ecology
  causes and effects of habitat fragmentation, 439–443
  connectivity, 443–444
  ecological theory and reserve design, 445–446
  edge effects, 443
  importance of population size and genetic variation, 250
  origins, 18
  restoration of biodiversity and, 573
Constant final yield, 266, *267*
Constitutive plant defenses, 314
Continental climate, 466
Continental scale, 423
Continentality, 500
Continents, patterns in plant biodiversity, 550–553
Continuous view of nature, 429, 430–431
"Control" treatments, 8–9
Controlled burns, *7*, 380
Controlled environment experiments, 6
Convection
  defined, 75
  energy flow and, 77–78
  leaf temperature and, 79
Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES), 399
Convergent evolution, 520–523, *534*
*Conyza canadensis* (horseweed), 372, 387
Cook Islands, *559*
Cool deserts. *See* Cold deserts
Cooper, William, 386
Copper, 97, *103*
Coral bean tree, 303
Cordgrass, 170, 405
Core habitats, 282, 283
Core species, 435
*Coreopsis lanceolata, 233*
Core-satellite hypothesis, 435
Coriolis effect, 460–461
Cork cambium, *145*
Cornales, *40*
*Cornus florida* (flowering dogwood), 529
*Cornus stolonifera* (red osier), 375
Correlation, 240
Correlation coefficient (*r*), 361
Correlational selection, 235
Corridors, 443, 445
Corymbs, *157*
*Coryphantha robbinsorum*
  demography of, 213
  life cycle graphs, 206
  matrix model of population growth, 212
  reproductive value, 215–216
  sensitivities and elasticities of population growth, 217, *218*
  stable growth rate, *214*
*Corythophora alta,* 442
Cosexuality, 163
*Cosmos* (von Humboldt), 523
*Cosmos bipinnatus, 157*
Coupled niche demographic models, 434
Covariance, 240
Cover, 356
Cowles, Henry C., 373, 375, 376
Cowslip (*Primula veris*), 308, *540*
Craine, Joseph, 265
Crassulaceae, 42
Crassulacean acid metabolism (CAM)
  convergent evolution and, 521–522
  discovery and elucidation of, 27
  evolution of, 39, 42
  overview and description of, 33, 38–39
  response to temperature, *36*
  *See also* CAM plants
*Crataegus,* 255
*Crataegus mongyne, 149*
Creeping bentgrass (*Agrostis stolonifera*), 148, 425–426
Creeping spearwort (*Ranunculus reptans*), 148
Creosote bush (*Larrea tridentata*), 45, *61,* 93, 177, 221
Cretaceous period
  continental land masses, 499, *500*
  dominance of the angiosperms, 501–503
  geological time scale, *496*
  K-T boundary, 503
  patterns in species richness, 502
Croll, James, *455*
Croll-Milankovitch cycles, 455–457
Crone, Elizabeth, 192–193, 229
Crops. *See* Agriculture
Crossosomatales, *40*
Crown fires, *377*, *378*
*Cryphonectria parasitica,* 322–323
Cryptic dioecy, 164, 165
Cryptic species, 255
*Cryptocarya alba, 534*
Cryptochromes, 47
Cryptogamic crust, 386
Cryptophytes, 146, *147,* 357
*C-S-R* model, 185–186, *187*
Cuba, *559*
*Cucumis sativus, 166*
*Cucurbita pepo* (zucchini), 165
Cucurbitales, *40*
Cultural ecological services, 572
Cumulative survivorship probability, 208
Currie, David, 415
Curtis, John, 355–356, 363, 565
Curtis, John T., 337
Curtis, Peter, 474
*Cuscuta europaea,* 148, *149*
Cuscutaceae, 317–318
Cushion plant (*Silene acaulis*), 147, 184
Cuticle, 54, 69–70
Cuticular conductance, 60
Cuticular resistance, 69–70
Cyanidin, *312*
Cyanobacteria
  global carbon cycle and, 473
  nitrogen fixing symbioses and, 105, 106, 107
Cyanogenic glycosides, 315
Cycads, *496, 497*
Cycles, 114
  *See also* Biogeochemical cycles
Cyclones, 381
Cymes, *157*
Cyperaceae, 39, *40, 509*
*Cypripedium,* 207
*Cypripedium acaule* (pink lady's slipper), 203–204
*Cypripedium reginae* (showy lady's slipper), 209
Cyprus, *559*
*Cystopteris bulbifera,* 150
*Cytisus scoparius* (Scotch broom), 219
Cytochrome $b_6f$, *23*


## D


*Dactylis glomerata* (orchard grass), 269
Daffodils (*Narcissus*), 150
Damping-off, 320
*Danaus,* 316–317
Dandelions
  agamospermy, 150, 255–256
  secondary succession and, 386
  seed dispersal, 171
*Danthonia spicata* (poverty grass), *234,* 235
D'Antonio, Carla, 406
Darwin, Charles
  development of plant ecology and, 14–15
  estimation of survival probability, 207
  natural selection and, 232
  *Ophrys* orchid pollination and, 159
  study of competition, 261, *262*
Data
  open science initiatives, 11–12
  statistical analyses in ecology, 9–10, 16

Daubenmire method, *356*
*Daucus carota* (wild carrot or Queen Anne's lace), 171
Dávalos, Andrea, 310
Davies, T. Jonathan, 549–550
Davis, Margaret, 338, 508, 509–510
Day length, plant responses to, 46–47
Dayflower (*Commelina communis*), *33*
DBH. *See* Diameter at breast height
de Jong, Gerdien, 179
de Kroon, Hans, 228
de Vries, Hugo, *159*
Death Valley, 538
December solstice, 451, 452, *453,* 454, 461
Deciduous forests, 507, 517–518
Decomposers, 124–125
Decomposition
  of leaves, 124, 132
  in pitcher plants, *130, 131*
  soil food webs and the carbon cycle, 122–125
Deere, John, 534
Deforestation
  anthropogenic global climate change and, 477, 480, 481–482
  biodiversity loss and, 568–569, 571
Degu (*Octodon degus*), 306
Delphinidin, *312*
*Delphinium, 154*
Demissin, *314*
Demographic life history theory, 184, 186–188
Demographic stochasticity, 224, 225
Demography
  additional approaches to modeling, 220
  basic equation describing population size, 200
  of *Coryphantha robbinsorum,* 213
  demographic processes across landscapes, 436–437
  estimating vital rates, 207–211
  heterogeneity of populations and, 220–221
  life cycle graphs, 206, *207*
  life table response experiments, 218–219
  life tables, 207, 208–209
  mark-recapture method, 209
  net reproductive rate, 215
  obtaining data for survival studies, 210
  population growth (*see* Population growth)
  population structure, 202–204
  reproductive value, 215–217
  strengths and limitations of demographic models, 228–229
  studies at larger spatial scales, 219–220
  studies of long-lived plant populations, 221–222
Dendrochronology, 376
*Dendroctonus ponderosae* (mountain pine beetle), 301–302
Denitrification, 128, 131
Density, 199, 355
Density-dependent mortality, 269–270
Density-dependent population dynamics, 199–200
Density-independent population dynamics, 199–200
Denslow, Julie, 414
*Deschampsia cespitosa* (hair grass), 286
Desert Indianwheat (*Plantago insularis*), 192
Desert mallow (*Sphaeralcea incana*), *70*
Desert mistletoe (*Phoradendron californicum*), 232
Desert plants
  adaptations for reduced water availability, 62–65
  cacti and convergent evolution, 521–522 (*see also* Cacti)
  CAM plants and, 43
  drought-deciduous phenologies, 195
  effects of temperature extremes on species distributions, 81
  environmental stochasticity and population growth in annuals, 223–224, *225*
  factors determining whole-plant temperature, 79–80
  nurse plants, 279
  pathways of water loss, *139*
  phenology of $C_3$ and $C_4$ plants, 43–44
  seed dispersal and, 174
  solar tracking, 48–49
Desert saltbush (*Atriplex polycarpa*), 62
Desertification, 96
Deserts
  cold deserts, 538 (*see also* Cold deserts)
  continental-scale air movements and, 465
  global patterns of air movement and, 459
  hot deserts, 537–538 (*see also* Hot deserts)
  physiognomy of, 519–520
  primary productivity, *524*
  rain shadows and, 464
  *See also* Chihuahuan Desert; Sonoran Desert
Determinate inflorescences, *157*
Deterministic processes, in succession, 393, 394
Detrended correspondence analysis, 362
Detritus, 122–125
Detto, Matteo, 293
Development
  developmental stages and plant demography, 202
  developmental stages in life cycle graphs, 206
  phenotypic variation and, 233
Devil's tongue (*Opuntia humifusa*), 81
Devonian period, *496,* 497–499
Dew point, 459
Dhundal (*Xylocarpus granatum*), 489
Diameter at breast height (DBH), 202, 357
*Dianthus, 68*
Diaspores, 170–175, 265
*Dicentra canadensis,* 220–221
*Dichaea pendula,* 158
Dickite, *91*
Dicots
  secondary growth, 145
  vegetative reproduction, 150
Differentiation diversity, 345, *346*
Diffuse competition, 292
Diffuse knapweed (*Centaurea diffusa*), 278
Diffusion porometers, 56
*Digitalis,* 177, 309, 313
*Digitalis purpurea, 178*
Dikes, *366*
Dilleniales, *40*
Dimension reduction, 362–363
Dioecy, 163
*Dioscorea,* 150
Dioscoreales, *40*
*Diospyros, 552*
Dipsacaceae, 320–321
Dipsacales, *40*
Diptera, 306
*Dipteryx panamensis* (almendro tree), *381*
Diptoerocarpaceae, *325,* 326
Direct gradient analysis, 363, 365
Directional selection, *234,* 235, 239
*Disa remota,* 545, *546*
Disassortative mating, 167
Diseases
  disturbance and, 384
  *See also* Plant pathogens
Disk florets, *157,* 180
Dispersal
  bet-hedging life histories and, 190
  dispersal kernels, 173–174
  effects of fruit and seed characteristics on, 170–174
  effects of human land use on, 572
  genetic variation and, 248
  genets and ramets, 201
  seed banks, 174–175
  seed size and, 181
  succession and the dispersal capacity of species, 385–387
  *See also* Migration
Dispersal kernels, 173–174
Dispersion, 348
Disruptive selection, 235
Dissected leaves, 253–254

Dissolved organic carbon (DOC), 119
Disturbance
abandoned agricultural field example, 372
animals and, 382, *384*
biological invasions and, 406
definition and overview of, 377
disease, 384 (*see also* Plant pathogens)
disturbance regime, 377
disturbance size affects which species can colonize, 377–379
earthquakes and volcanoes, 382, 384
fire, 379–381 (*see also* Fire; Forest fires)
human activities, 384–385
intermediate disturbance hypothesis, 415–416
invasive species and, 405
latitudinal diversity gradient and, 549
*r*- and *K*-strategy theory and, 185–186
species diversity and, 415–416
succession and, 371–372, 377–385
water and drought, 381–382, 383
wind, 381, *382*
Disturbance regime, 377
*Dithyrea californica* (spectacle pod), 181, 210
Diversity
alpha diversity, 344, 345, 347, 550
community stability and, 417, *418*
dominance-diversity curves, 411–412
effects of habitat fragmentation on genetic diversity, 442
effects of increased diversity on community productivity and stability, 416–418
effects on ecosystem properties, 417
epsilon diversity, 344
gamma diversity, 344, 345, 347
interactions between regional and local diversity, 556–561
phylogenetic diversity, 341, 342, 344, 345–347
point diversity, 344
trait diversity, 341–342
variability with spatial extent, 417–418
*See also* Beta diversity; Biodiversity; Global biodiversity; Latitudinal diversity gradient; Regional diversity; Species diversity
Diversity indices, 343–344
DNA bar coding, 238
DNA sequences
mutations and genetic variation, 248
using to track pollen and seeds and to identify species, 238
DOC. *See* Dissolved organic carbon
Dodders, 148, *149*, 317–318
Dodo (*Raphus cucullatus*), 171–172
Dogbane family, 159, 313
Dokuchaev, Vasily, 84
Dolomitic soils, 92
Dominance, 236, 237, 239
Dominance variance, 239
Dominance-diversity curves, 411–412
Dominant eigenvalue, 214, 222–223
Dominant eigenvectors, 222–223
Dominants (dominant species), 343, 367, 398
Domin-Krajina method, *356*
Donald, C. M., 266, *267*
Double fertilization, 151, *152*
Douglas fir (*Pseudotsuga menziesii*), *301*, 387, *511*
Downing, John, 417
Downy yellow violet (*Viola pubescens*), *28*, 29, 33, 48
*Dracena*, 145
*Dracena drago* (dragon tree), *146*
*Dracula lafleurii*, 156
Dragon tree (*Dracena drago*), *146*
*Dromaius novaehollandiae* (emu), 508
*Drosera*, *283*
Drought
adaptations to, 62–65, 65–67
as disturbance, 382, 383
El Niño Southern Oscillation and, 469, 470
global climate change and, 483, 488
Drought avoidance, 62–64
Drought-deciduous plants
defined, 63
hot deserts, 538
phenology, 195
thorn forests, 527
tropical deciduous forests, 195, 526
uncommon in tropical rainforests, 59
Drought-tolerance, 64–65
Dry adiabatic lapse rate, 459
Dry deposition of nitrogen, 132, 133
Dry weight, 102–103, 356
*Dryas drummondii*, 386, 392
*Dryas octopetala*, 541
*Drypetes glauca*, *123*
Duke Forest, 221, 376
Dukes, Jeffrey, 410
*Dulichium*, *284*
Dune communities
lakeshore dune communities, 375, 392, 396
sand dune communities, 324
Dune willow (*Salix glaucophjylloides*), 375
Dung beetles, 124
Dust Bowl, 382, 383
Dust storms, *96*
Dutch elm disease, 324
Dynamic stability, 417

## E

Earth
climate of (*see* Climate)
greenhouse effect, 450–451
long-term orbit cycles affecting radiant energy input, 455–457
radiant energy balance, 448–451
solar angle of incidence and seasonality, 451–452, *453*
Earthquakes, 382, 384
Earthworms, 124–125, 310
Eastern Cape region, 560
Eastern hemlock (*Tsuga canadensis*), 309, *339*, 507, 509–510, 529
Eastern redbud (*Cercis canadensis*), 184
Eastern white pine (*Pinus strobus*)
annual growth rings, *376*
common mycorrhizal networks, 281
Hartwick Pines State Park, *531*
pine pollen, 505
succession and, *388*
tree migration following the Pleistocene glaciation and, 508, 509, *510*
Eccentricity, *455*, 456
*Echinocactus grusonii* (golden barrel cactus), *43*
*Echinopsis chilensis*, 317
Eckert, Christopher, 167–168
ECM. *See* Ectomycorrhizae
Ecological gradients
altitudinal gradient in species frequencies, 337–338, *339*
direct gradient analysis, 363, 365
effects of productivity gradients on competition and facilitation, 282–288
indirect gradient analysis, 363–364
pattern diversity along, 564
stress-gradient hypothesis, 287
*See also* Latitudinal diversity gradient
Ecological network theory, 294
*Ecological Niches* (Chase & Leibold), 246
Ecological restoration. *See* Restoration ecology
Ecological Society of America, 15
Ecological specialists, 402
*Ecological Theater and the Evolutionary Play, The* (Hutchinson), 231
Ecology
basic principles of, 16–17
"control" treatments, 8–9
defined, 1
ethics and, 11
experiments, 5–8
features of scientific research, 5
heterogeneity of ecological phenomena, 12–13
internal and external consistency in science, 12
modern trends in, 18
nature of science and, 2
open science initiatives, 11–12
research syntheses, systematic reviews, and meta-analyses, 289
science as a human endeavor and, 18
scientific knowledge, 2–4
specific results and general understanding, 12

statistics and, 9–10
subdisciplines, 17–18
testing of theories, 10–12
Ecoregions, 566–567
Ecosystem, defined, 113
Ecosystem concept, 113
Ecosystem models, 118–119
Ecosystem processes
carbon cycle, 116–127
introduction, 111
nitrogen cycle, 127–132
nitrogen deposition and, 132–135
overview, 112–114
phosphorus, sulfur, and calcium cycles, 135–137
pools, fluxes, and biogeochemical cycles, 114–116
water cycle, 137–139
Ecosystem services, 572–573
Ecotypes, 250–254
*See also* Clines
Ectomycorrhizae (ECM), 281, *325,* 326, 327–328, 329
Eddy covariance method, 126, 127
Edge effects, 443
Edges, between patches, 429
Egg cell, 151, *152*
Egler, Frank, 390
Ehleringer, James, 48–49, 79–80
Ehrlich, Paul, 571
*Eichhornia crassipes* (water hyacinth), 150
Eigenvalues, 213–214, 222–223
Eigenvectors, 213–214, 216, 222–223
Ekoko, François, 571
El Niño Southern Oscillation (ENSO), 426, 468–470, 487, 568
Elaiosomes, 173, 548
Elasticity of population growth, 212–213, 217–218, 219
Electromagnetic spectrum, 427–429
*Eleocharis calva, 283*
Elevation
altitudinal gradient in species frequencies, 337–338, *339*
climate and, 455
geographic distributions of $C_3$ and $C_4$ plants and, 45
temperature and, 455
Elfin forest, 526
*See also* Tropical montane forest
Elk, 306–307
Ellis, Erle, 570
Ellison, Aaron, 268
Ellner, Stephen, 211, 292, 294
Elms, 502, 508, *509, 510,* 528
*Elodea canadensis* (Canadian pondweed), 150
Elton, Charles, 404, 406
Eluviation, 95
*Elymus canadensis* (wild rye), 375, 392, 396
*Elymus glaucus* (blue wildrye), 323–324
Emergent properties, 338–339, 340
Emergent trees, 523
Emery, Nancy, 228–229
Emissivity, 76
Empty niches, 405
Emu (*Dromaius novaehollandiae*), 508
*Encelia farinosa* (brittlebush), *38,* 79–80
Endangered species
concept of rarity and, 398–399
global climate change and, 492
Endemic species, 398, 557–561
Endomycorrhizae, 328–329
Endophytes, 324–325
Endosperm, 151, *152*
Energy balance, 75
*See also* Thermal energy balance
Energy budget, 75
*See also* Thermal energy balance
Energy industry, carbon emissions and, *479*
Engler, Adolf, 547
English ivy (*Hedera helix*), *314*
English ryegrass (*Lolium perenne*), 290, 398
Ensembles, 335
ENSO. *See* El Niño Southern Oscillation
Entisols, *87, 88*
Environment
effects of year-to-year variation on life histories, 189–193
genotype-environment interactions, 239–240
interaction of scale with environmental heterogeneity, 424–425
plant adaptations to terrestrial environments, 54–55
plant adaptations to xeric environments, 62–65
Environmental control hypothesis, 402–403
Environmental stochasticity
defined, 223
population growth rates and, 223–224, 225–226, 227–228
Eocene epoch, 41, *496, 500,* 503
*Eperua falcata,* 99
*Ephedra viridis* (Mormon tea), *511*
Epidermis, *70*
Epigenetic effects, 245
Epiphytes
CAM photosynthesis, 42–43, *44*
defined, 42
tropical montane forest, 526
tropical rainforest, *523,* 525
Epistasis, 239
Epistatic variance, 239
Epsilon (ε) diversity, 344
Equability, 348–349
Equalizing mechanisms, 291
Equatorial Countercurrent, *462*
Equinoxes, 451, *452, 453,* 461
*Equisetum, 259*
Equitaceae, 311
*Erica,* 547
*Erica arborea, 534*
*Erica mammosa, 328*
*Erica tetralix,* 135
Ericaceae, 93, 328–329, 547, 555, 559
Ericales, *40, 325,* 328–329
Ericoid mycorrhizae, *325, 326,* 328–329
*Erigeron heteromorphus* (fleabane), *242*
*Eriocaulon, 284*
*Eriogonum abertianum* (Abert's buckwheat), 195, 196, *197*
*Eriophorum vaginatum,* 541
*Eriophyllum lanatum, 361*
*Eriophyllum lanosum, 223, 224, 225*
*Erodium cicutarium, 223, 224, 225*
*Erodium texanum, 223, 224, 225*
Erosion, 96
Errors of development, 233
*Eryngium cuneifolium* (wedgeleaf eryngium), 228
*Erysimum,* 160
*Erythrina,* 303
*Erythronium grandiflorum* (glacier lily), 164, *165*
*Erythroxylum coca* (coca), *313*
*Eschscholzia californica* (California poppy), 191
*Espeletia, 555*
*Espeletia pycnophylla,* 521
Estuaries, 112
Ethanolic fermentation, 67
Ethics, 11
*Eucalyptus,* 313, 320, 387, 521
*Eucalyptus* forests, 379, 380
*Eucalyptus pauciflora* (snow gum), 302
*Eucalyptus saligna* (Sydney bluegum), 279, *280*
*Eucalyptus stellulata* (black sally), 302
Euclidean distance, 360–361
*Eucrypta micrantha, 224*
Eudicots
phylogeny of $C_4$ photosynthesis, *40*
xylem vessel elements, 73
Euglossidae, 158
*Euphorbia,* 42
*Euphorbia enopla,* 522
*Euphorbia pulcherrima* (wild poinsettia), *40,* 156, 522
Euphorbiaceae, *40,* 42, 311, 313
Euphyllophytes, *497*
Eurasia
Mesozoic era, 499, *500*
Pleistocene glaciation, *507*
Europe
carbon emissions, 478, 479
continental patterns in plant biodiversity, 551, 552
land use and anthropogenic carbon emissions, 482

European honeysuckle (*Lonicera periclymenum*), *299*
European rabbit (*Oryctolagus cuniculus*), 308–309
European spruce bark beetle, 301
European Union, carbon emissions, 478, *480*
European Vegetation Survey, 341
Evans, Margaret E. K., 209–210
Evaporation
  latent heat loss and, 78
  water cycle and, 138–139
Evapotranspiration, 138–139
*Evax verna, 223, 225*
Evenness, 343–344, *345*
Evergreens
  adaptations to reduced water availability, 63
  broad-leaved evergreen forest, 516–517
  mixed deciduous-evergreen forest, 528
  needle-leaved evergreen forest, 517, 518, 531
  *See also* Conifers; Temperate evergreen forest
Evolution
  ecotypes, 250–254
  genetic variation and, 248–250
  heritability and the genetic basis of phenotypic variation, 235–240
  life histories and, 181–184
  natural selection and evolutionary change, 232–235
  natural selection and patterns of adaptation, 240–246
  natural selection and the origin of new species, 254–256
  other factors causing evolutionary change, 232
  problem of Panglossian assumptions and adaptive speculations, 231–232
  seed banks and, 175
  variation among populations, 250
Exchangeable bases, 93–94
*Excoecaria agallocha* (genwa), 489
Experiments
  approaches for studying competition, 270–273
  "control" treatments, 8–9
  hypothesis testing, 4
  manipulative experiments, 5–8
  natural experiments, 8
  observational experiments, 8
  randomized experiments, 6, 9
  two-species substitutive experiments, 276
Exponential population growth, 205
Extent, scale and, 422, 423
Extinction debt, 566, 568
Extinction probability, 227
Extinction rate
  island biogeography theory and, 432
  metapopulation theory and, 434, 435, 436
  SLOSS debate and, 446
Extinctions
  global warming and, 511
  human-caused, 508
  K-T boundary, 503
  *See also* Local extinctions; Species extinctions

## F

Fabaceae. *See* Legumes
Fabales, *40*
FACE experiments, 474
Facelli, Evelina and José, 330
Facilitation
  of colonization during succession, 391, 392
  defined, 262, 273–274, 279
  discussion of, 279–281
  effects of productivity gradients on, 282–286
Facultative symbioses, 105
Fagaceae, 312–313, *325,* 327
Fagales, *40*
*Fagus,* 507
*Fagus grandifolia* (American beech), *339,* 375, *384,* 509, *510, 529*
*Fagus sylvatica, 47, 48*
Fahrig, Lenore, 434, 439–440
Faith's phylogenetic diversity, 346–347
*Falcataria moluccana* (Moluccan albizia), 279, *280*
Falkland Current, *462*
*Fallopia japonica* (Japanese knotweed), 405
False Solomon's seal (*Maianthemum japonicum*), 29
Falsification, 10
Fan designs, 272
Fang, Changxun, 98
Fanmustard (*Nerisyrenai camporum*), *70*
Far-red light, plant sensitivity to day length and, 47
Fast-slow continuum, 188–189
Fauth, John, 334
Favonol, *312*
Fecundity
  defined, 211, 234
  fitness and natural selection, 234
  population growth models and, 211–212
Female choice, 165
Female fitness, 179–181
Fence lines, 309
Feng, Gang, 557
Fens, 130–131
Ferns
  Devonian period, 498
  evolutionary relationships of land plants, *497*
  Mesozoic era, 499–500, 501
  polyploidy, 255
  vegetative reproduction, 150
Ferrel cells, *458,* 461, 462
Fertility
  defined, 211
  demographic life history theory, 184
  fitness and natural selection, 234
  population growth models and, 211–212
*Festuca rubra* ssp. *arenaria* (sand fescue), 324
Fibers, xeromorphic leaves, *70*
Fibrous root systems, 71, *72*
*Ficus,* 159, 264, *265,* 525
Field capacity, 99, *100*
Field experiments, 6
Field mustard (*Brassica rapa*), 160
Figs, 159
Fiji, *559*
Filaments, 151, *152, 154*
Fine scale, 423
Fire
  biological invasions and, 408
  controlled burns, *7,* 380
  as disturbance, 379–381
  effects on root suckering in aspens, 387
  forest burning and anthropogenic fluxes of carbon dioxide, 473–474, 482
  global climate change and, 483
  grasslands and, 519
  intensity of, 377, *378*
  landscape ecology and, 422
  nitrogen loss and, 131
  seed germination and, 192
  serotiny and, 387, 389
  temperate grassland and, 534, 535
  tree bark and, 391
  *See also* Forest fires
Fireweed (*Chamaenerion angustifolium*), 387, *389*
Firs, 532
Fischer, Murkus, 148
Fisher, R. T., 373
Fisher, Ronald A., 6
Fish-plant interactions, 160
Fitchner, Andreas, 279
Fitness
  concept of, 181
  effects of among-year demographic variation on, 189–190
  effects of herbivory on, *318*
  evolution of gender and, 164
  frequency-dependent selection, 167–168
  natural selection and fitness differences, 234–235
Flannelflower (*Actinotus helianthi*), 228–229
*Flaveria,* 34, *40,* 41

Flavone, *312*
Flavonoids, 50, *312,* 313
Flea beetles, 160
Fleabane (*Erigeron heteromorphus*), *242*
Floods and flooding
  adaptations to, 66–67
  as disturbance, 381
  *See also* Waterlogged soils
Floral bracts, 156
Floral odors, 156, 158
Florets, *154*
Florida dollar orchid (*Prosthechea boothiana*), *44*
Floristic-sociological school, 336, 338, 367, 374, 431
Flowerhead weevil (*Rhinocyllus conicus*), 304
Flowering
  overview, 144–145
  reproductive phenologies, 195–197
  synchronized flowering and masting, 192–193
Flowering dogwood (*Cornus florida*), 529
Flowers
  angiosperm sexual life cycle, 151–153
  aquatic plants, 161–162
  assortative mating and, 166–167
  evolution in the Cretaceous period, 502
  floral odors and acoustic guides for pollinators, 156, 158
  inflorescences, 156, *157*
  limiting unwanted visits to, 158
  pathogens, 320–321
  plant gender expression and, 163–165
  sexual selection and, 165
  specialized flowers and pollinators, 159
  structure and pollination, *154*
  visual displays attracting animal visitors, 154–156, *157*
Flux equations, 32
Flux towers, *56*
Fluxes
  biogeochemical cycles and, 114–116
  defined, 114
  quantifying, 114
  turnover time and, 120–121
FLUXNET, 126
Fly pollination, 159
Foamflower (*Tiarella cordifolia*), *28,* 29, 33, 48
Focus, scale and, 422, *423*
Fog, 457
Foote, Eunice, 477
Foothill palo verde (*Parkinsonia microphylla*), *31, 65,* 392
Foraging, clonal growth as, 148, *149*
Forbs, 44, 323
Forest, Félix, 560
Forest drying, 443
Forest fires
  anthropogenic fluxes of carbon dioxide and, 473–474, 482
  as disturbance, 377, *378,* 379, 380–381
  forest management and, 380–381
  human activity and, 384
  intensity of, 377, *378,* 379
  *See also* Fire
Forest fragmentation
  anthropogenic carbon emissions and, 482
  Biological Dynamics of Forest Fragments Project, 441–443, 444
  connectivity, 443–444
  edge effects, 443
Forests
  calcium depletion, 137
  chronosequence, 376
  competition for light during succession, 393
  disturbance and species diversity, 415
  effects of acid precipitation on, 132, *133*
  effects of herbivory on trees, 301–302
  effects of human activity on, 384–385
  effects of nitrogen deposition on, 134
  effects of regeneration on anthropogenic carbon dioxide emissions, 480–482
  fire and forest management, 380–381 (*see also* Forest fires)
  R. T. Fisher's view of, 373
  gap size and species richness, 378
  human impact on the carbon cycle and, 481–482
  land use histories and succession, 394, 395
  methods of studying competition in, 273
  mitigation of global climate change and, 492
  physiognomy, 516–518
  Pleistocene glaciation and, 507
  propagule pool and secondary succession, 387
  spatial patterns and ecological processes, 430, *431*
  tree line, 518
  wind as disturbance, 381, 382
  *See also individual forest types*
Forget-me-not, *264,* 268–269
Forman, Richard T. T., 422
Formations, 334
Forseth, Irwin, 48–49
Fossil fuels, 473–474, 477–481
Fossils, methods of studying, 504–505
Foster, Christopher, 422
Foster, Tammy, 335
*Fouquieria splendens* (ocotillo), *63,* 195
Fourier, Joseph, 477
Fourwing saltbush (*Atriplex canescens*), *40, 70,* 255, *511*
Fowler, Norma, 273
Fox, Gordon, 214, *215*
Foxgloves, 177, 309, 313
Fractal dimension, 429
*Fragaria,* 150, 247
Fragmentation, dispersal by, 201
*Frailea,* 169
Frailejones, *555*
Frank, Douglas, 306
*Frankia,* 107
Franklin, Benjamin, 15
Franson, Susan, 171, *172*
Franz Josef Glacier, *385*
Fraterrigo, Jennifer, 395
*Fraxinus, 509*
*Fraxinus americana* (white ash), *339*
Freckleton, Robert, 283
Free air $CO_2$ enrichment (FACE) experiments, 474
Freezing tolerance, 81
Frequency
  of a disturbance, 377
  of species in a quadrat, 355
Frequency-dependent selection, 167–168
Fretwell, Stephen, 180
Frost pockets, 80–81
Frostweed (*Helianthemum chamaecistus*), 308
Frugivores, 170–172, 197, 298
Fruits
  angiosperm sexual life cycle, 151, *152,* 153
  chloroplasts and, *31*
  evolution in the Cretaceous period, 502
  major functions of, 170
  pathogens, 320
  seed dispersal and, 170–174
  serotiny, 175
Fuel load, 379, 380–381
Fukami, Tadashi, 394
Fulvic acids, *97*
Functional diversity, 342, 344, 347–349
Functional ecology, 21
Functional group diversity, 417
Functional groups, 335, 348
Fundamental niche, 246, 402
Fungi
  endophytes, 324–325
  interactions with other organisms, 324
  mycorrhizae, 325–331 (*see also* Mycorrhizae)
  pathogenic, 318, 319–321
  saprophytic, 125
Funk, Jennifer, 30
*Fusarium subglutinans* f. sp. *Pini,* 399
Fynbos, *543*
  characteristics of, 533
  convergent evolution and, *534*

plant diversity, 547–548
species diversity and endemism, 559–560
Fynn, Richard, 274–275

## G

Galápagos Archipelago, *559*
*Galerucella calmariensis. See* Neogalerucella calmariensis
*Galerucella pusilla. See* Neogalerucella pusilla
*Galium ambiguum, 361*
*Galium saxatile,* 281
*Galium sylvestre* (*G. sterneri*), 281
Gallotannin, *312*
Galls, 298, *299*
Gametophytes
alternation of generations, 150–151, *152,* 153
early land plants, 498
plant adaptations to terrestrial environments and, 54, 55
Gametophytic SI, 162
Gamma (γ) diversity, 344, 345, 347
Gange, Alan, 306
Gap creation, 273
Gaps
caused by windthrows, 382
colonization in secondary succession, 385–386
disturbance and colonizing species, 377–378
species diversity in forests and, 415
Garden competition experiments, 272–273
Garden experiments, 6
Garden pea (*Pisum sativa*), 237
Garlic mustard (*Alliaria petiolata*), 310
Garrigue, 533
Garry oak (*Quercus garryana*), *504*
Garryales, *40*
Gasser, Patrick, 124–125
Gaston, Kevin, 403
Gelisols, 88–89
Gender expression, 163–165
Gene frequencies, genetic drift, 248–250
Gene-environment covariance, 240
General circulation models, 486–487
Generalist herbivores, 305
Generalist species, 245–246
Generative cell, 151
Genes
heritability and, 235–237
mutations and genetic variation, 248
using to track pollen and seeds and to identify species, 238
Genetic diversity, effects of habitat fragmentation on, 442
Genetic drift, 232, 248–250
Genetic response, 235
Genetic variation
evolution and, 248–250
genotype-environment interactions and, 239–240
significance for conservation, 250
strict additivity, 237
Genetically modified (GM) crops, 169–170
Genetics
genes and heritability, 235–237
genotype-environment interactions, 239–240
plant ecology and, 16
using genes to track pollen and seeds and to identify species, 238
Genets, 201, 246–247
*See also* Clonal plants
Genome, 233, 239–240
Genotype
defined, 233
genotype-environment interactions, 239–240
nonrandom distribution among environments, 240
systems of strict additivity and dominance, 237
Genotype-environment interactions, 239–240
Gentianales, *40*
Genwa (*Excoecaria agallocha*), 489
*Genyornis newtonii,* 508
Geoengineering, 492–493
Geographic distribution
abundance and, 400
$C_3$ and $C_4$ plants, 44–46
CAM plants, 45, 46
effects of herbivory on, 302–303
*See also* Species distribution
Geographic information systems (GIS), 426–427
Geographic isolation, 557–559
Geological time scale, *496*
Geometric mean, 226
Geometric population growth, 205
*Geonoma cuneata,* 29–30
Geostatistics, 427
Geraniales, *40*
*Geranium,* 241
*Geranium sanguineum,* 254
Germany, carbon emissions, 479
Germination. *See* Seed germination
GHGs. *See* Greenhouse gases
Ghost moth (*Phymatopus californicus*), 303
Giant horsetails, *498*
Giant sequoia (*Sequoiadendron giganteum*), *58*
Gigasporaceae, *325*
Gillman, Len, 563
Ginkgo (*Ginkgo biloba*), 163, 500, 502, 544
Ginkgos (Ginkgoaceae), *496, 497,* 502, 544
Giraffes, 316
GIS. *See* Geographic information systems
Gisekiaceae, 39
Glacial refuges, 507–508
Glaciations
Little Ice Age and modern vegetation, 510–511
vegetation patterns and Pleistocene glaciations, 505–508
*See also* Glaciers
Glacier Bay, 386, 392, *541*
Glacier lily (*Erythronium grandiflorum*), 164, *165*
Glacier National Park, *485*
Glaciers
continental patterns of biodiversity and, 551, *552*
global climate change and, 484–485, 488, 491–492
*See also* Glaciations
Glasswort (*Salicornia europea*), *62,* 268
Gleason, Henry A., 336, 337–338, 339, 372–373
Gleasonian ecology, 336, 337–338, 339, 340
*Glechoma hederacea* (ground ivy), 148, 290
Glenn, Susan, 435–436
Global biodiversity
continental patterns, 550–553
explanations for patterns in, 546–550, 561–566
interactions between regional and local diversity, 556–561
introduction, 543
latitudinal diversity gradient, 545–546
variability in, 544–546
Global climate change. *See* Anthropogenic global climate change
Global climate models, 486–487
Global human footprint, 571
Global Positioning System (GPS), 368
Global scale, 423
Global warming. *See* Anthropogenic global climate change
Glomeraceae, *325*
*Glomeris marginata, 124*
Glomeromycota, *325*
Glucosinolates, 313–314
Glycerate, *38*
Glycine, 41
*Glycine max* (soybean), *106*
Glycolysis, 67
GM crops. *See* Genetically modified crops
*Gnaphalium purpureum* (purple cudweed), 372
Gnetophytes, *497*
Gobi Desert, 464, 538
Goldberg, Deborah, 221–222, 266, 272–273, 283, 286–287, 398
Goldberg, Emma, 164
Golden banner (*Thermopsis montana*), *47*

Golden barrel cactus (*Echinocactus grusonii*), *43*
Goldenrods, *310*
Goldstein, Guillermo, 81
Gondwana, 499, *500,* 501
Goodall, David, 363
Gorse (*Ulex europaeus*), *410*
Gotelli, Nicholas, 435
Gould, Stephen Jay, 231
GPP. *See* Gross primary production
GPS. *See* Global Positioning System
Grace, James, 413
Grain (minimum resolution of data), 422, 423
Grains (fruit of grasses), 170
Grana stacks, 22, *23,* 50
Grand Canyon, 511
Granivores, 172–173, 192, 197, 538
Granivory, 181, 302–303
Grapefruit (*Citrus paradisi*), 255–256
Grasses
  adaptations to reduced water availability, 64
  agamospermy, 150
  $C_3$ photosynthesis, 44–46
  $C_4$ photosynthesis, 42, 44–46
  Cenozoic era, 503–504
  as clonal plants, 246
  competition along an environmental gradient, 284–286
  competitive strategies, 274–275
  direct effects of increasing levels of atmospheric carbon dioxide on, 474
  effects of mycorrhizae on competitive interactions, 330
  endophytes, 324
  evolutionary consequences of plant-herbivore interactions, 316
  geographic distributions of $C_3$ and $C_4$ grasses, 44–46
  heavy-metal tolerance, 241–243
  leaf rolling in response to water stress, 80
  physical defenses to herbivory, 311, 314
  succession in dune communities, 375
  wind pollination, 153
  *See also* Poaceae
Grasslands
  alkaline soils and, 93
  alpine grassland, *515, 516*, 539, *540*
  $C_4$ photosynthesis and, 42
  calcareous grasslands, 135
  carbon pools, 120, 121, *122*
  Cenozoic era, 503–504
  chalk grasslands, 308–309
  cold grassland, *514*
  competitive dominants, 398
  direct effects of increasing levels of atmospheric carbon dioxide on, 474
  effects of diversity on ecosystem properties, 417
  effects of invasive species of grass diseases, 323–324
  effects of livestock grazing on, 385, 569
  effects of mycorrhizae on competitive interactions, 330
  factors affecting biological invasions, 406, *407,* 408
  pathways of water loss, *139*
  physiognomy, 518–519
  rain shadows and, 470–471
  restoration after cattle grazing, 395
  temperate grassland, 534–535 (*see also* Prairies; Temperate grassland)
  tropical savanna, 536, *537*
Gravitational potential, 55, 57
Gray squirrel (*Sciurus carolinensis*), 303
Grazers, 298
  *See also* Cattle; Herbivores; Livestock grazing
Great Basin, 465, *466,* 511, 538
Great Lakes, 465
Great Plains, 383, *466*
Great Salt Lake, 511
Great Salt Lake basin, *466*
Great Sandy Desert, 459
Great Smoky Mountains, 134, 337–338, *339*
Great Victoria Desert, 459
Great white trillium (*Trillium grandiflorum*), 204, 221, 570
Green algae, 107, 497
Green Revolution, 105, 134
Green River basin, *466*
Greenhouse competition experiments, 271–272
Greenhouse effect
  Cretaceous period, 501
  fossil fuel combustion and, 477–481
  human-created greenhouse gases and, 475, 476
  Milankovitch cycles and, 456
  overview and description of, 450–451
  scientific recognition of, 476–477
Greenhouse gases (GHGs)
  agriculture as a source of, 482
  defined, 450
  efforts to reduce levels of, 492
  fossil fuel combustion and, 477–481
  greenhouse effect and, 450–451, 475–477
  nitrous oxide, 131
  radiant energy balance of the Earth and, 449
  *See also* Atmospheric carbon dioxide
Gregg, Katharine, 209
Gremer, Jennifer, 190
Grime, J. Philip, 185–186, 274, 275, 282, 288, 364, 413–414
Grime's triangular model, 185–186, *187*
Groom, Martha, 430
Gross photosynthesis, 28
Gross photosynthetic production, 116
Gross primary production (GPP)
  defined, 116
  global carbon cycle, 473
  net ecosystem production and, 119
Ground ivy (*Glechoma hederacea*), 148, 290
Ground-truthing, 368, 429
Groundwater, 93
Grubb, Peter, 265
Guard cells
  chloroplasts, *60*
  defined, 67
  mesophyte leaves, *68*
  responses to water availability, 67–68
  stimuli inducing the opening and closing of, 67
Guayacan trumpet tree (*Tabebuia guayacan*), *63*
Guerrilla growth, 148–149
Guilds, 335
Guinea grass (*Panicum maximum*), 274–275
Gulf Stream, 462, 463
*Gunnera insignis* (poor man's umbrella), *79*
Gunnerales, *40*
Gunter, Nicole, 124
Gurevitch, Jessica, 214, 253, 268–269, 272, 284–286, 394, 409, 437, 525
Gutiérrez, Julio, 306
*Guzmania monostachya, 44*
Gymnosperms
  adaptations to terrestrial environments, 54
  coal deposits and, 499
  dominance in the Mesozoic era, 499–501
  evolutionary history and the geological time scale, *496*
  evolutionary relationships of land plants, *497*
  forms of gender expression, 163
  mycorrhizae and, 325, 327 (*see also* Mycorrhizae)
  number of known species, 544
  Pleistocene extinctions, *552*
  sexual life cycle, 153
  tracheids, 73
Gynodioecy, 163, 164
Gynomonoecy, 163
Gypsum, 136
Gypsy moth (*Lymantria dispar*), 301, 382, *384*
Gyres, 462


## H


Haber, Fritz, 134
Haber process, 134
Habitat construction, 246
Habitat destruction
  biodiversity loss and, 568–569

impact on tropical rainforests, 525–526
Habitat fragmentation
causes and effects of, 439–443
connectivity, 443–444
edge effects, 443
effects on community assembly and succession, 444–445
global climate change and, 490–491
inbreeding depression and, 169
nestedness of communities and, 444
pollinator availability and, 169
Habitats
centrifugal theory of community organization, 282–283
core habitats, 282, 283
defined, 13, 424
effects of global climate change on, 490–491
environmental heterogeneity and, 424, *425*
heterogeneity of ecological phenomena and, 13
microhabitats, 13, 424, *425*
peripheral habitats, 282–283
during the Pleistocene glacial maximum, 507–508
Hadley cells, 458–461, 462, 464, 527
Haeckel, Ernst, 15
Hair grass (*Deschampsia cespitosa*), 286
Hairs, xeromorphic leaves, *70*
Hairston, Nelson, 297, 300
*Hakea,* 175
*Hakea sericea* (silky hakea), 548
Hale, Alison, 278–279
*Halesia monticola* (mountain silverbell), *339*
Halogenated hydrocarbons, 476
Halophytes, 62, 66
Halsey, Richard, 277
*Handroanthus guayacan,* 209
Hanski, Ilkka, 434
*Haplopappus squarrosus. See* Sawtooth goldenbush
Hapu'u (*Cibotium glaucum*), *386*
Harms, Kyle, 556
Harper, John, 17, 180, 200, 246, 269
Hartig net, *326,* 327
Hartwick Pines State Park, *531*
Hastings, Rodney, 222
Hatch, Marshall, 27
Hättenschwiler, Stephan, 124–125
Hawaii arrowhead (*Sagittaria sagittifolia*), *242*
Hawaii Volcanoes National Park, 30, 407
Hawaiian Islands, 464, 557, *559,* 567
Hawaiian silverswords, 177
Hawkweeds, 150
Hawthorns, *149,* 255
*Hazardia squarrosa* (sawtooth goldenbush), 302
Heads, 156, *157*
Heat
defined, 448
energy balance of the Earth and, 448–451
Heath
acidic soils, 93
definition and description of, 539
effects of global climate change on, 491–492
effects of nitrogen deposition on, 134, 135
Heath-leaved banksia (*Banksia ericifolia*), *502*
Heavy-metal tolerance, 241–243, 329, 521
Hector, Andy, 416
*Hedera helix* (English ivy), *314*
Hederagenin, *314*
*Hedychium gardnerianum,* 150
Heegaard, Einar, 491
Heimpel, George, 305
Helianieae, *40*
*Helianthemum chamaecistus* (frostweed), 308
*Helianthus annuus,* 180, 257
*Helianthus anomalus,* 257
*Helianthus deserticola,* 257
*Helianthus paradoxus,* 257
*Helianthus petiolarus,* 257
*Heliconia acuminata,* 218–219, 442
*Heliotropium, 40*
Hemicryptophytes, 146, *147,* 357
Hemiparasites, 317
Hemiptera, 306
Hemlock woolly adelgid (*Adelges tsugae*), 309
Hemlocks. *See* Eastern hemlock
Hemlock–white pine–northern hardwood forest, 528, 553
Herbaceous plants
adaptations to reduced water availability, 63–65
vegetative phenology in temperate deciduous forests, 194–195
Herbaceous seepweed (*Suaeda maritima*), 112
Herbivores
defined, 297
evolutionary consequences of plant-herbivore interactions, 316–317
generalist and specialist, 305
insect herbivores in biological control, 303–305
*See also* Cattle; Livestock grazing
Herbivory
biological invasions and, 408–409
chronic herbivory, 302
disturbance and, 382, *384*
effects on individual plants, 298–300
effects on plant communities, 305–310
effects on population composition and dynamics, 300–305
evolutionary consequences of plant-herbivore interactions, 316–317, *318*
introduction and overview, 297, 298
overcompensation and, 299–300
parasitic plants and, 318
plant adaptive responses to, 298–299
plant defenses against, 310–317
plant reproductive phenologies and, 197
*See also* Livestock grazing
Heritability
definition and concept of, 235–237
genetic and nongenetic components of phenotypic variation, 238–239
genotype-environment interactions, 239–240
natural selection and, 234–235
strict additivity, 237
Hermaphroditism, 163–164
Herrera, Carlos, 192
*Hesperis,* 256
*Hesperolinon,* 169
*Hesperostipa neomexicana,* 284–286
*Heterobasidion annosum,* 320
Heterocysts, *106*
Heterogeneity, 12–13
*Heteromeles arbutifolia, 534*
*Heterorhabditis hepialus,* 303
Heterostyly, 167
Heterotrophs, 114
Hexagonal designs, 272
Hickories, 327, 372, 507, *509,* 529
*Hieracium aurantiacum* (orange hawkweed), 207
Hiesey, William, 251–253
High Plains, *466*
Hill, Mark, 344
Hill metric, 343, 344
Hillebrand, Helmut, 546
Himalayas, 464, 484
Histisols, *87, 88*
Hitchcock, A., 545
Hnatiuk, Roger, 365
Hobbs, Richard, 407
Holah, Jenny, 324
Holocene epoch, *496*
Honey mesquite (*Prosopis glandulosa*), 299
Honeybees, 161, 169
Hooper, David, 407, 417
Horizons, 94–97
Hornworts, *497*
Horse chestnuts, 529
Horses, 503
Horsetails, *259,* 311, *497,* 498
Horseweed (*Conyza canadensis*), 372, 387
Horvitz, Carol, 158
Host shifts, 321
Hot deserts
characteristics of, 537–538
dominant growth forms and general climatic conditions, *516*
global distribution, *514*
impact of human activity on, 538
physiognomy, 520

temperature and precipitation, *515*
β-diversity, 550
Houle, David, 179
Howard, Timothy, 398
"Howie's Fire Glow," *231*
Hubbard Brook Experimental Forest, 358
Hubbell, Stephen, 209, 562
Hughes, Randall, 415
Human activity
biodiversity loss and, 567–569, 570–572
causes and effects of habitat fragmentation, 439–443
disturbance and, 384–385
global climate change (*see* Anthropogenic global climate change)
impact on alpine regions, 539
impact on biomes, 570
impact on hot deserts, 538
impact on taiga, 533
impact on temperate deciduous forest, 529
impact on temperate grassland, 534–535
impact on temperate rainforests, 530
impact on temperate shrubland, 533
impact on the fynbos, 548
impact on tropical rainforests, 525–526
impact on tropical savanna, 536
impact on tundra, 541
nitrogen deposition and acid precipitation, 132–135, 136, 137
soil erosion, 96
*See also* Agriculture; Deforestation; Logging
Humans
benefits of biodiversity, 572–573
migrations during the Pleistocene, 508
population growth and biodiversity loss, 570–572
Humboldt Current, 462, 464, 465
Humic acid, 97
Humidity, 457
Hummingbirds, *155*
Hump-shaped distribution, 563, 565
Humus, 97
Huntington, Thomas, 137
Huntly, Nancy, 305
Hurricanes, 381, 483
Hurtt, George, 562
Huston, Michael, 412
Hutchings, Michael, 148
Hutchinson, G. Evelyn, 231, 402
Hybrid swarms, 255
Hybridization
adaptation and speciation through, 255, 256–257
concerns about genetically modified crops and, 169–170
creation of invasive species and, 405
between invasive species and native relatives, 170
Hydraulic lift, 58, 101
Hydraulic redistribution, 58–59, 101
Hydrocharitraceae, *40*
Hydrogen and hydrogen ions
organic matter and, 97
plant nutrition, 101, *102*
root uptake of nutrients and, 130
soil pH and, 93, 94
in soils, 92
Hydrogen bonds, *93*
Hydrogen sulfide, 136
Hydroxy aluminum ions, 93
Hygrophytes, 61
*Hylobius transversovittatus,* 304
*Hyparrhenia hirta* (thatching grass), 274–275
*Hypericum, 283, 284*
*Hypericum fasciculatum* (St. Johnswort), 160
Hyphae
mycorrhizae, 326–327, *328,* 329
saprophytic fungi, 125
Hypothesis, 4
Hypothesis testing, 10
Hypsithermal, *505*
Hysteresis, 395

## I

Iberian Peninsula, 553, *554*
Ice storms, 382
Ida, Takashi, 29
Identity, biodiversity and, 342
Igneous rock, 85
Illicit trade in rare species, 525–526
Illite, *91*
Illuviation, 95
ILTER. *See* International Long-Term Ecological Research Network
Imaizumi, Toshiyuki, 245
Imbibition, 191
Immigration rate
island biogeography theory and, 432
metapopulation theory and, 434, 435, 436
SLOSS debate and, 446
Immobilization, of nitrogen, 132
*Impatiens,* 169
*Impatiens capensis* (jewelweed), *234,* 243–245
Importance value (IV), 355–356
Inbreeding, 162–163, 166
Inbreeding depression, 162, 169
Inceptisols, *87, 88*
Inclusive fitness, 181
Indeterminate inflorescences, *157*
India, carbon emissions, *478, 480*
Indian paintbrush (*Castilleja indivisa*), 317, 318
Indian pipe (*Monotropa uniflora*), *335*
Indian ricegrass (*Achnatherum hymenoides*), 302
Indiana Dunes, 181
Indicator species, 367
Indirect gradient analysis, 363–364
Individual-based model of metapopulations, 434
Induced defenses, 314–316
Induced resistances, 314–316
Induced responses, 314–316
Industry
carbon emissions and, 479, 480
industrial nitrogen fixation, 133
Inflorescences, 144, 145, 156, *157*
Infrared gas analyzer (IRGA), 56, 57
Infrared (IR) radiation, 450–451, 475
Inhibition, of colonization during succession, 391, 392
Initial floristic composition theory, 390, 393–394
Inorganic nitrogen, 105
Insects
in biological control, 303–305
decomposers, 124–125
decomposition in carnivorous plants, *130, 131*
evolutionary consequences of plant-herbivore interactions, 316–317
mutualisms in the Cretaceous period, 502
symbioses with plant pathogens, 324
*See also* Pollinators
Institute for Crops Research-Rothamsted, 358
Instituto Nacional de Pesquisas da Amazônia, 441
Integral projection models (IPMs), 211–212, 213
Intensity, of a disturbance, 377, *378*
Intercalary meristems, 144
Interference competition, 276
*See also* Allelopathy
Intergovernmental Panel on Climate Change, 573
Intermediate disturbance hypothesis, 415–416
Intermountain Desert. *See* Great Basin
International Long-Term Ecological Research (ILTER) Network, 358
International Union for Conservation of Nature (IUCN), 399, 573
Internodes, 144, *145*
Intertropical Convergence Zone (ITCZ), *458,* 461, 527
Intracellular conductance, 32
Intraspecific competition, 266–270
Intrinsic rate of population growth, 185
Invasive species
allelopathy and, 278
archaeophytes and neophytes, 403
causes of invasiveness, 404–405, 409
defined, 403

effects on communities and biodiversity, 403, 410–411
effects on grasslands and grass diseases, 323–324
factors making a community vulnerable to, 405–409
hybridization with native relatives and, 170
rules for predicting, *405*
seed dispersal and rate of spread, 174
weedy and ruderal species, 403–404
Inventory diversity, 342–345
IPMs. *See* Integral projection models
*Ipomoea,* 146
*Ipomoea purpurea, 234,* 249–250
*Ipomoea wolcottiana, 504*
IR radiation. *See* Infrared radiation
IRGA. *See* Infrared gas analyzer
Iridaceae, 547, 559
Iridescence, 49, 50
*Iridomyrmex humilis* (Argentine ant), 548
Iridoplasts, 49
*Iris bracteata, 361*
Irish potato famine, 319
Iron, 97, *103,* 116
Iron oxide, 135
Irrigation
impact on hot deserts, 538
impact on temperate grassland, 535
ISAR. *See* Island species-area relationship
Island biogeography theory
community restoration and, 395
neutral model of species interactions and, 293
overview and description of, 431–433
SLOSS debate and, 445–446
Island species-area relationship (ISAR), 352
Islands
defined, 431
geographic isolation and endemics, 557–559
island biogeography theory, 431–433
metapopulation theory and, 434
seed dispersal and, 174
*See also* Archipelagos; Oceanic islands
Isobilateral plants, 69
*Isoetes,* 42
Isoprene, 313
ITCZ. *See* Intertropical Convergence Zone
Iteroparous organisms, 177
IUCN. *See* International Union for Conservation of Nature
IUCN Red List, 399
IV. *See* Importance value
Iwasa, Yoh, 193

## J

Jaccard, P., 361
Jaccard index, *359,* 361, *362*
Jack pine (*Pinus banksiana*), 154, 389, 505, *509,* 532
Jack-in-the-pulpit (*Arisaema triphyllum*), 159
Jack-of-all-trades strategy, 241
Jacobs, Matthew, 49
Jansson, Roland, 549–550
Janzen, Daniel, 322
Janzen-Connell hypothesis, 322
Japan, land use and anthropogenic carbon emissions, 482
Japanese barberry (*Berberis thunbergii*), 310
Japanese knotweed (*Fallopia japonica*), 405
Japanese stiltgrass (*Microstegium vimineum*), 310
Jasmonate signaling pathway, 329
Jasper National Park, 365
Jennings, Michael, 565–566
Jenny, Hans, 84
Jet streams, 461–462
Jetz, Walter, 546
Jewelweed (*Impatiens capensis*), *234,* 243–245
Jia, Shihong, 301
Johnson, Anna, 161
Johnson, Edward, 375–377
Jojoba (*Simmondsia chinensis*), *80*
Jolliffe, Peter, 271
Jones, Sharon, 565
Jongejans, Eelke, 228
Jonkershoek Nature Reserve, *547*
Joshua tree (*Yucca brevifolia*), *53*
Jost, Lou, 344
*Juglans,* 528
*Juncus, 284*
*Juncus dichotomus* (path rush), 324
June solstice, 451, *452, 453,* 461
*Juniperus, 511,* 531, *532*
*Juniperus oxycedrus, 534*
Jurassic period, *496,* 499, *500,* 501
Just-so stories, 232

## K

Kaempferol, *312*
*Kageneckia oblonga, 534*
Kalahari Desert, 459, 537
*Kalanchoe daigremontianum* (mother of thousands), 150
Kalisz, Susan, 163, 208, 210, 279
*Kalmia latifolia* (mountain laurel), 529
Kamen, Martin, 26
Kankra (*Bruguiera gymnorrhiza*), 489
Kansas Fragmentation Experiment, 444–445
Kaolinite, *91*
Kaplan-Meier estimator, 208
Karoo, 533
Karpilov, Yuri, 27
Kazakho-Dzungarian Desert, 538
Keck, David, 251–253
Keddy, Paul, 274, 275, 282–283, 284
Keeley, Jon, 277
Keever, Catherine, 372
Kellner, James, 209
Kelly, Colleen, 148, 403
Kelly, David, 308
Kendall, Bruce, 219
Kentucky bluegrass (*Poa pratensis*), *154,* 286
Kéry, Marc, 209
Khalsi (*Aegiceras corniculatum*), 489
Kharouba, Heather, 161
Kier, Gerold, 557
Kilauea volcano, 386
Kinetic energy, 75
"King Clone" (creosote bush), 177
King protea (*Protea cynaroides*), *502*
Kinlock, Nicole, 294, 546
Kipukas, 386
Kira, Tatuo, 266, *267*
Kirchner Marsh, *509*
Kitajima, Kaoru, 303
Knapp, Alan, 406
Knapweed, 278
Knight, Tiffany, 160, 221
*Kobresia myosuroides,* 286
Koenig, Walter, 192
Konza Prairie Biological Station, 573
Konza Prairie Research Natural Area, *7,* 8, 354, 376, 382
Köppen-Geiger climate classification maps, *490*
Korb, Julie, 395
Kortschak, Hugo, 27
Koyukuk National Wildlife Refuge, *532*
Krakatoa, 384, 386
Kranz anatomy, 35–36
Kreft, Holger, 546
Kristiansen, Thea, 556
Krummholz, 518, *519*
*K*-strategists, 185
*See also* r- and K-strategy theory
K-T boundary, 503
Kudo, Gaku, 29
Kudzu, 247, *265,* 320
Kunin, William, 403
Kuroshio Current, *462*
Kutiel, Pua, 398
Kuyper, Thom, 281
Kwongan, 533

## L

La Niña, *468,* 469, 470
La Paz Waterfall Gardens Nature Park, *526*
La Roi, George, 365
La Selva Biological Station, *381*
Labrador Current, *462*
Lacey, Elizabeth, 245
*Lachenalia mutabilis, 157*
Lag phase, 405
Lagan region, 570

LAI. *See* Leaf area index
Lake effect, 465
Lakes, 465, *524*
Lakeshore dune communities, 375, 392, 396
LaMarche, Valmore, 302
Lamb's quarters (*Chenopodium album*), *504*
Lamiaceae, 66, 313, 533
Lamiales, *40*
Laminated root rot, 320
*Lamium amplexicaule,* 169
Land plants
  evolution in the Paleozoic era, 496–499
  evolutionary relationship to other eukaryotes, *497*
  key characteristics in the evolution of, 54–55
  major groups and their relationships, *497*
  Mesozoic era, 499–503
Land use changes
  anthropogenic global climate change and, 480, 481–482
  biodiversity loss and, *439,* 568–569, 570–572
Land use histories, 394, 395
Landa, Keith, 272
Landscape demography, 220, 437
Landscape ecology
  analysis of data by ordination or classification, 430–431
  categorical versus continuous views of nature, 429, 430–431
  causes and effects of habitat fragmentation, 439–443
  concern with questions of pattern and process, 422
  debate on community assembly rules, 433–434
  definition of and introduction to, 421–422
  demographic processes, 436–437
  dynamic equilibrium of communities, 395–396
  edge effects, connectivity, and nestedness, 443–445
  metacommunity theory, 437–439
  metapopulation theory, 434–436, *437*
  methods for analyzing spatial patterns, 426–427
  methods for describing the distribution of communities, 362–368
  origins, 18
  patches, 427–430
  plant conservation and, 439–446
  scale and, 422–427
  spatial patterns and ecological processes, 430, *431*
  theory of island biogeography, 431–433
'Landscape of fear' hypothesis, 307
Landscape scale, 423
*Lappula occidentalis, 224*
Larches, 194, 507, 532
Large-scale manipulative experiments, 7–8
Larios, Eugenio, 181, 210
*Larix,* 507, 532
*Larix kaempferi* (ochudo), *194*
*Larix laricina,* 194, *509,* 532
*Larix siberica* (Siberian larch), *397,* 399, 409
Larkspurs, *154*
*Larrea divaricata, 199*
*Larrea tridentata* (creosote bush), 45, *61,* 93, 221
Latent heat loss, 75, 78, 79
Latent heat of vaporization, 78
Lateral roots, *145*
Latex, 313, 316–317
*Lathyrus maritimus, 549*
Latitude
  climate and, 455
  solar angle of incidence and seasonality, 451–452, *453*
Latitudinal diversity gradient
  available energy hypothesis and, 563
  bounded ranges model and, 562
  explanations for, 546, 548–550
  overview, 545–546
Lauenroth, William, 44–45, 354
Laurance, William, 571
Laurasia, 499, *500*
Law, R., 289
Lawrence Berkeley Laboratory, 26
Leaching
  defined, 92
  of nitrate, 131
  nitrogen saturation and, 133
  of phosphorus, 108
Leading eigenvalue, 214
Leaf anatomy
  adaptations in dry environments, 68–71
  $C_3$ plants, *59*
  $C_4$ plants, 35–36
  light environment and photosynthesis, 29, *30*
  mesophytes, 67, *68*
Leaf area index (LAI), 139
Leaf conductance
  to carbon dioxide, 32
  to water vapor, 59, 60
Leaf economic spectrum, 275, 276
Leaf mass per leaf area (LMA), 47
Leaf miners, 298, *299,* 311
Leaf shape, ecotypes, 253–254
Leaf water potential, 67, *68, 101*
Leaves
  boundary layer, 78
  characteristic dimension of, 78
  decomposition of, 124, 132
  dissected, 253–254
  effects of light environment on photosynthesis, 29–30
  energy balance equation, 75
  energy flow via conduction and convection, 77–78
  factors affecting carbon dioxide and plant carbon uptake, 31–32, *33*
  factors determining leaf temperature, 78–80
  latent heat loss, 78
  leaf primordia, *145*
  leaf rolling in response to water stress, 80
  pathogens, 320
  phenotypic plasticity in aquatic plants, 241, *242*
  physical defense to herbivory, 310–311
  radiant energy, 76–77
  senescence and nitrogen mobilization, 132
  soil–plant–atmosphere continuum of water movement, 57, *58*
  solar tracking, 48–49
  structural coloration, 49, 50
  succulence and CAM photosynthesis, 39
  sun and shade leaves, 47–48
  temperature extremes for damage from different environments, *81*
  transpiration rate, 59–60
  vegetative phenologies and leaf expansion, 194–195
  worldwide leaf economic spectrum, 275, 276
*Leersia, 283*
Left eigenvectors, 216
Legumes (Fabaceae)
  cryptic dioecy, 165
  ectomycorrhizae, *325*
  nitrogen fixing symbioses, 106–107
  secondary chemicals, 312–313
  thorn woods, 527
  tropical woodlands, 528
Leibold, Matthew, 246, 437
Leishman, Michelle, 404, 407
Lemus-Albor, Aurora, 443
*Lepidium virginicum* (pepperweed), 315
Lepidopterans, 301, 502
Lesica, Peter, 192–193
*Lesquerella vicina,* 545
Lesser snow geese, 307
Lethal limit, 80
*Leucadendron salignum, 534*
Levine, Jonathan, 406, 407, 425–426
Levins, Richard, 434
Lewontin, Richard, 231
Lianas, 525
Lichens, 107, 134
*Licuala arbuscula,* 30

LiDAR (light detection and ranging), 427–429
Liebig's law of the minimum, 564
Life cycle graphs, 206, *207*, 213, 216
Life cycles, sexual, 150–153
Life history
  defined, 177
  demographic life history theory, 184, 186–188
  effects of year-to-year environmental variation on, 189–193
  evolution and, 181–184
  fast-slow continuum, 188–189
  Grime's triangular model, 185–186
  patterns of succession and, 390–391
  phenology, 193–197
  *r*- and *K*-selection theory, 184–185
  *r*- and *K*-strategy theory, 185
  trade-offs and, 178–181
  variation and, 177–178
Life table response experiments (LTREs), 218–219
Life tables, 207, 208–209
Light
  photons, light wavelengths, and photosynthesis, 22, *23*, 24, *25*
  plant adaptations to, 46–50, 243–245
  plant competition for, 263, 264, *265*, 392–393
  Rayleigh scattering, 77
  *See also* Light environment; Solar energy
Light compensation point, 28–29, 48
Light detection and ranging (LiDAR), 427–429
Light environment
  effects on photosynthesis, 28–31
  introduction to and importance of, 22
  ozone and, 30–31
  photosynthetic response in $C_3$ and $C_4$ plants to light intensity, *38*, 39
  plant adaptations to, 46–50, 243–245
  plant competition for light, 263, 264, *265*, 392–393
  plant reproductive phenologies and, 195–196
  seed germination and, 191
  shoot architecture and light interception, 146–147
Light reactions, 22, *23*, 24
Light saturation, 32
Lignins, 313
Lignite, *97*
Likens, Gene, 376
Liliaceae, 150, 320–321
Liliales, *40*
*Lilium catesbaei*, *152*
Limber pine (*Pinus flexilis*), *511*
Limestones, 108
Lin, Chi-Hua, 220–221
Linden, 528
Linder, Hans Peter, 545
*Lindera*, *552*
Lineage divergence, 346–347
Line-intercept method, 355, 356
Linnaeus, Carl, 544
*Liquidambar*, 528, *552*
*Liriodendron*, *552*
*Liriodendron tulipifera* (tulip poplar), *73*, 326, *339*, 529
*Lithraea casustica*, *534*
Litter, 93, 124
Little Ice Age, *505*, 510–511
Liverworts, *497*, *498*
Livestock grazing
  grassland restoration and, 395
  greenhouse gases and, 482
  impact on grasslands, 385, 569
  impact on hot deserts, 538
  impact on plant communities, 309
  impact on shrublands, 385
  impact on tropical rainforests, 525
  plant pathogen interactions and, 324
  *See also* Herbivory
Llanos, 536
LMA. *See* Leaf mass per leaf area
Loamy sand, *90*
Loamy soils, 89, *90*, *101*
Loasaceae, 311
*Lobelia cardinalis* (cardinal flower), *155*
*Lobelia rhynchopetalum*, 177, 521
Loblolly pine (*Pinus taeda*), 301, 372, 531
Local abundance
  commonness and rarity, 398–403
  competitive dominants, 398
  introduction, 397–398
Local communities, 334
Local extinctions
  available energy hypothesis of species richness and, 563
  island biogeography theory and local extinction rates, 432
Local guilds, 335
Local scale, 423
Locoweed, 309
Lodgepole pine (*Pinus contorta*)
  in boreal forests, 532
  cluster analysis of stands, *368*
  direct gradient analysis of stands, 365
  distribution patterns and ecological processes, 430, *431*
  pollen grains, *504*
  polythetic-agglomerative analysis of stands, *367*
  serotiny, 389
  southern pine beetle and, 301
*Lodoicea maldivica* (coco de mer or sea coconut), 179
Loess, 85, *86*
Log response ratio (LRR), 271, 287, *288*
Logging
  effects of clear-cutting, 384–385
  fire and forest management, 380–381
  impact on temperate deciduous forest, 529
  impact on the carbon cycle, 481
  impact on tropical rainforests, 525
Logging roads, 385
Logistic population growth, 185
*Lolium perenne* (ryegrass), 290, 398
Londsdale, Mark, 406
Long-distance dispersal, 173–174
Longevity of plants, 147
Longleaf pine (*Pinus palustris*)
  adaptations to fire, 391
  demographic studies, 221
  impact of hurricanes on population growth rates, 222
  longevity, 184
  pine barrens and piney woods, 531
  self-pruning, 147
  southern pine beetle and, 301
Long-lived plant populations, demographic studies of, 221–222
Long-term community studies, 357–358
Long-Term Ecological Research (LTER) Network, 358, 376
Long-term manipulative experiments, 7–8
Longwave radiation, 77
*Lonicera maackii* (Amur honeysuckle), 405
*Lonicera periclymenum* (European honeysuckle), *299*
Loranthaceae, 317
Lortie, Christopher, 287, 288
Los Alamos (NM), 380
Lotka-Volterra competition model, 290, 291
Lottery model, 291–292
Louda, Svata, 304, 313–314
Love lies bleeding (*Amaranthus caudatus*), *40*
Lovett-Doust, Jon, 270
Lovett-Doust, Lesley, 148–149
LRR. *See* Log response ratio
LTER Network. *See* Long-Term Ecological Research Network
LTREs. *See* Life table response experiments
*Lupinus*, 108
*Lupinus albus*, 318
*Lupinus arboreus* (yellow bush lupine), 303
*Lupinus latifolius* (broadleaf lupine), 387
*Lupinus texensis* (bluebonnet), 166
Lutein, *25*
Luteolin, *312*
Lycopene, 24
Lycophytes, *497*
Lycopodiophyta, 498
Lycopods, 66
*Lymantria dispar* (gypsy moth), 301, 382, *384*
Lyons, S. Kathleen, 562

*Lysimachia, 284*
*Lythrum salicaria* (purple loosestrife), 167–168, 282, 304, *404*

## M

MacArthur, Robert, 184, 185, 264, 274, 291, 411–412, 431, 432, 433
Mack Lake Fire, 379, 380
Mackey, Robin, 415
Macrofossils, 504, 505
Macronutrients, *102,* 103–104
Macropores, *92*
Madagascar
  carbon emissions, 479
  geographic isolation and endemics, 557–558, *559*
  human-caused extinction of large mammals, 508
Madrone, 323
Maestre, Fernando, 287
Magnesium, 92, 93–94, *102,* 116
*Magnolia,* 529, *552*
*Maianthemum japonicum* (false Solomon's seal), 29
Mal, Tarun, 282
Malate, *35, 38*
Malcolm, Stephen, 316–317
Male-male competition, 165
Malic acid, *38*
Malmstrom, Carolyn, 323–324
Malpighiales, *40*
*Malva pusilla* (round-leaved mallow), 175
Malvales, *40*
Mammals
  evolutionary consequences of plant-herbivore interactions, 316
  herbivory and its effects on plant community composition, 306–307
  *See also* Cattle; Livestock grazing
*Mammillaria dioica* (pincushion cactus), 164
*Mammillaria geminispina* (twin spined cactus), *43*
Mandelbrot, Benoit, 429
Manganese, 97, *103*
Mangrove forests, 489
Manipulative experiments, 5–8
Mantle, *326,* 327, *328,* 329
Manú National Park, 401
Manure, 134, 482
Manzanita (*Arctostaphylos manzanita*), 323
Maples
  decomposition of leaves, 132
  Pleistocene glaciation and, 507, *509*
  seed dispersal, 171
  temperate deciduous forests, 528, 529
Mapping, 356
Maputaland, 560
*Marchantia, 498*
Marigold (*Tagetes patula*), *268*
Marijuana (*Cannabis sativa*), *68*
Mariposa lilies, 227
Maritime climate, 467
Mark-recapture method, 207, 209
Marram grass (*Ammophila arenaria*), 324
Marshes, *524*
  *See also* Salt marshes
Marsico, Travis, 304
Martinez, Kelsey, 436
Marvier, Michelle, 310
Mass balance, 114
Mass effects, 438, *439*
Masting, 177, 192–193, 303
Maternal effects, 245
Mathematical models, 4
Mating systems
  assortative mating, 166–167
  competition among plants and among pollen grains, 165
  defined, 162
  forms of gender expression, 163–165
  frequency-dependent selection, 167–168
  inbreeding, 162–163
  mating between neighboring individuals, 166
  other important consequences of, 168–170
Matorral, 533, *534*
Matric potential, 55, 57
Matrices, multiplying by a population vector, 214
Matrix models, 188
Matrix projection models
  analyzing for population growth rates and population composition, 213–215
  COMPADRE database, 219
  constructing, 212
  effects of transition probabilities on population growth rate, 219
  overview and description of, 211–212
  population properties derived from, 212–213
  sensitivities and elasticities of population growth, 212–213, 217–218, 219
  strengths and limitations of, 229
Mauna Kea, 464
Mauna Kea silversword (*Argyroxiphium sandwicense*), 220, *221*
May, Robert, 417
Mazer, Susan, 181
Mazzola, Mark, 98
McAuliffe, Joseph, 290
McDaniel, Sierra, 30
McGill, Brian, 572
McGinley, Mark, 180
McIntosh, Robert, 355–356
McPeek, Mark, 210, 222
McPherson, Guy, 299
Meadow rues, 153
Meadows, 395
Mead's milkweed (*Asclepias meadii*), 209
Mean field approach, 351–352
Mean field models, 434
Mean residence time, 114, 120–121
Mediterranean climate, 466, 519
Megagametophytes, 151
*Megalodonta beckii* (Beck's water marigold), *242*
Megapascals (MPa), *55*
Megaspores, 55
Mendel, Gregor, 237
Menges, Eric, 228
*Mentha cervina,* 400
Meristems
  of grazed grasses, 316
  plant growth, 144–145
  propagule pool and succession, 387
Mesic habitats, 61
Mesophyll
  $C_3$ leaves, 29, *30, 36*
  $C_4$ leaves, *35, 36*
Mesophyll conductance, 32
Mesophytes
  behavior of stomata in, 67–68
  leaf anatomy, 67, *68*
  properties of stomata, *69*
  water availability and, 60–61
Mesophytic flora, *496*
Mesozoic era, 41, 499–503
Mesquite, *65,* 299
Meta-analyses, 12, 287, 289
Metabolic energy, 75
Metabolic heat, 75
Metacommunities, 437–439
Metamorphic rock, 85
Metapopulation models, 435
Metapopulation theory, 434–436, *437*
*Metasequoia,* 502
Methane
  a greenhouse gas, 475, 476
  net ecosystem carbon balance and, *118*
  net ecosystem production and, 119
  produced by livestock, 482
2-Methoxy-4-vinylphenol, 158
*Metrosideros angustifolia, 534*
Mexican bark beetle, 301
Michalet, Serge, 99
*Microbotryum violaceum, 408*
Microfossils, 504–505
Microgametophytes, 151
Microhabitats, 13, 424, *425*
Micrometeorology, 126, 127
Micronutrients, *103,* 104
Microspores, 55, *92*
*Microstegium vimineum* (Japanese stiltgrass), 310
*Microtus agrestis,* 309
Middens, 505
Midgley, Guy, 474
Migration
  effects of human land use on, 572
  emergence of modern plant communities and, 508–510

evolutionary change and, 232
genetic variation and, 248
human migrations during the Pleistocene, 508
vertical, 511
Milanković, Milutin, 455–456
Milankovitch cycles, 455–457
Milkweeds, 313, *314,* 316–317
Millennial model, 119
Millennium Ecosystem Assessment, 572, 573
Miller, Kathryn, 572
Miller, Thomas, 291
Milne, Bruce, 429
*Mimosa luisana,* 279
*Mimulus,* 169
*Mimulus cardinalis,* 169
*Mimulus lewisii,* 169
Mine waste, heavy-metal tolerance, 241–243, 329, 521
Mineralization
defined, 123
of nitrogen, 129–131
of sulfur, 136
Minimum viable population (MVP), 228
Mint family, 313
Miocene epoch, *496,* 503–504
Miombo, 527
Misodendraceae, 317
Mistletoes, *171,* 232, 317, 318
Mites, 125
Mitochondria, photorespiration and, 34
Mittelbach, Gary, 563
Mixed deciduous-evergreen forest, 528
Miyanishi, Kiyoko, 375–377
Modular growth, 144–145
Moist tropical forest, 523–526
*See also* Tropical montane forest; Tropical rainforest
Mojave Desert, 520, 537
*Molinia caerulea,* 135
Mollicutes, 321
Mollisols, *87, 88*
*Mollugo,* 41
Moluccan albizia (*Falcataria moluccana*), 279, *280*
Molybdenum, *103*
Monarch butterflies, 316–317
Monilophytes, *497*
Monkey puzzle tree (*Araucaria araucana*), 500
Monocarpic organisms, 177
Monocarpic ramets, 183–184
Monocots
phylogeny of $C_4$ photosynthesis, *40*
secondary growth, 145, *146*
vegetative reproduction, 150
Monocultures, 266–270
Monoecy, 163, 164
*Monoptilon bellioides, 223, 224, 225*
Monoterpenes, 301, 315, 316
Monothetic-divisive analysis, 366–367
*Monotropa uniflora* (Indian pipe), *335*
Montane species
global climate change and, 491–492
*See also* Mountains and mountain ranges
Monterey pine (*Pinus radiata*), *328,* 399, 548
Montmorillonite, *91*
Mooney, Hal, 46–47, 410, 540
Moore, Joslin L., 207
Moosewood (*Acer spicatum*), *339*
*Morella faya,* 407
Mori, Akira, 562
*Moricandia arvensis* (colletxó or purple mistress), 41, *42*
Mormon tea (*Ephedra viridis*), *511*
Morning glories, 146, *234,* 249–250
Morrison, Janet, 324
Morrow, Patrice, 302
Mortality, 199, 269–270
*Morus, 552*
Mosses, 134, *497*
Moth mullein (*Verbascum blattaria*), 175
Moth pollination, 159
Mother of thousands (*Kalanchoe daigremontianum*), 150
Mt. Kilimanjaro, 485
Mount St. Helens, 384, 387, 394
Mount Washington, 518, *519*
Mountain big sagebrush (*Artemisia tridentata vaseyana*), *61*
Mountain Fynbos, 547
Mountain laurel (*Kalmia latifolia*), 529
Mountain maple (*Acer spicatum*), *339*
Mountain pine beetle (*Dendroctonus ponderosae*), 301–302
Mountain silverbell (*Halesia monticola*), *339*
Mountain sorrel (*Oxyria digyna*), 46–47, 540, 541
Mountains and mountain ranges
altitudinal gradient in species frequencies, 337–338, *339*
effects of global climate change on montane species, 491–492
effects of nitrogen deposition on forests, 134
effects on patterns of precipitation, 464–465, *466,* 470–471
effects on species diversity, 554–556
montane glaciers and global climate change, 484–485
temperature and elevation, 455
Mpassa forest, 381
*Mucuna holtonii,* 158
Multidimensional niche, 402
Multiple regression, 364
Multivariate statistics, 359, 360–361, 362–364
Musk thistle (*Carduus nutans*), 304
Muskeg, 507, 532
Mustard family, 313–314
Mutations
evolutionary change and, 232
genetic variation and, 248
Mutualisms
Cretaceous period, 502
defined, 105
facilitation and, 279
mycorrhizae, 325, 329 (*see also* Mycorrhizae)
MVP. *See* Minimum viable population
Mycoplasmas, 321
Mycorrhizae
arbuscular mycorrhizae, *325,* 326–327, 328, 329, 330
arbutoid ectendomycorrhizae, *325,* 329
definition and importance of, 325–326
ectomycorrhizae, *325,* 326–327–328, 329
effect on plant-plant interactions, 330–331
ericoid mycorrhizae, *325, 326,* 328–329
facilitation and, 280–281
major types and characteristics of, *325, 326*
as mutualists or parasites, 329–330
orchidaceous mycorrhizae, *325, 326,* 328, 329
other functions of, 329
plant-plant interconnections and, 331
as symbioses, 105
Myers, Jonathan, 556
*Myosotis alpestris* (alpine forget-me-not), 399–400
*Myosotis micrantha, 264,* 268–269
*Myrica faya. See* Morella faya
Myrosinase, 160
Myrtaceae, 313, *325*
Myrtales, *40*
*Myrtillocactus geometrizans,* 522
Myrtle oak (*Quercus myrtifolia*), *311*
Myxomatosis, 308


## N


NAD-malic enzyme, 35
NADP-malic enzyme, 35
Naeem, Shahid, 416
Nakagawa, Shinichi, 11
Nama Karoo, 547
Namib Desert, 537
*Narcissus,* 150
Narrowleaf plantain (*Plantago lanceolata*), 166, *167,* 219–220, 245, 330
Narrow-sense heritability, 239
*Nasutitermes triodiae* (cathedral termite), *537*
National Ecological Observatory Network (NEON), 126

National Institute for Research in the Amazon, 441
National Vegetation Information System (Australia), 341
Natural experiments, 8, 273
Natural gas, 477, 479
*See also* Fossil fuels
Natural selection
competition and, 261
evolutionary change and, 232–235
heavy-metal tolerance, 241–243
at levels other than the individual, 246–247
life history and, 181–184
origin of new species and, 254–256
patterns of adaptation and, 240–246
phenotypic variation and, 233
problem of Panglossian assumptions and adaptive speculations, 231–232
reproductive value and, 216–217
The Nature Conservancy, 341
Nature reserves, 445–446
NDVI. *See* Normalized difference vegetation index
Nearest neighbor probability, 430
Nebish soil series, 89
NECB. *See* Net ecosystem carbon balance
Necrosis, 321
Nectar guides, 155
Nectar rewards, 154, 502
Nectar robbers, 158
NEE. *See* Net ecosystem exchange
Needle-leaved evergreen forests, 517, 518, 531
Negative assortative mating, 167
Nematodes, *124,* 125
*Neobuxbaumnia tetetzo* (cardón), 279
*Neogalerucella calmariensis,* 304
*Neogalerucella pusilla,* 304
Neogene period, *496*
NEON. *See* National Ecological Observatory Network
Neophytes, 403
*Neotoma,* 505
NEP. *See* Net ecosystem production
*Nerisyrenai camporum* (fanmustard), *70*
Nestedness, 444, 445
Net carbon dioxide exchange, 28–29
Net ecosystem carbon balance (NECB), *118,* 119–120, 122
Net ecosystem exchange (NEE), *118,* 126
Net ecosystem production (NEP), 119–120, 122
Net primary production (NPP)
defined, 116
factors limiting, 117–118
global carbon cycle, 473
methods of estimating, 125, 126, 127
versus nitrogen input, *128*
ratio of aboveground to belowground NPP, 127
salt marshes, 122
soil microbial respiration and, 121–122
tropical rainforest, 523, *524*
variability over the Earth, 116–117
Net reproductive rate, 215
Nettles, 150
Network theory. *See* Ecological network theory
Neutral model of metacommunities, 438, *439*
Neutral model of species interactions, 292–293
Neutral theory of biodiversity and biogeography, 562
New Caledonia, 557, *559*
New Zealand
biodiversity loss, 567
exotic species and, 410
geographic isolation and endemics, 557, *559*
Newman, Edward, 282
Niche, 246
Nicholson, John, 300
Nickel, *103*
Nicobar, *559*
*Nicotiana tabacum* (tobacco), *23, 313,* 315
Nicotinamide adenine dinucleotide phosphate (NADP+/NADPH), *23,* 24–25, *38*
Nicotine, 313, 314–315
Nitrate
denitrification, 131
leaching, 131
nitrification, 129–130
nitrogen cycle and, 129–130
nitrogen deposition and, 134
plant nutrition and, 105
tropical rainforest soils, 525
wetland soils, 130
Nitric acid, 132, 137
Nitric oxide, 132, 133
Nitrification, 129–130, 131
Nitrifying bacteria (nitrifiers), 129–130
Nitrite, 129–130, 131
*Nitrobacter,* 129–130
Nitrogen
anthropogenic sources of, 133
atmospheric pool, 115
C:N ratios, 132
forms found in the atmosphere, 132–133
limiting nutrient for plant growth, 104–105
nitrogen deposition and acid precipitation, 132–135, 137
nitrogen fixing symbioses, 105–108
organic and inorganic forms in soil, 105
plant nutrition, *102,* 104
uptake by roots, 130
Nitrogen cycle
bacteria and nitrogen mineralization, 129–131
biological immobilization of nitrogen, 132
interconnection to other biogeochemical cycles, 115
nitrogen deposition and acid precipitation, 132–135, 137
nitrogen loss through denitrification and leaching, 131
overview and characteristics of, 127–129
wetland soils, 130–131
Nitrogen deposition, 132–135, 137, *439*
Nitrogen dioxide, 132, 133
Nitrogen fertilizers, 133, 134, 480
Nitrogen fixation
industrial fixation and nitrogen deposition, 133, 134
nitrogen cycle and, 128
nitrogen fixing symbioses, 105–108
Nitrogen oxide, 476
Nitrogen saturation, 133
Nitrogen use efficiency, 37
Nitrogenase, 106, 108
Nitrogen-fixing plants
nitrogen fixing symbioses, 105–108
succession and, 392
*Nitrosomonas,* 129–130
Nitrous oxide, 131, 475, 476, 482
Noble fir (*Abies procera*), 387
Nodes, 144, *145*
Nonmetric multidimensional scaling, 362
Non-native species, 403, 567
*See also* Invasive species
Non-nested quadrats, 354–355
Nonrandom mating, 232
Nonvascular land plants, *497*
Norfolk Island pine (*Araucaria excelsa*), 500
Normalized difference vegetation index (NDVI), 126
North America
carbon emissions, 478–479
continental patterns in plant biodiversity, 550–553
continental-scale air movements and, 465, *466*
dominance of angiosperms in the Cenozoic era, 503–504
emergence of modern plant communities, 508–510
endemic species, 557
forests in the Cretaceous period, 502–503
human migration to, 508
K-T boundary, 503
Mesozoic era, 499, *500*
Pleistocene glaciation, 506, 507–508

seasonal variation in precipitation, 466, *467*
transition zones, 553–554
North Atlantic Current, *462,* 463
North Equatorial Current, *462*
*Nothofagus,* 530
*Notholithocarpus densiflorus* (tanoak), 323
Novoplansky, Ariel, 283
NPP. *See* Net primary production
Null hypothesis, 562
Null model, 562
Null treatment, 8
Nunes, Carlos, 158
*Nuphar, 72,* 161
*Nuphar lutea* (yellow pond lily), *242*
Nurse plants, 279–280, 392
Nutrient Network (NutNet), 358
Nutrients
absorption by clay particles, 92
basic building blocks of plants, 101–103
beneficial mineral elements, *103,* 104
cycling (*see* Biogeochemical cycles)
effects of leaf nutrient levels on herbivory, 311
effects of soil nutrients on biological invasions, 407
effects of soil pH on availability, 92, 93
essential mineral elements, *102–103,* 103–104
mycorrhizae and plant uptake of nutrients, 325, 326–328, 329, 330, 331
nitrogen as a limiting nutrient, 104–105
phosphorus as a limiting nutrient, 108
plant competition for, 263, 264, 265
plant nutrition stoichiometry, 104
soil organic matter and, 97
sources of, 115–116
symbiotic nitrogen fixation, 105–108
*Nymphaea,* 161
*Nyssa, 552*

## O

O horizons, 94, 95
OAA. *See* Oxaloacetate
Oak-pine woodlands, 531
Oaks
Chestnut blight and, 323
decomposition of leaves, 124, 132
effects of herbivory on, 301
effects of leaf nutrient levels on herbivory, 311
hybridization and, 255
masting, 177
migration following the glaciers, 509
during the Pleistocene glaciation, 507
squirrel granivory and, 303
succession and, 390, 391
sudden oak death, 323
temperate deciduous forests, 528, 529
temperate evergreen forests, 531
temperate woodlands, 531
Oats, *36*
Obligate symbioses, 105
Obliquity, *455,* 456
Observational experiments, 8
Oceanic currents, 462–464, 465
Oceanic islands
biodiversity loss and, 566
geographic isolation and endemics, 557–559
island biogeography theory and, 433
*See also* Archipelagos; Islands
Oceans
calcium cycle and, 136–137
effects on air masses and precipitation, 465
evaporation and the global water cycle, 139
global carbon cycle, 472, 473
global climate change and, 484, 487, 488–489
sea level rise, 484, 488–489
sulfur pool, 136
Ochioi index, *362*
*Ochotona princeps* (pika), 305
Ochudo (*Larix kaempferi*), *194*
Ocotillo (*Fouquieria splendens*), *63,* 195
*Octodon degus* (degu), 306
*Odocoileus virginianus* (white-tailed deer), 298, *299,* 309–310, 316, 570
Odum, Eugene P., 339
Odum, Howard T., 339
*Oecology,* 15
*Oenothera arizonica,* 209–210
*Oenothera californica,* 209–210
Off-road vehicles, 538
Ogallala Aquifer, 535
Oil, 477, 479, 499
*See also* Fossil fuels
Oksanen, Lauri, 283
Öland Island, 358
Old-field succession, 372, 377, 394
Old-growth temperate rainforests, *529,* 530
Oli, Madan, 218–219, 442
Oligocene epoch, *496,* 503
*On the Origin of Species by Means of Natural Selection* (Darwin), 15, 232, 261
*Ondinea purpurea* (purple water lily), *242*
O'Neill, Robert V., 339
Onguene, Nerre Awana, 281
Onions, 150
*Onopordum illyricum,* 211
Oomycota, 318
*See also* Water molds
Oosting, Henry, 372, 376, 394
Open science initiatives, 11–12
*Ophioglossum reticulatum,* 255
*Ophiostoma ulmi,* 324
*Ophrys,* 159
Opium, *313*
Opium poppy (*Papaver somniferum*), *313*
Optimal control theory, 188
*Opuntia,* 201, 314
*Opuntia engelmannii, 511*
*Opuntia fulgida, 201*
*Opuntia humifusa* (devil's tongue), 81
*Opuntia inermis,* 303
*Opuntia rastrera,* 279
*Opuntia stricta,* 303
Orange hawkweed (*Hieracium aurantiacum*), 207
Orchard grass (*Dactylis glomerata*), 269
Orchid bees, 158
Orchidaceae, *325,* 329, 544
Orchidaceous mycorrhizae, *325, 326,* 328, 329
Orchids
human exploitation of, 525–526
number of species, 544
orchidaceous mycorrhizae, *325, 326,* 328, 329
pollinia and pollinators, 159
Ordination, 337, 362–364, 366, 368, 430–431
Ordovician period, *496*
Organic acids, 130
Organic matter, 97–98, 125, 127
Organic nitrogen, 105
Orinoco River, 536
Orobanchaceae, 317
Orthophosphate, 136
*Oryctolagus cuniculus* (European rabbit), 308–309
Osmoregulation, 66
Osmotic potential, 55, 66
*Osmunda,* 500
*Osteospermum pinnatum* (African daisy), *62*
"Out of the tropics" hypothesis, 552
Outbreeding depression, 163
Outcrossing, 163, 243
Ovaries, 151, *152, 154*
Overcompensation, 299–300
Overexploitation, biodiversity loss and, 570
Overgrazing. *See* Livestock grazing
Ovules, 151, *152, 154*
*Oxalis,* 169
*Oxalis purpurea, 143*
Oxaloacetate (OAA), 35, 36, *38*
Oxidales, *40*
Oxisols, *87, 88*
Oxygen
atmospheric levels in the Carboniferous period, 499
atmospheric pool, 115
flooded soils and aerenchyma, 72

plant nutrition, 101, *102*
released during photosynthesis, 24
rubisco and, 34
waterlogged soils and, 66–67
*Oxyria digyna* (alpine sorrel or mountain sorrel), 46–47, 540, 541
Oyashio Current, *462*
Ozone, 475, 476
Ozone hole, 30–31, 476
Ozone layer, 30–31

## P

Pacala, Stephen, 220
*Pachycereus pringlei* (cardón), *222*
Pacific Islands, 557, *559*
Pacific Ocean, El Niño Southern Oscillation, 468–470
Pacific silver fir (*Abies amabilis*), 127, *387*
Packrats, 505
Paine, Kellen, 156
Paleocene epoch, *496*
Paleoecology
Cenozoic era and dominance of the angiosperms, 503–504
definition of and introduction to, 495–496
emergence of modern plant communities, 508–510
K-T boundary, 503
Mesozoic era, 499–503
methodologies, 504–505
Paleozoic era and the rise of land plants, 496–499
Pleistocene glaciations, 505–508
Paleogene period, *496*
Paleophytic flora, *496*
Paleozoic era, 496–499
Palisade parenchyma, *30, 68, 70*
Palm oil plantations, 568
Palmer agave (*Agave palmeri*), *43*
Palms, 145, 311, 528
Palo verde (*Parkinsonia microphylla*), *31, 65,* 392
Palynology, 504–505
Pampas, 464, 534
Pandanales, *40*
*Pandanus,* 145, *146*
Pangaea, 499, 500, 501
Panglossian assumptions, 231
Panic grass (*Panicum sphaerocarpon*), 330
*Panicum*, 41
*Panicum maximum* (guinea grass), 274–275
*Panicum sphaerocarpon* (panic grass), 330
*Papaver somniferum* (opium poppy), *313*
Paper birch (*Betula papyrifera*), 532
Paper chromatography, 26
PAR. *See* Photosynthetically active radiation
*Paraberlinia bifoliolata* (African beli), 281
Paradox of enrichment, 413
Paramo, 539, *555*
Parapatric speciation, 254
Parasites
mycorrhizae and, 329–330
parasitic plants, 317–318
Parasitism, 105, 317–318
Parayil, Govindan, 571
Parent material, 85–86
Parental effects, 245
Park Grass Experiment, 358
Parker, Ingrid, 219
Parker, Timothy, 11
*Parkinsonia microphylla* (palo verde or foothill palo verde), *31, 65,* 392
Parmesan, Camille, 489
Parry's agave (*Agave parryi*), *64*
Pärtel, Meelis, 556
Partial pressure of carbon dioxide, 67
Particulate carbon, *118*
Paruelo, José, 44–45
Pastore, Melissa, 474
Patch dynamic models of metacommunities, 438, *439*
Patches
defined, 427
effects of habitat fragmentation on patch size, 444–445
ground-truthing, 429
island biogeography theory, 431–433
metapopulation theory and, 434
methods of aggregating into "natural units," 427–429
quantification of, 429–430
species coexistence and, 290
Path rush (*Juncus dichotomus*), 324
Pathogenicity, 321
Pathogens. *See* Plant pathogens
Patterns, 2
PCA. *See* Principal components analysis
PE. *See* Phylogenetic endemism
Pea family. *See* Legumes
Peacock begonia (*Begonia pavonina*), *49*
Pearcy, Robert, 49–50
Pearson product-moment correlation, 361
Peat, 131
Peat bogs, 122
*Pectis, 40*
*Pectocarya recurvata, 223, 225*
Pedicle, *154*
Peet, Robert, 221, 376
Pelargonidin, *312*
*Pelargonium myrrhifolium, 177*
Peng, Shijia, 407
Pennine Hills, 134
*Pentaclethra macroloba, 202*
PEP. *See* Phosphoenolpyruvate
PEP carboxykinase, 35
PEP carboxylase, 35, 37, 38
Pepperweed (*Lepidium virginicum*), 315
Per capita growth rate, 185
Percentage base saturation, 94
Perennials
adaptations to reduced water availability in grasses, 64
Charnov-Schaffer model of perennial life histories, 188
demographic life history theory, 184
life history variations, 184
semelparous perennials, 183
Peres, Carlos, 570
Perfect blackbody radiator, 76
Perfect flowers, 163
Perforations, xylem vessel elements, *74*
Perianth, 155
Perihelion, *455,* 456
Perimeter-to-area ratio, 429
Peripheral habitats, 282–283
Permafrost, 88–89, 485, 540
Permian period, *496,* 499
Peroxisomes, 34
*Persea, 552*
Peru Current, 462
PET. *See* Potential evapotranspiration
Petals, *154*
Pfaff, Alexander, 571
3PGA. *See* Phosphoglycerate
pH of soils. *See* Soil pH
*Phakopsora pachyrhizi,* 320
Phalanx growth, 148–149
Phanerophytes, 146, *147,* 357
*Phellinus weirii,* 320
PhenoCam Network, 358
Phenolics, 312–313, 315
Phenology
of $C_3$ and $C_4$ plants, 43–44
definition and concept of, 43, 193–194
global climate change and, 492
Phenotype
assortative mating and, 166–167
defined, 233
genetic systems of strict additivity and dominance, 237
genotype-environment interactions, 239–240
Phenotypic plasticity
adaptive plasticity to light conditions, 243–245
defined, 48, 233
in flowering times, 196
genotype-environment interactions, 239–240
importance for understanding other ecological concepts, 245–246
as a pattern of adaptation, 241, *242*
responses to environmental conditions and, 53–54
shoot architecture and light interception, 146–147
sun and shade leaves, 48
transgenerational, 245
Phenotypic selection, 234–235
Phenotypic variance, 239

Phenotypic variation
complete equation for, 240
gene-environment covariance, 240
genetic and nongenetic components, 238–239
heritability and the genetic basis of, 235–240
natural selection and, 233, 234–235
Pheromones, 301
Philippines, *559*
*Phillyrea angustifolia, 534*
*Phillyrea media, 534*
Phloem, *68*
Phloem plugging, 321
*Phoradendron californicum* (desert mistletoe), 232
*Phoradendron leucarpum, 171*
Phosphatases, 136
Phosphoenolpyruvate (PEP), 35, *38*
Phosphoglycerate (3PGA), 26, 34
Phosphorus
global phosphorus cycle, 115, 135–136
limiting nutrient for plant growth, 108
mycorrhizae and plant uptake of, 326–327, 328, 330, 331
plant nutrition, *102,* 104
sources of, 116
Photon flux, 76
Photoperiod, 46–47
Photorespiration, 34, 35, 41
Photorespiratory $CO_2$ pump, 41
*See also* $C_2$ photosynthesis
*Photorhabdus luminescens,* 303
Photosynthesis
$C_3$ photosynthesis, 33–35
$C_4$ photosynthesis, 33, 35–37, *38*
carbon isotopes and, 37
carbon reduction, 22, 24–25, 26–27
crassulacean acid metabolism, 33, 38–39
defined, 22
effects of the light environment on, 28–31
evolution of photosynthetic pathways, 39–42
introduction, 21–22
light reactions, 22, *23,* 24
measuring, 56
overview of the photosynthetic pathways, 33
quantum yield, 25
response to temperature, *36*
summary of, 25–26
Photosynthetic induction, 50
Photosynthetic photon flux density (PPFD), 22, 30, *38*
Photosynthetic rate
carbon dioxide uptake and, 32
effects of the light environment on, 28–31
light quality and, 22
light saturation level, 32
response to temperature, *37*
variation among species in different habitats, 32–33
Photosynthetically active radiation (PAR), 22
Photosystem I, 22, *23, 24*
Photosystem II, 22, *23,* 24, 34
Phreatophytes, 65
Phycobilins, 24
Phylogenetic diversity, 341, 342, 344, 345–347
Phylogenetic endemism (PE), 557, *558*
Phylogenetic relatedness, 341
Phylogeny
biodiversity and, 341, 342
of $C_4$ photosynthesis, *40*
defined, 39, 335
*Phymatopus californicus* (ghost moth), 303
Physical plant defenses
to herbivory, 310–311, 314
to pathogens, 321
Physiognomy
definition and concept of, 349, 357
forests, 516–518
grasslands and woodlands, 518–519
shrublands and deserts, 519–520
tree line, 518, *519*
using to classify landscapes, 367
variability across the globe, 514–515, *516*
Physiological ecology, 17–18, 21
Physiological stress, 185–186
Phytoalexins, 321
Phytochromes, 47
Phytoecdysones, 313, *314*
*Phytophthora,* 320
*Phytophthora cinnamomi,* 320
*Phytophthora infestans,* 319
*Phytophthora ramorum,* 323
Phytoplankton, 473
Phytoplasmas, 321
Pianka, Eric, 185
*Picea,* 507, *509,* 532
*Picea glauca* (white spruce), 154, 532
*Picea likiangensis,* 555
*Picea maritima* (black spruce), 518, *519,* 532
*Picea obovata* (Siberian spruce), 399
*Picea sitchensis* (Sitka spruce), 392
Pickett, Steward, 377, 440
Pigments, light absorption in photosynthesis, 22, 24
Pika (*Ochotona princeps*), 305
Pill millipedes, 124–125
Pin cherry (*Prunus pensylvanica*), 393
Pinaceae, *325,* 327
Pincushion cactus (*Mammillaria dioica*), 164
Pine barrens, 279–280, 301, 387, 531
Pine forests, 376, 517, 531
Pine savannas, 536
Piñero, Daniel, 228
Pines
bark beetles and, 301–302
causes of invasiveness, 404–405
gender expression, 163
habitat construction, 246
mating between neighboring individuals, 166
pollen, *504,* 505, *509*
serotiny, 387, 389
taiga, 532
Piney woods, 531
Pink lady's slipper (*Cypripedium acaule*), 203–204
Pinks, *68*
*Pinus albicaulis* (whitebark pine), 172–173, 301–302
*Pinus banksiana* (jack pine), 154, 389, 505, *509,* 532
*Pinus contorta. See* Lodgepole pine
*Pinus echinata* (shortleaf pine), 301
*Pinus edulis* (pinyon pine), 302, 531, *532*
*Pinus elliottii* (slash pine), *123,* 301, 531
*Pinus flexilis* (limber pine), *511*
*Pinus lambertiana* (sugar pine), *73*
*Pinus palustris. See* Longleaf pine
*Pinus ponderosa* (ponderosa pine), 391, 395, 430, *431*
*Pinus radiata* (Monterey or radiata pine), *328,* 399, 548
*Pinus resinosa* (red pine), 505, *509*
*Pinus rigida* (pitch pine), *95,* 280, 301, 387, 389, 531
*Pinus sibirica* (Siberian pine), 399
*Pinus strobus. See* Eastern white pine
*Pinus sylvestris* (Scots pine), 399
*Pinus taeda* (loblolly pine), 301, 372, 531
Pinyon pine (*Pinus edulis*), 302, 531, *532*
Pinyon-juniper woodland, 531, *532*
Pirnat, Janez, 443–444
Pisgah forest, 381
Pistillate flowers, 163
Pistils, *154*
*Pisum sativa* (garden pea), 237
Pitch, 301, 315–316
Pitch canker, 399
Pitch pine (*Pinus rigida*), *95,* 280, 301, 387, 389, 531
Pither, Jason, 552–553
Pitman, Nigel, 386, 401
Pits, xylem vessel elements, *74*
Pixels, 368
Plant defenses
to herbivory, 310–317
to pathogens, 321–322
*See also* Chemical plant defenses; Physical plant defenses
Plant ecology
basic principles of ecology and, 16–17
development of, 14–18

introduction, 1
modern trends in, 18
science as a human endeavor and, 18
science of ecology, 2–12 (*see also* Science)
subdisciplines, 17–18
Plant growth
basic structures in modular growth, 144–145, *146*
effects of soil pH on, 92–93
effects on resource acquisition, 146–149
introduction, 143
longevity of plants, 147
nitrogen as a limiting nutrient, 104–105
phenology, 194–195
phosphorus as a limiting nutrient, 108
seed germination (*see* Seed germination)
Plant litter, 93, 124
Plant pathogens
biological invasions and, 408–409
disturbance and, 384
effects on communities and populations, 322–324
effects on human communities, 319
examples of, 319–321
interactions with other organisms, 324
introduction and overview, 318–319
plant defenses and evolutionary responses to, 321–322
potato blight, 319
Plant physiological ecology, 17–18
*Plantago insularis* (desert Indianwheat), 192
*Plantago lanceolata* (narrowleaf or ribwort plantain), 166, *167,* 219–220, 245, 330
*Plantago major* (common plantain), 425–426
*Plantago ovata, 223, 224, 225*
*Plantago patagonica, 223, 224, 225*
Plantations, 385
*Plantesamfund* (Warming), 337
Plantlets, 148
Plant–plant interactions
allelopathy, 262, 273, 276–279
competition (*see* Competition)
facilitation (*see* Facilitation)
spectrum of, 262
*See also* Species interactions
PlantPopNet, 219–220
Plant-soil feedbacks, 278–279
Plasmodesmata, 36
Platt, William, 221
Platte thistle (*Cirsium canescens*), *183,* 302, 304–305
Playfair, William, 16
Pleistocene epoch
continental patterns of biodiversity and glaciation, 551, *552*
geological time scale, *496*
Little Ice Age and modern vegetation, 510–511
vegetation patterns and glaciations, 505–508
*Pleuraphis rigida, 38*
Pliocene epoch, *496*
PO. *See* Polar ordination
*Poa compressa* (bluegrass), 398
*Poa pratensis* (Kentucky bluegrass), *154,* 286
Poaceae, *509*
agamospermy, 150
$C_4$ photosynthesis, 39, *40*
number of species, 544
resurrection plants, 66
wind pollination, 153
*See also* Grasses
Poales, *40*
Podocarpaceae, 530
Poinsettia (*Euphorbia pulcherrima*), 156
Point diversity, 344
Point-intercept method, 356
Point-quadrat method, 356
Poison gas, 134
Polar cells, *458,* 461
Polar fronts, 461–462
Polar jet streams, *458,* 461–462
Polar ordination (PO), 362, 363
Policha, Tobias, 156
Pollen tubes, 151, *152,* 162, *166*
Pollen/Pollen grains
angiosperm sexual life cycle, 151, *152*
competition among, 165
from Kirchner Marsh, *509*
as a limiting resource, 164, 165
migration and genetic variation, 248
palynology, 504–505
self-incompatibility systems, 162–163
using genes to track, 238
*See also* Pollination
Pollination
aquatic plants, 161–162
association of floral characteristics with particular pollinators, 158–161
buzz pollination, 158
competition among plants for pollinators, 165
competition among pollen grains, 165
complex plant-pollinator interactions, 160
declining pollinator populations, 161
flower structure and, *154*
limiting unwanted visits to flowers, 158
plant adaptation to terrestrial environments and, 54
self-pollination, 164
specialized pollinators, 159
visual displays attracting animal visitors, 154–156, *157*
wind pollination, 153–154
*See also* Pollinators
Pollination syndromes, 159–160
Pollinators
association of floral characteristics with particular pollinators, 158–161
competition among plants for, 165
complex plant-pollinator interactions, 160
declining pollinator populations, 161
evolution in the Cretaceous period, 502
floral odors and acoustic guides, 156, 158
floral visual displays and, 154–156, *157*
habitat fragmentation and pollinator availability, 169
limiting unwanted visits to flowers, 158
plant reproductive phenologies and, 196–197
specialized pollinators, 159
Pollinia, 159
Polycarpic organisms, 177
Polygonaceae, 313
*Polygonum, 283*
*Polygonum viviparum,* 150
Polyploidy, 255–256
*Polypodium vulgare,* 255
Polythetic-agglomerative analysis, 367
Ponderosa pine (*Pinus ponderosa*), 391, 395, 430, *431*
Poole, Robert, 197
Pools
biogeochemical cycles and, 114–116
carbon pools, 119–122
defined, 114
quantifying, 114
turnover time and, 120–121
Poor man's umbrella (*Gunnera insignis*), *79*
Popper, Karl, 10
*Population Bomb, The* (Ehrlich), 571
Population density
defined, 199
density-dependent or density-independent population dynamics, 199–200
*r*- and *K*-selection theory, 185
Population dynamics
defined, 199
density-dependent or density-independent population growth, 199–200
effects of herbivory on, 300–302
habitat fragmentation and, 442
population growth (*see* Population growth)
Population ecology, 17

Population growth
basic equation of demography, 200
defined, 199
demographic models, 213–215
density-dependent or density-independent, 199–200
features of population models, 205
life cycle graphs, 206, *207*
life table response experiments, 218–219
methods of estimating vital rates, 207–211
net reproductive rate, 215
ramets and genets, 201
random fluctuations over time, 222–228
reproductive value, 215–217
sensitivities and elasticities of, 212–213, 217–218, 219
simplest mathematical models of, 205
types of models for structured populations, 211–213
Population growth rate
changes in transition probabilities and, 219
deriving from demographic models, 213–215
effects of competition on, 262
effects of habitat fragmentation on, 442
stable growth rate, 214
transient dynamics, 214–215
Population size
basic equation for, 200, 202
genetic drift and, 249–250
population growth and, 205
significance for conservation, 250
Population structure
age-structured and size-structured populations, 203
definition and concept of, 202
effects of plant variability in size and form on, 203–204
importance in plant demography, 202–203
models of population growth and, 205
stable structure, 213–215
Population vectors, 212, 213, 214
Populations
demographic studies of long-lived plant populations, 221–222
effects of competition on, 262
effects of habitat fragmentation on genetic diversity, 442
effects of herbivory on, 300–305
effects of plant pathogens on, 322–324
effects of population heterogeneity on demographics, 220–221
plant pathogen interactions with other organisms and, 324
*See also* Structured populations
*Populus balsamifera* (balsam poplar), 532
*Populus grandidentata* (big-toothed aspen), 387, *388,* 390
*Populus tremuloides* (quaking aspen), 149, 179, 246–247, 306–307, 387, 532
*Populus trichocarpa* (black cottonwood), *73,* 392
Porometers, 56
Portulaceae, 320–321, *325*
Postreproduction censuses, 206
Potassium
biogeochemical cycling, 115
plant nutrition, *102*
in soils, 92, 93–94
sources of, 116
Potato blight, 319
Potential evapotranspiration (PET), 138–139
*Potentilla,* 150
*Potentilla glandulosa,* 251
Poverty grass (*Danthonia spicata*), *234,* 235
PPFD. *See* Photosynthetic photon flux density
Prairies
effects of diversity on community stability, 417, *418*
fire and, 379
propagule pool and secondary succession, 387
soil, succession, and community restoration, 394
*See also* Temperate grassland
Precambrian, 39, 41
Precession, *455,* 456
Precipitation
acid precipitation, 132, 133
biomes and, 514, *515*
climate and, 448
climate diagrams, *467*
continental-scale patterns in, 464–465, *466*
defined, 457
deserts and, 519–520
effects of plants on, 96, 138
El Niño Southern Oscillation, 468–470
forest physiognomy and, 517
global climate change and, 483, 488
global patterns of air movement and, 457–464
global patterns of mean precipitation, *465*
hot deserts, 537
latitudinal gradient in dry and rainy seasons, *527*
Milankovitch cycles and, 457
net primary production and, *117*
physiognomy of grasslands and woodlands, 518–519
physiognomy of shrublands, 519
seasonal variation in, 465–468
significance for plant ecology and evolution of predictability in patterns, 470–471
water movement through soil and, 99
Predation, 298
*See also* Herbivory
Preregistration of studies, 11–12
Prereproduction censuses, 206
Prescribed burns, 380
Presence/absence data, *359,* 361, *362,* 365–366
Pressure potential, 55, 57
Preston, Frank, 412
Prickly pear cacti, 279, 303–304, 314, *511*
Primary growth, 145
Primary metabolites, 312
Primary productivity
defined, 116
of major biome types, *524*
measuring or estimating, 125–127
Primary research, 2–3
Primary succession
defined, 377
dispersal capacity of species and, 385, 386–387
facilitation of colonization, 392
Primary wall, xylem vessel elements, *74*
*Primula veris* (cowslip), 308, *540*
Principal components analysis (PCA), 189, 362, 363, 364
Principal coordinates analysis, 362
Productivity
defined, 116
effects of diversity on, 416–417
effects of global climate change on, 489
levels of, 116
net primary production, 116–118 (*see also* Net primary production)
salt marshes, 112–113, 122
standing crop, 412
*See also* Productivity gradients
Productivity gradients
effects on competition and facilitation, 282–288
species diversity and, 412–414, 563–566
Progymnosperms, *496*
Proline, 66
*Promerops cafer* (Cape sugarbird), *548*
Prop roots, *146*
Propagule rain, 435
Propagules
plant competition and, 265
succession and propagule pools, 387–389

*Prosopis glandulosa* (honey mesquite), 299
*Prosopis velutina, 65*
*Prosthechea boothiana* (Florida dollar orchid), *44*
*Protea acaulis,* 547
*Protea arborea, 534*
*Protea cynaroides* (king protea), *502*
*Protea lepicarpodendron* (black-beard sugarbush), *548*
Proteaceae
effects of soil type on distribution in South Africa, 281
fynbos, 547, 548, 559
Pangea and, 501
proteoid roots, 108
rareness of mycorrhizal associations, 325
serotiny, 175
Proteales, *40*
Proteoid roots, 108, 325
Protists, 125
Provisioning ecological services, 572
*Prunus pensylvanica* (pin cherry), 393
*Prunus pumila* (sand cherry), 375
Pseudoflowers, 321
*Pseudotsuga menziesii* (Douglas fir), *301,* 387, *511*
Psilophyta, 498
*Psoralea obliqua, 534*
Psychrometers, 56–57
Pteridophyta, 498, 499–500
Pteridospermatophyta, 499–500
Pubescence, 78, 79, 80
Public policy, science and, 11
*Puccinia monoica,* 321
*Pueraria,* 247
*Pueraria lobata, 265*
*Pueraria montana,* 320
Puerto Rico, 465
Purple clover (*Trifolium pratense*), 324
Purple cudweed (*Gnaphalium purpureum*), 372
Purple loosestrife (*Lythrum salicaria*), 167–168, 282, 304, *404*
Purple mistress (*Moricandia arvensis*), 41, *42*
Purple pitcher plant (*Sarracenia purpurea*), 130, *131*
Purple sage (*Salvia leucophylla*), 277
Purple water lily (*Ondinea purpurea*), *242*
Puszta, 534
Pyrite, 136
Pyrogenic species, 380
Pyruvate, *35, 38*
Pyšek, Petr, 410
*Pythium,* 320

## Q

Qatar, 479
Qian, Hong, 550, 552
Quadrats, 349–350, 351, *352,* 354–355
Quaking aspen (*Populus tremuloides*), 149, 179, 246–247, 306–307, 387, 532
Quantum yield, 25
Quaternary period, *496*
Queen Anne's lace (*Daucus carota*), 171
Quercetin, *312*
*Quercus. See* Oaks
*Quercus agrifolia, 534*
*Quercus alba* (white oak), 303, 327, *338*
*Quercus dumosa, 534*
*Quercus garryana* (garry oak), *504*
*Quercus ilex, 534*
*Quercus ilicifolia* (scrub oak), 280, 387
*Quercus lobata* (valley oak), 169
*Quercus myrtifolia* (myrtle oak), *311*
*Quercus rubra. See* Red oak
*Quercus velutina* (black oak), 391
*Quillaja saponaria, 534*
Quillworts, 42
Quintana-Ascencio, Pedro, 228
Quiver tree (*Aloe dichotoma*), *333*

## R

*R**, 264–265, 291
*r*- and *K*-selection theory, 184–185
*r*- and *K*-strategy theory, 185–186, *187*
*R* selection, 186
Raaijmakers, Jos, 98
Rabbit viral hemorrhagic disease, 309
Rabbits, 308–309
Rabinowitz, Deborah, 399
Racemes, 156, *157*
Radiant energy
defined, 75
energy balance of the Earth, 448–451
exchange between plants and surroundings, 76–77
long-term cycles in the Earth's orbit and, 455–457
plant control of radiant heat input, 79–80
solar angle of incidence and, 451–452, *453*
Radiata pine (*Pinus radiata*), *328,* 399, 548
Radicle, 144
Raffa, Kenneth, 302, 315–316
Rain shadows, 464–465, 470–471
Rainfall. *See* Precipitation
Rainforests. *See* Temperate rainforest; Tropical rainforest
Ramankutty, Navin, 570
Ramets
defined and distinguished from genets, 201, 246–247
population dynamics and, 201
semelparous plants, 183–184
vegetative reproduction and, 150
*See also* Clonal plants
Randomized experiments, 6, 9
Ranomafana National Park, *559*
Ranunculales, *40*
*Ranunculus,* 150
*Ranunculus adoneus* (snow buttercup), 195–196
*Ranunculus reptans* (creeping spearwort), 148
*Ranunculus trichophyllus* (threadleaf crowfoot), *242*
*Raphanus sativus* (wild radish), 160
*Raphus cucullatus* (dodo), 171–172
Rara Avis Rainforest, *523*
Rare species, illicit trade in, 525–526
Rarefaction, 353
Rarity
causes of, 402–403
concept and description of, 398–399
scheme for describing, *399*
in space or time, 399–402
species characteristics associated with, 403
variability with spatial extent, 417–418
Raspberries, 150, 372
Rathcke, Beverly, 197
Raunkiaer, Christen, 357
Raunkiaer system, *147,* 357
Ray florets, *157,* 180
Rayleigh scattering, 77
RCI. *See* Relative competition index
Reaction center, 24
Reader, R., 286, 288
Realized fertility, 223–224
Realized niche, 246
Receptacles, *154*
Reciprocal averaging, 362
Recreational off-road vehicles, 538
Recruitment, 209
Red alder (*Alnus rubra*), 387
Red algae, 24
Red light, plant sensitivity to day length and, 47
Red mangrove (*Rhizophora mangle*), *72*
Red maple (*Acer rubrum*), *173,* 281, 326, *388, 504*
Red oak (*Quercus rubra*)
changes in importance along a moisture gradient, *338*
ectomycorrhizae, 327
granivory by squirrels, 303
life history patterns and succession, 390, 391
succession in abandoned agricultural fields, 372
succession northern forests, *388*
Red osier (*Cornus stolonifera*), 375
Red pine (*Pinus resinosa*), 505, *509*
Red ribbons (*Clarkia concinna*), 430
Redwood (*Sequoia sempervirens*), 470, 530
Reekie, Edward, 188
Rees, Mark, 190–191, 193, 211, 283–284, 287–288

Regeneration niche, 265, 292
Regional diversity
  interactions with local diversity, 556–561
  patterns in, 551
Regional scale, 423
Regulating ecological services, 572
Rejmánek, Marcel, 404–405
Relative competition index (RCI), 270, 283, 286, 287, 288
Relative growth rate, 274
Relative humidity, 457
Relay floristics, 390
Relevé, 354
Rembrandt tulips, 319
Remote sensing
  aggregating patches into "natural units," 427–429
  classification of communities in a landscape, 368
  using to differentiate vegetation, 428
  using to estimate net primary production, 126
Renosterveld, 547
Replication, in science, 11
Reproduction
  introduction, 143, 149
  net reproductive rate, 215
  sexual life cycles, 150–153
  vegetative, 149–150
  *See also* Sexual reproduction
Reproductive allocation, 186–188
Reproductive effort, 187–188
Reproductive isolation
  hybridization and, 255
  polyploidy and, 256
  speciation and, 254–255
Reproductive output, estimating, 210–211
Reproductive phenologies
  abiotic factors, 195–196, *197*
  biotic factors, 196–197
Reproductive value, 186–187, 212, 215–217
Repulse Bay, 154
Rescue effect, 435
Research, primary and secondary, 2–3
Research syntheses, 286–287, 289
Reserves, 445–446
Resistance
  to herbivory, 298, 299
  induced resistance, 314
  systemic resistance, 321
Resistance genes, 321, 322
Resonance transfer, 24
Resource availability
  biological invasions and, 407, 408–409
  succession and, 392–393
Resource competition
  introduction, 262–263
  mechanisms of, 263–266
  plant size and, 266
  between seedlings, 266–270
  *See also* Competition
Resource depletion, 264–265
Resource-ratio hypothesis, 264–265
Resource-ratio model of species coexistence, 291
Resources
  defined, 263
  effects of plant growth on resource acquisition, 146–149
  pollen as a limiting resource, 164, 165
  spatial dispersion of clonal plants and resource uptake, 147–149
  trade-offs and resource allocation, 179
Response surface experiments, 272
Restinga de Jurubatiba National Park, 163
Restionaceae, 547, 559
Restoration ecology
  biodiversity loss and, 573
  community restoration and successional processes, 394–395
  effects of wetland restoration on anthropogenic carbon dioxide emissions, 480–481
  mycorrhizae and, 331
Resurrection plants, 66
Retention time, 114
Return interval, 377
Reveal, James, 545
Reverse fertilization, 394
Rey-Benayas, José, 401, 436, 550, 553
*Rhamnus alaternus, 534*
*Rhexia, 283*
*Rhinocyllus conicus* (flowerhead weevil), 304
*Rhizobium,* 106–107
Rhizomes, 150, 246
Rhizomorphs, 327
*Rhizophora mangle* (red mangrove), *72*
Rhizosphere, 98–99, 108, 125
*Rhododendron,* 529
*Rhododendron traillianum,* 555
*Rhoeo spatifolia, 60*
*Rhus,* 372
*Rhus ovata, 534*
*Rhynchospora, 283, 284*
*Rhynia gwynne-vaughanii,* 498
Rhyniophytes, *496*
Ribulose bisphosphate (RuBP), 26, 33–34
Ribwort plantain (*Plantago lanceolata*), 166, *167,* 219–220, 245, 330
Rice
  *Azolla*-cyanobacterium nitrogen fixation and, 107
  rice farming and greenhouse gases, 482
Richards, M. B., 281
Richards, O., 302
Richardson, David, 404–405
Ricketts, Taylor, 566
Ricklefs, Robert, 412, 550
Rieseberg, Loren, 257
Right eigenvectors, 216
Ring porous trees, 73
Riparian habitats, 381
Roads and road building
  biodiversity loss and, 569, 571
  habitat fragmentation and, 440–441
*Robinia, 552*
Rock cress, 321
Rocks, weathering, 116, 135, 136
Rocky Mountain National Park, *518, 540*
Rocky Mountains, 134, 465, *466,* 471
Rodents, granivory and, 302, 303
Rodman, James, 314
Root, Terry, 489
Root apical meristems, 145
Root cap, *145*
Root exudates, 98, 99
Root hairs, 145
Root suckering, 387
Roots
  adaptations to reduced water availability, 65
  adaptations to water conditions, 71–72
  competition for resources and, 264
  evolution of, 498
  growth, 145
  methods of estimating growth and productivity, 127
  nitrogen uptake, 130
  pathogens, 320
  proteoid roots, 108
  rhizosphere, 98–99
  root-to-shoot ratio in xerophytes, 65
  soil–plant–atmosphere continuum of water movement, 57, 58
  water movement through soil and, 101
*Rorippa,* 405
*Rorippa austriaca,* 405
*Rorippa sylvestris,* 405
*Rorippa* x *armoracioides,* 405
*Rosa acicularis* (Alberta wild rose), *61*
Rosaceae, 150
Rosales, *40*
Rosebud orchid (*Cleistes bifaria*), 209
Rosenzweig, Michael, 186–187, 282, 412, 413
Rosette plants, 144
Ross, M., 269
Rothamsted Experiment Station, 358
Rothstein, David, 28–29
Round-leaved mallow (*Malva pusilla*), 175
Roy, Barbara, 321
*r*-Strategists, 185
  *See also* r- and K-strategy theory
Ruben, Sam, 26

Rubiaceae, 544
Rubisco (RuBP carboxylase/oxygenase)
  $C_4$ photosynthesis, 35–37
  CAM photosynthesis, 38, 39
  photosynthetic rates in spring ephemerals and, *28*, 33
RuBP. *See* Ribulose bisphosphate
*Rubus*, 150, 372
*Rudbeckia hirta* (black-eyed Susan), *155*
Ruderal habit, 186, *187*
Ruderal species, 403–404
Rumen, 316
Ruminants, 316
Runners, 150
Runoff, 99
Russian thistle (*Salsola kali*), *70*
Rusts, 321
Rutaceae, 150
Ryan, Casey, 195
Ryegrass (*Lolium perenne*), 290, 398

## S

*S loci*, 162–163
*S* selection, 186
*Sabal*, *552*
*Sabatia*, *283*
*Saccharum officinarum* (sugarcane), 27, *36*
Sacchi, Christopher, 324
Sacramento Valley, 535
Safe sites, 393
Sagebrush (*Artemisia tridentata*), *511*, *538*
*Sagittaria sagittifolia* (Hawaii arrowhead), *242*
Saguaro (*Carnegiea gigantea*), 45, *64*, 81, 392, 557
Sahara Desert, 537
Sahara sand, 116
Sahel region, 382, 570
Sale, Peter, 291–292
Salguero-Gómez, Roberto, 189, 436–437
Salicaceae, *325*, 327
*Salicornia europaea* (glasswort), *62*, 268
*Salicornia maritima* (slender glasswort), 112
Salinization, 535, 538
Salisbury, Edward, 181
*Salix*, 539, 541
*Salix exigua* (sandbar willow), 381
*Salix fragilis* (brittle willow), 59
*Salix glaucophjylloides* (dune willow), 375
*Salix lasiandra*, 387
*Salix nigra* (black willow), *338*
*Salix scouleriana*, 387
Salk, Carl, 194
Salmon, 128–129
*Salsola kali* (tumbleweed or Russian thistle), *70*
Salt marshes
  ecosystem processes, 112–113
  effects of bird herbivory on Arctic salt marshes, 307
  global distribution, *112*
  hybridization between invasive species and native relatives, 170
  nitrogen cycle in, 130
  productivity, 112–113, 122
Salt pannes, *113*
Saltbush, 255, 256
Saltmeadow cordgrass (*Spartina patens*), 112
*Salvia leucophylla* (purple sage), 277
Salzman, Amy, 148
Sampling, of communities, 354–357
Sampling error, 224
San Bernardino Mountains, 134
San Gabriel Mountains, 134
Sand cherry (*Prunus pumila*), 375
Sand dune communities, 324
Sand fescue (*Festuca rubra* ssp. *arenaria*), 324
Sand particles, 89–90, *91*, *92*
Sandbar willow (*Salix exigua*), 381
Sandwort (*Arenaria serpyllifolia*), 387
Sandy soils, 89, 90–91, *101*
Sangamonian interglacial, *505*
Sankary, Mohamed, 62
Santa Catalina Mountains, *554*
Santalaceae, 317
Santalales, *40*
Sapindales, *40*
*Sapindus*, *552*
*Saponaria ocymoides*, *320*
Saponins, 313, *314*
Saprophytes, 125, 324
*Sarothamnus scoparius* (broom), 302
*Sarracenia purpurea* (purple pitcher plant), 130, *131*
*Sassafras*, *552*
Satake, Akiko, 193
Satellite species, 435
Saturated adiabatic lapse rate, 459
Saturated soil, 99, *100*
*Satureja gilliesii*, *534*
Savanna River Site Corridor Experiment, 444–445
Savannas
  acacia savanna, 536
  direct effects of increasing levels of atmospheric carbon dioxide on, 474
  Mesozoic era, 501
  pine savannas, 536
  *See also* Tropical savanna
Sawtooth goldenbush (*Hazardia squarrosa*), 302
*Saxifraga cernua*, 150
Saxifragales, *40*
Scalariform perforation plates, *73*
Scale
  defined, 422
  explanations for patterns in species diversity and, 561
  grain, extent, and focus, 422–423
  heterogeneity of ecological phenomena and, 13
  interaction with environmental heterogeneity, 424–425
  levels of, 423–424
  manipulative experiments and, 6–8
  methods for analyzing spatial patterns, 426–427
  variability of processes and patterns at different levels of scale, 425–426
Scale dependence, 429
Scale trees, *498*
Scarab beetles, 124
Scarification, 171–172
Scavengers, 124
Schaffer, William, 184, 186–187
Scheiner, Samuel, 17, 352, 436, 550, 553, 564, 565
Schemske, Douglas, 158, 169
Schimper, Andreas, 15
*Schismus barbatus*, *223*, *224*, *225*
*Schizachyrium condensatum*, 379, 408
Schluter, Dolph, 412
Schmalzel, Robert, 213
Schmitt, Johanna, 243–245
Schmitt, Susanne, 386
Schnitzler, Annik, 551
Scholander, P. F., 56
Scholander pressure bombs, 56
Schwartz, Mark, 402
Schwinning, Susanne, 268
Science
  ethics and, 11
  experiments, 5–8
  features of scientific research, 5
  as a human endeavor, 18
  internal and external consistency, 12
  nature of, 2
  open science initiatives, 11–12
  public policy and, 11
  scientific knowledge, 2–4
  specific results and general understanding, 12
  testing of theories, 10–12
Scientific method, 3–4
*Scirpus acutus*, *283*
*Scirpus americanus*, *283*
*Scirpus fluviatilis*, *283*
*Sciurus carolinensis* (gray squirrel), 303
Sclereids, 311
Sclerenchyma, 311
Sclerophyllous leaves, *68*, 70–71, 531
Scolytidae, 301–302
Scotch broom (*Cytisus scoparius*), 219
Scots pine (*Pinus sylvestris*), 399
Screwpines, 145, *146*
Scrophulariaceae, *40*, 66, 313
Scrub oak (*Quercus ilicifolia*), 280, 387
Sea beet (*Beta vulgaris maritima*), 182
Sea coconut (*Lodoicea maldivica*), 179
Sea grasses, 162, 201

Sea ice, global climate change and, 484, *485,* 488
Sea level rise, 484, 488–489
Sea rocket (*Cakile edentula*), *234*
Seasonal tropical forests, 63
Seasons and seasonality
  biomes and, 514
  patterns of variation in precipitation, 465–468
  solar angle of incidence and, 451–452, *453*
  temporal species coexistence and, 554
Secondary chemicals (secondary metabolites)
  evolutionary consequences of plant-herbivore interactions, 316–317
  plant defenses to herbivory, 312–314
  plant defenses to pathogens, 321
Secondary growth, 145, *146*
Secondary research, 2
Secondary succession
  defined, 377
  dispersal capacity of species and, 385–386
  facilitation of colonization, 392
  propagule pool, 387–389
Secondary wall, xylem vessel elements, *74*
Secondary xylem, 145
Sedges, 539, 541
Seed banks
  as a bet-hedging strategy, 190–191
  defined, 174, 201
  estimating seed survival, 175, 209–210
  evolution and, 175
  seed dispersal and, 174–175
  serotiny and, 175
  succession and, 387
Seed dormancy
  bet-hedging life histories, 190–191
  factors ending, 191–192
  plant growth and, 144
Seed ferns, 54, *496,* 499–500
Seed germination
  estimating seed survival, 209–210
  factors triggering, 191–192
  plant growth and, 144
  stabilizing selection and, 269
Seed plants, 54–55, *497,* 499–503
Seed predation and seed predators, 181, 192, 298
Seedlings
  competition, 266–270
  damping-off, 320
  estimations of reproductive output and, 210, 211
  life tables, *208*
  obtaining data for survival studies, 210
  self-thinning, 269–270
Seeds
  agamospermy, 150, 255–256
  angiosperm sexual life cycle, 151, *152,* 153
  bet-hedging life histories, 190–191
  desert annuals, 538
  dispersal, 170–175
  dormancy (*see* Seed dormancy)
  elaiosomes, 173
  fungal pathogens, 319–320
  germination (*see* Seed germination)
  granivory, 302–303
  as individuals, 201
  life tables, *208*
  maternal effects, 245
  methods for estimating seed output, 210–211
  methods for estimating survival probability, 209–210
  migration and genetic variation, 248
  nondispersing, 168–169
  plant adaptation to terrestrial environments and, 54–55
  predation and predators, 181, 192, 298
  propagule pool and succession, 387, 389
  serotiny and, 387, 389
  size hierarchies and, 269
  synchronized flowering and masting, 192–193
  trade-offs between size and number, 179–181
  using genes to track, 238
Seismic events, 382, 384
*Selaginella,* 49
*Selaginella lepidophylla, 66*
Selaginellaceae, 66
Selective logging, 384
Self-incompatibility (SI) systems, 162–163, 167
Self-pollination, 164, 168–169
Self-pruning, 147
Self-thinning, 269–270
Semelparous plants, 177, 183–184, 190–191
Senescence, nitrogen mobilization and, 132
Sensible heat, 75
Sensible heat exchange, 77–78
Sensitivity of population growth, 212, 217–218, 219
Sepals, *31, 154*
Sequential hermaphrodites, 163
*Sequoia,* 500, 502, *552*
*Sequoia sempervirens* (redwood), 470, 530
*Sequoiadendron giganteum* (giant sequoia), *58*
Seral communities, 395
Serengeti National Park, 306
Serengeti Plains, 536
Serotiny, 175, 387, 389
Serpentine barrens, 85
Serpentine soils, 85
Sesquiterpenes, 315
Sessions, Laura, 308
Sexual life cycles, 150–153
Sexual reproduction
  alternation of generations, 150–153
  genetic variation and, 248
  importance of water to early land plants, 54
  mating systems, 162–170 (*see also* Mating systems)
  movement of pollen and, 153–162 (*see also* Pollination)
Sexual selection, 162, 165
Seychelles, *559*
Shade leaves, 47–48
Shade-adapted plants, 49–50
Shadscale (*Atriplex canescens*), *40, 70,* 255, *511*
Shannon-Wiener index, 343, 344
Shmida, Avi, 292
Shoots
  plant growth, 144, 145
  shoot architecture and light interception, 146–147
*Shorea siamensis, 165*
Shortgrass prairie, 379, 519
Shortleaf pine (*Pinus echinata*), 301
Shortwave radiation, 77
Showy lady's slipper (*Cypripedium reginae*), 209
Shreve, Edith, 61
Shreve, Forrest, 64, 81, 222
Shrub oak, 531
Shrublands
  alpine shrubland, *515, 516,* 539
  cold shrubland, *514*
  effects of livestock grazing on, 385
  nurse plants, 279–280
  physiognomy of, 519
  primary productivity, *524*
  *See also* Temperate shrubland
Shrubs
  competitive exclusion and, 281
  geographic distributions of $C_3$ and $C_4$ plants, 44–46
Shurin, Jonathan, 556–557
Siberia, 507
Siberian fir (*Abies sibirica*), 399
Siberian larch (*Larix siberica*), *397,* 399, 409
Siberian pine (*Pinus sibirica*), 399
Siberian spruce (*Picea obovata*), 399
*Sideroxylon grandiflorum* (tambalacoque), 171
Sierra Nevada, 251–253, *466*
*Sigillaria, 498*
Sigmoid curves, 352, 353
Signal-to-noise problem, 366
*Silene acaulis* (cushion plant), 147, 184

*Silene latifolia* (bladder campion), 321, 408
Silica, 311, 314
Silicon, *91, 103*
Silky hakea (*Hakea sericea*), 548
Silt loam, *90*
Silt particles, 89–90
Silty soils, 89
Silurian period, *496,* 497–498
Silva, Juan, 228
Silvertown, Jonathan, 270
Simberloff, Daniel, 402, 435–436
Similarity measures, 361, *362*
*Simmondsia chinensis* (jojoba), *80*
Simple perforation plates, *73*
Simpson's index, 343, 344
Sink population, 434
Siskiyou Mountains, *361*
Sitka spruce (*Picea sitchensis*), 392
Sitka spruce forest, 530
Size
  of a disturbance, 377
  effects of plant size and form variability on population structure, 203–204
  plant size and resource competition, 266
  population structure and, 202, 203
Size hierarchies, 268–269
Size-based methods, 203
Size-structured populations, 203
Skole, David, 441
Skunk cabbage (*Symplocarpus foetidus*), 75, 159
Slack, Roger, 27
Slade, Andrew, 148
Slash pine (*Pinus elliottii*), *123,* 301, 531
Slatyer, Ralph, 391
Slauson, William, 283
Slender glasswort (*Salicornia maritima*), 112
Slippery elm (*Ulmus rubra*), *338*
Slobodkin, Lawrence, 300
SLOSS (single large or several small ) debate, 445–446
Small cordgrass (*Spartina maritima*), 405
Smallwood, Peter, 303
*Smilacina japonica. See* Maianthemum japonicum
Smilanich, Angela, 314
*Smilax glauca* (catbrier), 316
*Sminthurinus elegans, 124*
Smith, Christopher, 180
Smith, Gordon, 160
Smith, Melinda, 406
Smith, Roger, 395
Smithsonian Institution, 441
Smooth cordgrass (*Spartina alterniflora*), 112, *113,* 170
Smuts, 320–321
Snow avalanches, 381–382
Snow buttercup (*Ranunculus adoneus*), 195–196
Snow gum (*Eucalyptus pauciflora*), 302
Snow trillium (*Trillium grandiflorum*), 204, 221, 570
Snow tussock (*Chionochloa pallens*), 193
Social-ecological regions, 570
Sodium, 92, 93–94, *103*
Soil crusts, 96
Soil erosion, 96
Soil food webs, 122–125
Soil microorganisms
  biological immobilization of nitrogen, 132
  denitrification, 131
  net primary production and, 121–122
  nitrogen mineralization and, 129–130
  rhizosphere, 98–99
  soil food webs and decomposition, 122–124
  soil microarthropods, 125
  soil microfauna and mesofauna, 125
Soil order, 87–89
Soil organic matter, 97–98, 125, 127
Soil pH
  defined, 92
  effects on nutrient availability, 92, 93
  effects on plant growth, 92–93
  factors determining, 93–94
  organic matter and, 97
Soil porosity
  clay-dominated soils, 92
  defined, 84
  organic matter and, 97
  water movement through soil and, 99, 100–101
Soil profiles, 94–97
Soil series, 89
Soil solution, 92
Soil structure, 84
Soil texture, 89–92
Soil type, 89
Soil water potential, 99, 100, *101*
Soils
  adaptations to waterlogged soils, 66–67
  calcium cycle and, 137
  classification of, 87–89
  components of, 84
  effects of nutrients on biological invasions, 407
  erosion, 96
  food webs and decomposition, 122–125
  global carbon cycle, 472–473
  global climate change and, 485
  horizons and soil profiles, 94–97
  introduction and overview, 84–85
  movement of water through, 99–101
  mycorrhizae and, 329
  net primary production and, 117–118, 121–122
  nitrogen cycle, 129–130
  nitrogen leaching, 131
  organic and inorganic forms of nitrogen in, 105
  organic matter, 97–98, 125, 127
  origins and formation, 85–87
  permafrost, 88–89, 485, 540
  permanent wilting point, 100
  pH (*see* Soil pH)
  phosphorus as a limiting nutrient for plant growth, 108
  phosphorus cycle and, 135–136
  plant adaptations to frozen soils, 65
  plant nutrition stoichiometry, 104
  plant-soil feedbacks, 278–279
  porosity (*see* Soil porosity)
  primary succession and, 377
  rhizosphere, 98–99
  salinization from irrigation, 535, 538
  serpentine soils, 85
  soil–plant–atmosphere continuum of water movement, 57–59
  successional processes and community restorations, 394–395
  texture and its effect on soil properties, 89–92
  tropical rainforests, 525
  in tundra, 540, 541
  variability in depth, 95
  water potential, 55, 57
Solanaceae, 158
Solanales, *40*
*Solanum demissum, 314*
Solar energy
  alpine regions and, 539
  angle of incidence and, 451–452, *453*
  energy balance of the Earth and, 448–451
  *See also* Light
Solar radiation modification (SRM), 492
Solar tracking, 48–49
*Solidago, 310*
*Solidago altissima* (tall goldenrod), *299*
*Solidgao canadensis* (Canada goldenrod), *61*
Soliveres, Santiago, 294
Solstices, 451, 452, *453,* 454, 461
Somatic mutations, 247, 248
Sonoran Desert
  characteristics of, 537, 538
  environmental stochasticity and population growth of annuals, 223–224, *225*
  facilitation of colonization, 392
  geographic distributions of $C_3$ and $C_4$ plants, 45
  global patterns of air movement and, 459
  physiognomy of hot deserts and, 520
  plant communities and predictability of temperature and precipitation patterns, *471*

reproductive phenology of *Eriogonum abertianum, 197*
seed dispersal and, 174
specialization and species diversity, 415
Sørensen-Dice index, *362*
Sork, Victoria, 169
Source population, 434
South America
deforestation and biodiversity loss, 568–569
Mesozoic era, 499, *500*
South Equatorial Current, *462*
South Island (New Zealand), 484
Southeast Asia, deforestation and biodiversity loss, 568
Southern Oscillation, 470
Southern pine beetle, 301
Soybean (*Glycine max*), *106*
Soybean rust, 320
Spanish moss (*Tillandsia usneoides*), 42–43, *44*
*Sparaganium, 283*
*Spartina,* 170
*Spartina alterniflora* (smooth cordgrass), 112, *113,* 170
*Spartina anglica,* 405
*Spartina foliosa,* 170
*Spartina maritima* (small cordgrass), 405
*Spartina patens* (saltmeadow cordgrass), 112
Spasojevic, Marko, 438–439
Spatial autocorrelation, *249,* 250, 427
Spatial statistics, 427
Spatially explicit approach, to species richness relationships, 350–351
Specialist herbivores, 305
Specialist species, 245–246
Specialization
as a pattern of adaptation, 240–241
species diversity and, 414–415
Speciation
assortative mating and, 169
available energy hypothesis of species richness and, 563
natural selection and, 254–256
through hybridization, 256–257
Species
biological species concept, 254–255
cryptic species, 255
ecotypes, 250–254
extinction debt, 566
geographic isolation and endemics, 557–559
natural selection and the origin of new species, 254–256
niche concept, 246
using genes to identify, 238
Species abundance, measures of, 356
Species coexistence
competition models and, 289–293
effects of mycorrhizae on, 330
modern coexistence theory, 290–291
new research and theoretical challenges, 293–294
Species density
defined, 350, 353
estimating, 350, 355
Species distribution
competitive exclusion and, 281–282
coupled niche demographic models, 434
effects of global climate change on, 489
effects of temperature extremes on, 81
environmental causes of patterns in, 364–365, *366*
transition zones, 553–554
*See also* Geographic distribution
Species diversity
disturbance and, 415–416
effects on ecosystem properties, 417
effects on productivity, 416–417
productivity gradients and, 412–414, 563–566
richness and evenness components, 343
specialization and, 414–415
Species extinctions
glaciation and continental patterns of biodiversity, 551, *552*
global climate change and, 492
introduction, 543
invasive species and, 410–411
Species frequency, estimating, 355
Species identity, biodiversity and, 343
Species interactions
allelopathy, 276–279
facilitation, 279–281
species coexistence, 289–294
species replacement during succession and, 391–392, 396
spectrum of, 273–274
variation along environmental gradients, 282–294
*See also* Competition
Species pool, 375, 556
Species richness
abundance-weighted measures of, 343, 344
average richness in North America over the last 20,000 years, *507*
biological invasions and, 406–407
defined, 342
effects of colonization on changes during succession, 386–387
effects of nitrogen deposition on, 134–135
explanations for global patterns in, 546–550
explanations for patterns along geographic or environmental gradients, 561–566
fynbos, 547–548
gap size and, 378
interactions between regional and local processes affecting, 556–561
inventory diversity and, 342–344
latitudinal diversity gradient, 545–546
measures of, 349–354
patterns in the Cretaceous period, 502
productivity gradients and, 563–566
transition zones, 553–554
variability across the Earth, 544–546
variability with spatial extent, 417–418
*See also* Biodiversity
Species richness relationships (SSRs), 350–354
Species sorting, 520
Species sorting model of metacommunities, 438, 439
Species turnover, 345
Species-area curves, 350–354
Specific heat, 454
Specific leaf area, 47
Specific leaf mass, 47
Specific leaf weight, 47
Spectacle pod (*Dithyrea californica*), 181, 210
Sperm cells, 151
*Sphaeralcea incana* (desert mallow), *70*
*Sphagnum,* 134
Sphenophyta, 498
Spikemoss, 49, *66*
Spikes, *157*
Spiroplasmas, 321
Spodosols, *87, 88, 94*
Spongy mesophyll, *30, 68*
Spores
angiosperm sexual life cycle, 151
palynology, 504–505
plant adaptation to terrestrial environments and, 54
*Sporobolus anglicus. See* Cordgrass
Sporophytes
alternation of generations, 150–151, *152,* 153
early land plants, 498
plant adaptations to terrestrial environments and, 54, 55
Sporophytic SI, 162–163
Spotted knapweed (*Centaurea stoebe*), 278
Spring annuals, 182
Spring ephemerals, 28–29
Springtails, *124,* 125
Spruce beetle, 301
Spruces, 507, *509,* 532
Spurge family, 313
Square quadrats, 355
Squirrels, 303
Sri Lanka, *559*
SRM. *See* Solar radiation modification
SSRs. *See* Species richness relationships
St. Johnswort (*Hypericum fasciculatum*), 160

St. Louis, Excedera, 304
Stabilizing mechanisms, 291
Stabilizing selection, *234,* 235, 239, 269
Stable growth rate, 214
Stable population structure, 213
Stable stage distribution, 213, 223
Stages
  life cycle graphs, 206
  life tables, *208*
  plant demography and, 202
Stage-structured populations, 215–216
Stamens, *154*
Staminate flowers, 163
Standing biomass, 120, 127
Standing crop, 412
Stands, 334
State-structured, 202
Statistics, ecology and, 9–10, 16
Staude, Ingmar, 566
Stebbins, G. Ledyard, 404
Steele, Michael, 303
Stem elongation, 244–245
Stefan-Boltzmann equation, 76
Steppe, 534
Sternberg, Marcelo, 280
Steven, Janet, 153
Stigmas, 151, *152, 154*
*Stigmella, 311*
*Stilbosis, 311*
Stiling, Peter, 311
Stinging nettles, *149,* 310–311
Stipules, 522
Stochasticity, population growth rates and, 223–228
*Stoebe plumosa, 534*
Stohlgren, Thomas, 406
Stoichiometry, 104, 114
Stolons, 150
Stomata
  in CAM photosynthesis, 38–39
  conductance, 32, 60, 67
  distribution and properties in mesophytes and xerophytes, 69, *71*
  dynamic nature of opening, 32, *33*
  evolution of, 41
  leaf temperature and, 79
  responses to water availability, 67–68
  soil–plant–atmosphere continuum of water movement, 57
  stimuli inducing the opening and closing of, 67
  transpiration rate and, 60
Stomatal conductance, 32, 60, 67
Stomatal crypts, 69, *71*
Storage effect, 292
Storm surges, 488–489
Strangler figs, 264, *265,* 525
Strategies
  for adapting to water availability, 60–62
  defined, 60
Stratosphere, *458,* 476
Strauss, Sharon, 318
Strawberries, 150, 247
Streams, *524*
Stress
  anaerobic stress proteins, 67
  Grime's triangular model and, 186, *187*
  leaf rolling in response to water stress, 80
  *r*- and *K*-strategy theory and, 185–186
  stress-gradient hypothesis, 287
Strict additivity, 237
Stroma, 24–25, *27*
Stroma lamellae (stroma thylakoids), 22, *23*
Strong, Donald, 303
Structural coloration, 49, 50
Structured populations
  age-structured populations, 203
  models of population growth, 211–213
  size-structured populations, 203
  stage-structured populations, 215–216
*Strychnos mitis,* 320
Styles, *154,* 167–169
*Stylocline micropoides, 223, 224, 225*
*Suaeda maritima* (herbaceous seepweed), 112
Sub clover (*Trifolium subterraneum*), *267,* 330
Subalpine forest, 532
Substitutive designs, 271–272
Substomatal cavities, *68*
Subterranean trefoil (*Trifolium subterraneum*), *267,* 330
Subtropical areas, 516–517
Subtropical forest, 517
Subtropical jet stream, 462
Suburbanization, 572
Succession
  abandoned agricultural field example, 372
  Clements-Gleason dichotomy, 372–374
  climax concept, 395
  community restoration and, 394–395
  conflicting views of, 372
  current theory of, 374, *375*
  defined, 371–372
  dispersal capacity of species and, 385–387
  disturbance and, 371–372, 377–385 (*see also* Disturbance)
  dynamic equilibrium of communities, 395–396
  effects of habitat fragmentation on, 444–445
  effects of pathogen interactions on, 324
  filter model of, *374*
  general causes of, 377
  history of the development of modern succession theory, 373–374
  life histories and, 390–391
  predictable and unpredictable trajectories, 393–394
  primary and secondary, 377
  propagule pools, 387–389
  resource availability and, 392–393
  role of methodology in understanding the process of, 374–376
  species interactions and species replacement during, 391–392, 396
  vegetation dynamics, 374
Succulent Karoo, 547
Succulents
  adaptations to reduced water availability, 64–65
  CAM photosynthesis and, 39
  leaf water potential and stomatal behavior, *68*
  secondary growth, 145, *146*
  thorn forests, 527
Suckers, 150
Sudden oak death, 323
Suding, Katharine, 278
Sugar maple (*Acer saccharum*)
  altitudinal gradient in species frequencies, *339*
  arbuscular mycorrhizae, 326
  changes in importance along a moisture gradient, *338*
  impact of acid precipitation on, 132
  Pleistocene glaciation and, 507
  succession and, 375, 393
  temperate deciduous forests, *529*
Sugar pine (*Pinus lambertiana*), *73*
Sugarcane (*Saccharum officinarum*), 27, *36*
Suicide tree (*Tachigali versicolor*), *172,* 303
*Suillus brevipes, 328*
Sulfate, 136
Sulfur
  anthropogenic sources of, 133
  atmospheric deposition of, 132, 136
  global sulfur cycle, 115, 136
  plant nutrition, *102*
  sources of, 116
Sulfur dioxide, 132, 136
Sulfuric acid, 132
Sumacs, 372
Summer annuals, 183
Sun leaves, 47–48
Sundarbans, 489
Sunflecks, 49–50
Sunflower family. *See* Asteraceae
Sunflowers, 180
Sunlight. *See* Light; Radiant energy; Solar energy
Superior organism hypothesis, 402–403
Superorganism concept, 336–337, 391
  *See also* Clementsian ecology

Supervised classification, 368
"Superweeds," 169–170
Supporting ecological services, 572
Surface fires, 377, *378*
Survival probability
  life tables, 208
  methods of estimating, 207–210
  obtaining data for survival studies, 210
Survivorship, 184, 234, 240–241
Swamps, *524*
Sweetgum, 528
Swiss Alps, 482, 484
Sydney bluegum (*Eucalyptus saligna*), 279, *280*
Sykes, Martin, 290
Symbioses
  definition and types of, 105
  endophytes, 324–325
  insects with plant pathogens, 324
  mycorrhizae, 325–331 (*see also* Mycorrhizae)
  nitrogen fixing symbioses, 105–108
  ruminants with cellulose-digesting bacteria, 316
Sympatric speciation, 254
*Symplocarpus foetidus* (skunk cabbage), 75, 159
Synecology, 17
*Syringa vulgaris* (common lilac), 409
Systematic reviews, 289
Systemic resistance, 321

## T

Tabebuia (*Tabebuia guayacan*), *63*
*Tachigali versicolor* (suicide tree), *172*, 303
*Tagetes patula* (marigold), *268*
Tags, 210
Taiga
  characteristics of, 532–533
  dominant growth forms and general climatic conditions, *516*
  global distribution, *514*
  impact of human activity on, 533
  low species richness, 399
  physiognomy, 516
  primary productivity, *524*
  temperature and precipitation, *515*
  *See also* Boreal forest
Tailings, heavy-metal tolerance and, 241–243, 329, 521
Takenaka, Akio, 30
Tall bellflower (*Campanula americana*), 220
Tall goldenrod (*Solidago altissima*), *299*
Tallgrass prairie
  bimodal patch occupancy, 435–436, *437*
  effects of large mammals on community composition, 306
  physiognomy, 519
  soil, succession, and community restoration, 394
Tamarack, 194, 532
Tambalacoque (*Sideroxylon grandiflorum*), 171
Tanner, Edmund, 273
Tannins, 312–313
Tanoak (*Notholithocarpus densiflorus*), 323
Tansley, Sir Arthur, 113, 281, 308
Taproots, 71
Tarangire National Park, *536*
*Taraxacum. See* Dandelions
Taxa, 335
Taxodiaceae, 502
*Taxodium, 552*
*Taxodium distichum* (bald cypress), *66*
Taxonomic guilds, 335
Taxonomic species, 255
*Taxus baccata, 122*
TBMs. *See* Terrestrial biosphere models
Teak, 570
*Tectona,* 570
*Tegeticula,* 159
Temperate alpine regions, 539
Temperate deciduous forest
  characteristics of, 528–529
  dominant growth forms and general climatic conditions, *516*
  global distribution, *514*
  impact of human activity on, 529
  primary productivity, *524*
  temperature and precipitation, *515*
  vegetative phenology of herbaceous plants, 194–195
Temperate evergreen forest
  characteristics of, 530–531
  convergent evolution and, 521
  dominant growth forms and general climatic conditions, *516*
  global distribution, *514*
  primary productivity, *524*
  temperature and precipitation, *515*
Temperate forests
  competition for light during succession, 393
  Cretaceous period, 502
  land use histories and succession in hardwood forests, 394
  pathways of water loss, *139*
Temperate grassland
  characteristics of and impact of human activity on, 534–535
  dominant growth forms and general climatic conditions, *516*
  global distribution, *514*
  primary productivity, *524*
  temperature and precipitation, *515*
  *See also* Prairies
Temperate habitats, temperature extremes for leaf damage, *81*
Temperate rainforest
  characteristics of, 529–530
  deforestation and anthropogenic carbon emissions, 482
  dominant growth forms and general climatic conditions, *516*
  impact of human activity on, 530
  temperature and precipitation, *515*
Temperate shrubland
  characteristics of, 533, *534*
  dominant growth forms and general climatic conditions, *516*
  global distribution, *514*
  impact of human activity on, 533
  temperature and precipitation, *515*
Temperate woodland, *515, 516,* 531, *532*
Temperature
  activity limit and lethal limit, 80
  adaptations to temperature extremes, 80–81
  biomes and, 514, *515*
  causes of seasonal variation across the globe, 451–452, *453*
  climate and, 448
  climate diagrams, *467*
  Cretaceous period, 501
  defined, 448
  energy balance of the Earth and, 448–451
  factors affecting the magnitude of daily variation in, 454–455
  factors determining leaf and whole-plant temperature, 78–80
  forest physiognomy and, 517, 518
  global climate change and, 483–484, 486–488
  mean global temperatures in January and July, *454*
  net primary production and, *117*
  optimum for photosynthesis, *36*
  over geological time periods, *505*
  during the Pleistocene glacial maximum, 507
  reproductive phenologies and, 195–196
  shrublands and, 519
  significance for plant ecology and evolution of predictability in patterns, 470–471
  Stephan-Boltzmann equation, 76
  temperature lag, 452, 454–455
  tree line and, 518
  vegetative phenologies and, 194–195
  *See also* Cold temperature; Thermal energy balance
Temperature lag, 452, 454–455
Temporal niches, 292
ter Steege, Hans, 400
Termites, 124, 536, *537*
Terpenes, 312, 313, *314,* 315, 316
Terrestrial biosphere models (TBMs), 119
Tertiary period, *496,* 503

Tetraploidy, 255
Thailand, *569*
*Thalassia,* 162
*Thalictrum,* 153
*Thalictrum alpinum,* 153
*Thalictrum dioicum,* 153
*Thalictrum fendleri,* 153
*Thalictrum sparsiflorum,* 153
Thar Desert, 459
Thatching grass (*Hyparrhenia hirta*), 274–275
THC. *See* Thermohaline circulation
Theodose, Theresa, 286
Theories, 2, 3–4, 10–12
Theory of island biogeography, 431–433
  *See also* Island biogeography theory
Thermal energy balance
  energy flow via conduction and convection, 77–78
  factors determining leaf and whole-plant temperature, 78–80
  introduction and overview, 53–54, 75
  plant adaptations to temperature extremes, 80–81
  radiant energy exchange, 76–77
Thermohaline circulation (THC), 463–464, 488
*Thermopsis montana* (golden banner), *47*
–3/2 Thinning law, 269–270
Thistle, 206, *207*
Thomas, Meirion, 27
Thomson, James, 164
Thomson, Vivien, 407
Thorn forest
  characteristics of, 527, *528*
  dominant growth forms and general climatic conditions, *516*
  global distribution, *514*
  temperature and precipitation, *515*
Thornwoods, 527, *528*
Threadleaf crowfoot (*Ranunculus trichophyllus*), *242*
Threatened species, concept of rarity and, 398–399
Threshold effects, 430
Throughflow, 137
*Thuja,* 507, *552*
Thylakoid disks, 22, *23,* 24
*Tiarella cordifolia* (foamflower), *28,* 29, 33, 48
*Tilia,* 528
*Tilia americana* (basswood), 375, 391, 529
*Tilia heterophylla* (white basswood), *339*
*Tillandsia usneoides* (Spanish moss), 42–43, *44*
Tillers, 246
Tilman, David, 264–265, 274, 275, 282, 288, 291, 376, 412, 414, 416, 417
Time-based species richness relationships, 354
Tobacco (*Nicotiana tabacum*), *23, 313,* 315
Tolerance
  of colonization during succession, 391, 392
  of a disease organism, 322
  drought-tolerance, 64–65
  heavy-metal tolerance, 241–243, 329, 521
  to herbivory, 298–299
  to temperature extremes, 80, 81
  tolerance model of succession, 390
Tong, Florence, 571
Top-down population control, 300, 301
Tornadoes, 381, *382*
*Torreya, 552*
Toxins, allelopathy and, 276–279
Tracheids, 73
Trade winds, 460–461, 469
Trade-offs
  causes of, 178
  difficulty in measuring, 178–179
  emergence of competitive strategies and, 274–275
  in heavy-metal tolerance, 241
  life histories and, 178–181
  between seed size and number, 179–181
Traits
  evolutionary change by natural selection, 232–235 (*see also* Evolution; Natural selection)
  trade-offs and the emergence of competitive strategies, 274–275
  trait diversity, 341–342
  trait function and biodiversity, 342
  worldwide leaf economic spectrum, 275, 276
Transects, 251, 355
Transgenerational plasticity, 245
Transgenes, 169–170
Transient analysis, 215
Transition matrix models, 212, 213, 216, 290
Transition probabilities, 219
Transition zones, 553–554
Transpiration
  latent heat loss and, 78
  leaf temperature and, 79
  soil–plant–atmosphere continuum of water movement, 58
  water cycle and, 138–139
  water loss and, 59–60
  water movement through soil and, 100–101
Transpiration rates, 56, 59–60
Transportation, carbon emissions and, 480
Tree ferns, *386*
Tree line, 518, *519*
Trees
  architecture and light interception, 146–147
  common mycorrhizal networks, 281
  continental patterns in species richness, 552–553
  dendrochronology, 376
  diameter at breast height, 202, 357
  direct effects of increasing levels of atmospheric carbon dioxide on, 474
  ectomycorhhizae, 327
  effects of herbivory on, 301–302
  effects of land use on dispersal and migration, 572
  emergence of modern communities following Pleistocene glaciation, 508–510
  estimating survival probability, 207
  facilitation and, 279–280, 281
  obtaining data for survival studies, 210
  overgrowth by vines, 264, *265*
  self-pruning, 147
  thick bark and fire, 391
  trade-off between wood growth and toughness, 179
Trembling aspen. *See* Populus tremuloides
Trenching, 273
*Trevoa trinervis, 534*
Triantis, Kostas, 433
Triassic period, *496*
*Trichocereus pascana, 199*
Trichomes, *70, 131,* 310–311
*Trientalis latifolia, 361*
*Trifolium pratense* (purple clover), 324
*Trifolium repens* (white clover), 290, 315
*Trifolium subterraneum* (sub clover or subterannean trefoil), *267,* 330
*Trillium grandiflorum* (great white or snow trillium), 204, 221, 570
*Tristerix aphyllus,* 317
Tristyly, 167–169
Triterpenes, *314*
*Triticum aestivum* (wheat), *72*
Trochodendrales, *40*
Troll, Carl, 421
Trophic cascades, 160, 309
Trophic interactions
  herbivory (*see* Herbivory)
  overview, 297
  parasitic plants, 317–318
Tropic of Cancer, 458, 461
Tropic of Capricorn, 458, 461
Tropical alpine regions, 539
Tropical deciduous forest
  characteristics of, 526–527
  deforestation and biodiversity loss, 568
  dominant growth forms and general climatic conditions, *516*
  drought-deciduous phenologies, 195
  global distribution, *514*
  primary productivity, *524*
  temperature and precipitation, *515*

Tropical forests
anthropogenic carbon emissions and, 482
biodiversity loss and, 566
Cretaceous period, 502
deforestation and biodiversity loss, 568–569
effects of human activities on, 385
moist, 523–526
physiognomy, 517
rarity and, 400
seasonal, 526–527
wind as disturbance, 381
Tropical habitats, temperature extremes for leaf damage, *81*
Tropical mangrove forest, 489
Tropical montane forest, *515, 516,* 526
Tropical rainforest
acidic soils, 94
characteristics of, 523–526
deforestation and biodiversity loss, 568–569
deterministic succession in, 394
dominant growth forms and general climatic conditions, *516*
drought-deciduous phenologies, 195
global distribution, *514*
Hadley cells and precipitation, 459
human population growth and biodiversity loss, 571
impact of human activity on, 385, 525–526
pathways of water loss, *139*
plant pathogens and, 322
Pleistocene glaciation and, 507
temperature and precipitation, *515*
*See also* Amazon Basin
Tropical savanna
characteristics of, 536, *537*
dominant growth forms and general climatic conditions, *516*
global distribution, *514*
impact of human activity on, 536
primary productivity, *524*
temperature and precipitation, *515*
Tropical woodland, *515, 516,* 527–528
Tropics
latitudinal diversity gradient, 545–546, 548–549
seasonal patterns in precipitation, 468
species richness in the Cretaceous period, 502
Troposphere, 455, 458, 476
True yams, 150
TRY database, 276
*Tsuga, 552*
*Tsuga canadensis* (eastern hemlock), 309, *339,* 507, 509–510, 529
*Tsuga heterophylla* (western hemlock), 387
Tsuyuzaki, Shiro, 280
Tuck, Sean, 293
Tucker, Compton, 441
Tulip poplar (*Liriodendron tulipifera*), *73,* 326, *339,* 529
Tulips, 150, 319
Tumbleweed (*Salsola kali*), *70*
Tundra
characteristics of, 540–541
dominant growth forms and general climatic conditions, *516*
effects of bird herbivory on salt marshes, 307
effects of global climate change on, 491
impact of human activity on, 541
Pleistocene glaciation and, 507
primary productivity, *524*
temperature and precipitation, *515*
Tuomisto, Hanna, 344, 345
Turesson, Göte, 251
Turner, Raymond, 221–222
Turnip (*Brassica septiceps*), *30*
Turnover time, 97–98, 114, 120–121
Turtle grass, 162
Tussock tundra, 541
Tussocks, 425–426
Twin spined cactus (*Mammillaria geminispina*), *43*
Two-species substitutive experiments, 276
Tyndall, John, 477
Type I species richness relationships, 352, *353*
Type II species richness relationships, *351, 353,* 354
Type III species richness relationships, *351, 353,* 354
Type IV species richness relationships, 352, *353*
*Typha,* 247
*Typha angustifolia,* 282
*Typha latifolia,* 283
Typhoons, 381
Typological thinking, 186
Tyukavina, Alexandra, 482

## U

Uesaka, Shoheo, 280
Uinta Mountains, *466*
*Ulex europaeus* (gorse), *410*
Ulmaceae, 502
*Ulmus,* 502, 508, *509, 510,* 528
*Ulmus americana* (American elm), 324, *338*
*Ulmus rubra* (slippery elm), *338*
Ultisols, *87, 88*
Ultraviolet (UV) radiation
alpine regions and, 539
ozone hole and, 476
UVB radiation, 30–31
Uluru-Kata Tjuta National Park, *528*
Umbels, *157*
Understory plants, 49–50
Ungulates, 536
Unified theory, 3–4
United Kingdom, carbon emissions, 479
United States
agriculture and greenhouse gases, 482
carbon emissions, 478–479, 480, 481
Univariate statistics, 359, 360
Unsupervised classification, 368
UNSVC. *See* U.S National Vegetation Classification system
Urban, Dean, 430, *431*
Urban ecology, 18
Urbanization, biodiversity loss and, 567, 572
Urban-wildland interface, 572
*Uromyces trifolii,* 324
*Ursinia pinnata, 534*
*Urtica,* 150, 310–311
*Urtica dioica, 149, 310*
Urticaceae, 310–311
U.S. LTER Network, 358
U.S National Vegetation Classification (USNVC) system, 341, 367
*Ustilago violacea,* 321
UV radiation. *See* Ultraviolet radiation

## V

*Vaccinium,* 158
Valley oak (*Quercus lobata*), 169
*Vallisneria,* 161–162
Valone, Thomas, 395
Valverde, Javier, 160
van der Maarel, Eddy, 290, 358
van Dijk, Henk, 182
van Kleunen, Mark, 148
van Noordwijk, Arie, 179
VanderPutten, Wim, 324
Vandvik, Vigdis, 491
Vapor pressure deficit, 59–60
Variance partitioning, 239
Vascular bundles, *68, 70*
Vascular cambium, 145
Vascular plants, 54–55, *497,* 498
Vascular tissue, 54
Vegetation dynamics, 374
Vegetation lag, 511
Vegetative cell, 151
Vegetative phenology, 194–195
Vegetative propagules, 387
Vegetative reproduction, 149–150, 201
Veldt, 534
Venable, D. Lawrence, 181, 190, 192, 223–224, *225*
*Verbascum blattaria* (moth mullein), 175
*Verbascum thapsus* (common mullein), 175
Vernalization, 182
Vertical migration, 511
Vertisols, *87, 88*

Vilà, Montserrat, 410
Vines, 264, *265*
*Viola,* 169
*Viola lobata, 361*
*Viola pubescens* (downy yellow violet), *28,* 29, 33, 48
Virulence, 321
Viruses
broken tulips and, 319
pathogenic, 318, 320
Vital rates, 200, 207–211
Vitales, *40*
Vitousek, Peter, 407, 417
Vivanco, Jorge, 278
VOC. *See* Volatile organic carbon
Vogler, Donald, 163
Volatile hydrocarbons, 476
Volatile organic carbon (VOC), *118,* 119
Volatilization, 131, 133
Volcanoes, 382, 384, 386, 387
Voles, 309
*Vombatus ursinus* (wombat), 508
von Humboldt, Alexander, 14, 15, 522–523, 544
Voodoo lily (*Amorphophallus titanum*), 75, 159

## W

Wagner, David, 161
Walker, Lawrence, 407
Wallace, Alfred Russel, 232
Waller, Donald, 153
Wallflowers, 160
Walnuts, 528
Waloff, Nadia, 302
Wand, Stephanie, 474
Wang, Guiling, 58–59
Wang, Ji-Hua, 554–555
Wang, Xianzhong, 474
Wardle, David, 398, 416
Wardle, Glenda, 220
Warm deserts. *See* Hot deserts
Warming, J. Eugenius, 15, 337
Warner, Robert, 292
Wasatch Mountains, *466*
Waser, Nickolas, 159
Wasp pollination, 159
Water
disturbance and, 381–382
hydrogen bonds, *93*
movement through soil, 99–101
plant competition for, 263
soil–plant–atmosphere continuum of water movement, 58
split during photosynthesis, 24
water content of plants, 102
water cycles, 137–139
*See also* Water relations
Water hyacinth (*Eichhornia crassipes*), 150
Water lilies, 161
Water molds, 318, 319, 320, 323
Water potential
components of, 55, 57
defined, 55
measuring, 56–57
soil–plant–atmosphere continuum of water movement, 57–59
water movement through soil and, 99, 100
Water relations
adaptations for reduced water availability, 62–65
evolution of land plants and, 54–55
introduction, 53–54
leaf adaptations in arid environments, 68–71
mycorrhizae and, 329
physiological adaptations to drought, 65–67
root adaptations to water conditions, 71–72
soil–plant–atmosphere continuum of water movement, 57–59
strategies for adapting to water availability, 60–62
water loss and latent heat loss, 78
water loss and stomatal responses, 67–68
water loss and transpiration, 59–60
water use efficiency, 62 (*see also* Water use efficiency)
xylem and, 73–74
Water use efficiency (WUE)
$C_4$ plants, 37
CAM plants, 39
definition and description of, 62
direct effects of increasing levels of atmospheric carbon dioxide on, 474
measuring, 57
Water vapor, 59–60, 475–476
Waterlogged soils, 66–67, 72
Watkins, James, 555
Watkinson, Andrew, 267, 289
Watt, Alexander S., 290, 373
Weather
defined, 448
distinguished from climate, 448
global climate change and, 483, 487–488
history of the study of, 15–16
Weathering, 116, 135, 136
Weaver, John, 274, 373
Wedgeleaf eryngium (*Eryngium cuneifolium*), 228
Weedy habit, 186, *187*
Weedy species, 403, 404
Weigelt, Alexandra, 271
Weighted species endemism, 557, *558*
Weiner, Jacob, 268
Welden, Charles, 283
Weltzin, Jake, 299
*Welwitschia mirabilis,* 147
Werner, Patricia, 266
West Wind Drift, *462*
Westerlies, *458,* 461
Western balsam bark beetle, 301
Western Cape region, 560
Western hemlock (*Tsuga heterophylla*), 387
Western pine beetle, 301
Westoby, Mark, 275, 276
Wet deposition of nitrogen, 132, 133
Wet sedge tundra, 541
Wetlands
competitive strategies, 274
effects of restorations on anthropogenic carbon dioxide emissions, 480–481
nitrogen cycle in, 130–131
Wheat (*Triticum aestivum*), *72*
Whisk ferns, 498
White, J., 270
White, Peter, 377
White ash (*Fraxinus americana*), *339*
White aster (*Aster ericoides*), 372
White basswood (*Tilia heterophylla*), *339*
White clover (*Trifolium repens*), 290, 315
White fir (*Abies concolor*), 430, *431, 511*
White oak (*Quercus alba*), 303, 327, *338*
White pine. *See* Eastern white pine
White spruce (*Picea glauca*), 154, 532
Whitebark pine (*Pinus albicaulis*), 172–173, 301–302
Whitehead, Michael, 164
White-tailed deer (*Odocoileus virginianus*), 298, *299,* 309–310, 316, 570
Whittaker, Robert H., 119, 337–338, 342, 344, 364
Whittaker, Robert J., 386, 433
Whole-canopy carbon flux, *56*
Whole-plant temperature, 78–80
Wild carrot (*Daucus carota*), 171
Wild leek (*Allium tricoccum*), 28–29, 33
Wild oats (*Avena fatua*), 323–324
Wild poinsettia (*Euphorbia pulcherrima*), *40,* 156, 522
Wild radish (*Raphanus sativus*), 160
Wild rye (*Elymus canadensis*), 375, 392, 396
Willig, Michael, 17, 562, 564
Willows, 387, 539, 541
Wilson, Chester, 268–269
Wilson, Edward O., 185, 431, 432, 433
Wilson, J. Bastow, 291
Wilson, Mary, 165
Wilson, Scott, 284
Wilting coefficient, 100
Wind
as disturbance, 381, 382
seed dispersal and, 171, 173
trade winds, 460–461, 469
Wind chill, 77–78
Wind pollination, 153–154, 166
Windthrows, 381, 382

Winter annuals, 182
Wisconsin glaciation, 506
Witches brooms, 321
Wog Wog experiment, 445
Wolfe, Lorne, 408
*Wollemia nobilis* (Wollemi pine), 557
Wolves, 306–307
Wombat (*Vombatus ursinus*), 508
Wood, 145, 179, 498
Woodlands
  oak-pine woodlands, 531
  physiognomy, 519
  pinyon-juniper woodland, 531, *532*
  primary productivity, *524*
  temperate woodland, *515, 516,* 531, *532*
  tropical woodland, *515, 516,* 527–528
Woodwell, George, 119
World War I, 134
Worldwide leaf economic spectrum, 275, 276
Wright, David, 412–413
Wright, Sewall, 434
Wright, Shane, 563
WUE. *See* Water use efficiency (WUE)
Wulf, Andrea, 14

## X

*Xanthium strumarium* (cocklebur), *171*
*Xanthomonas axonopodis* pathovar *citri,* 320
Xanthophylls, 22
Xeric environments, 62–65
*Xerophyllum tenax, 361*
Xerophytes
  adaptations to xeric environments, 62–65
  behavior of stomata in, 68
  defined, 61
  drought-tolerance, 64–65
  factors determining whole-plant temperature, 79–80
  leaf structure and anatomy, 69–71
  xylem, 73
Xylem
  cavitation, 74
  description and characteristics of, 73, *74*
  mesophyte leaves, *68*
  plant adaptation to terrestrial environments and, 54
  soil–plant–atmosphere continuum of water movement, 57, 58
Xylem vessel elements, 73, *74*
Xylem vessels, 73
*Xylocarpus granatum* (dhundal), 489
*Xyris, 283*

## Y

Yamuna Biodiversity Park, 573
Yareta (*Azorella compacta*), *539*
Yarrow, 251–253, 398
Yellow birch (*Betula allegheniensis*), *339*
Yellow buckeye (*Aesculus octandra*), *339*
Yellow bush lupine (*Lupinus arboreus*), 303
Yellow pond lily (*Nuphar lutea*), *242*
*YELLOW UPPER* gene, 169
Yellow water lily, *72*
Yellowcress, 405
Yellows diseases, 321
Yellowstone National Park, 306–307, 380
Yoda, Kyoji, 269–270
Yohe, Gary, 489
*Yucca brevifolia* (Joshua tree), *53*
Yuccas, *53,* 64–65, *71,* 159

## Z

Zak, Donald, 28–29
Zalucki, Myron, 316–317
Zeng, Zengzhong, 568
Zinc, 97, *103*
Zingiberales, *40*
Zooplankton, 473
*Zostera,* 162, 201
ZTL blue-light receptor, 47
Zucchini (*Cucurbita pepo*), 165
Zurich-Montpellier school, 354, 431
Zygomycota, *325*
Zygophyllales, *40*
Zygote, 151, *152*