Ecology

From Individuals to Ecosystems

Michael Begon, Colin R. Townsend,
and John L. Harper

ECOLOGY
From Individuals to Ecosystems

ECOLOGY
From Individuals to Ecosystems

MICHAEL BEGON

School of Biological Sciences,
The University of Liverpool, Liverpool, UK

COLIN R. TOWNSEND

Department of Zoology, University of Otago, Dunedin, New Zealand

JOHN L. HARPER

Chapel Road, Brampford Speke, Exeter, UK

FOURTH EDITION

Blackwell
Publishing

Contents

Part 3: Communities and Ecosystems

Color plate section between pp. 000 and 000

Preface

A science for everybody – but not an easy science

This book is about the distribution and abundance of different types of organism, and about the physical, chemical but especially the biological features and interactions that determine these distributions and abundances.

Unlike some other sciences, the subject matter of ecology is apparent to everybody: most people have observed and pondered nature, and in this sense most people are ecologists of sorts. But ecology is not an easy science. It must deal explicitly with three levels of the biological hierarchy – the organisms, the populations of organisms, and the communities of populations – and, as we shall see, it ignores at its peril the details of the biology of individuals, or the pervading influences of historical, evolutionary and geological events. It feeds on advances in our knowledge of biochemistry, behavior, climatology, plate tectonics and so on, but it feeds back to our understanding of vast areas of biology too. If, as T. H. Dobzhansky said, 'Nothing in biology makes sense, except in the light of evolution', then, equally, very little in evolution, and hence in biology as a whole, makes sense except in the light of ecology.

Ecology has the distinction of being peculiarly confronted with uniqueness: millions of different species, countless billions of genetically distinct individuals, all living and interacting in a varied and ever-changing world. The challenge of ecology is to develop an understanding of very basic and apparent problems, in a way that recognizes this uniqueness and complexity, but seeks patterns and predictions within this complexity rather than being swamped by it. As L. C. Birch has pointed out, Whitehead's recipe for science is never more apposite than when applied to ecology: seek simplicity, but distrust it.

Nineteen years on: applied ecology has come of age

This fourth edition comes fully 9 years after its immediate predecessor and 19 years after the first edition. Much has changed – in ecology, in the world around us, and even (strange to report!) in we authors. The Preface to the first edition began: 'As the cave painting on the front cover of this book implies, ecology, if not the oldest profession, is probably the oldest science', followed by a justification that argued that the most primitive humans had to understand, as a matter of necessity, the dynamics of the environment in which they lived. Nineteen years on, we have tried to capture in our cover design both how much and how little has changed. The cave painting has given way to its modern equivalent: urban graffiti. As a species, we are still driven to broadcast our feelings graphically and publicly for others to see. But simple, factual depictions have given way to urgent statements of frustration and aggression. The human subjects are no longer mere participants but either perpetrators or victims.

Of course, it has taken more than 19 years to move from man-the-cave-painter to man-the-graffiti-artist. But 19 years ago it seemed acceptable for ecologists to hold a comfortable, objective, not to say aloof position, in which the animals and plants around us were simply material for which we sought a scientific understanding. Now, we must accept the immediacy of the environmental problems that threaten us and the responsibility of ecologists to come in from the sidelines and play their full part in addressing these problems. Applying ecological principles is not only a practical necessity, but also as scientifically challenging as deriving those principles in the first place, and we have included three new 'applied' chapters in this edition, organized around the

three sections of the book: applications at the level of individual organisms and of single-species populations, of species inter-actions, and of whole communities and ecosystems. But we remain wedded to the belief that environmental action can only ever be as sound as the ecological principles on which it is based. Hence, while the remaining chapters are still largely about the principles themselves rather than their application, we believe that the *whole* of this book is aimed at improving preparedness for addressing the environmental problems of the new millennium.

Ecology's ecological niche

We would be poor ecologists indeed if we did not believe that the principles of ecology apply to all facets of the world around us and all aspects of human endeavor. So, when we wrote the first edition of *Ecology*, it was a generalist book, designed to overcome the opposition of all competing textbooks. Much more recently, we have been persuaded to use our 'big book' as a springboard to produce a smaller, less demanding text, *Essentials of Ecology* (also published by Blackwell Publishing!), aimed especially at the first year of a degree program and at those who may, at that stage, be taking the only ecology course they will ever take.

This, in turn, has allowed us to engineer a certain amount of 'niche differentiation'. With the first years covered by *Essentials*, we have been freer to attempt to make this fourth edition an up-to-date guide to ecology *now* (or, at least, when it was written). To this end, the results from around 800 studies have been newly incorporated into the text, most of them published since the third edition. None the less, we have shortened the text by around 15%, mindful that for many, previous editions have become increasingly overwhelming, and that, clichéd as it may be, less is often more. We have also consciously attempted, while including so much modern work, to avoid bandwagons that seem likely to have run into the buffers by the time many will be using the book. Of course, we may also, sadly, have excluded bandwagons that go on to fulfil their promise.

Having said this, we hope, still, that this edition will be of value to all those whose degree program includes ecology and all who are, in some way, practicing ecologists. Certain aspects of the subject, particularly the mathematical ones, will prove difficult for some, but our coverage is designed to ensure that wherever our readers' strengths lie – in the field or laboratory, in theory or in practice – a balanced and up-to-date view should emerge.

Different chapters of this book contain different proportions of descriptive natural history, physiology, behavior, rigorous laboratory and field experimentation, careful field monitoring and censusing, and mathematical modeling (a form of simplicity that it is essential to seek but equally essential to distrust). These varying proportions to some extent reflect the progress made in different areas. They also reflect intrinsic differences in various aspects of ecology. Whatever progress is made, ecology will remain a meeting-ground for the naturalist, the experimentalist, the field biologist and the mathematical modeler. We believe that all ecologists should to some extent try to combine all these facets.

Technical and pedagogical features

One technical feature we have retained in the book is the incor-poration of marginal es as signposts throughout the text. These, we hope, will serve a number of purposes. In the first place, they constitute a series of subheadings highlighting the detailed struc-ture of the text. However, because they are numerous and often informative in their own right, they can also be read in sequence along with the conventional subheadings, as an outline of each chapter. They should act too as a revision aid for students – indeed, they are similar to the annotations that students themselves often add to their textbooks. Finally, because the marginal notes generally summarize the take-home message of the paragraph or paragraphs that they accompany, they can act as a continuous assessment of comprehension: if you can see that the signpost is the take-home message of what you have just read, then you have understood. For this edition, though, we have also added a brief summary to each chapter, that, we hope, may allow readers to either orient and prepare themselves before they embark on the chapter or to remind themselves where they have just been.

So: to summarize and, to a degree, reiterate some key features of this fourth edition, they are:

- marginal notes throughout the text
- summaries of all chapters
- around 800 newly-incorporated studies
- three new chapters on applied ecology
- a reduction in overall length of around 15%
- a dedicated website (www.blackwellpublishing.com/begon), twinned with that for *Essentials of Ecology*, including inter-active mathematical models, an extensive glossary, copies of artwork in the text, and links to other ecological sites
- an up-dating and redrawing of all artwork, which is also avail-able to teachers on a CD-ROM for ease of incorporation into lecture material.

Acknowledgements

Finally, perhaps the most profound alteration to the construction of this book in its fourth edition is that the revision has been the work of two rather than three of us. John Harper has very rea-sonably decided that the attractions of retirement and grand-fatherhood outweigh those of textbook co-authorship. For the two of us who remain, there is just one benefit: it allows us to record publicly not only what a great pleasure it has been to have

collaborated with John over so many years, but also just how much we learnt from him. We cannot promise to have absorbed or, to be frank, to have accepted, every one of his views; and we hope in particular, in this fourth edition, that we have not strayed too far from the paths through which he has guided us. But if readers recognize any attempts to stimulate and inspire rather than simply to inform, to question rather than to accept, to respect our readers rather than to patronize them, and to avoid unquestioning obedience to current reputation while acknowledging our debt to the masters of the past, then they will have identified John's intellectual legacy still firmly imprinted on the text.

In previous editions we thanked the great many friends and colleagues who helped us by commenting on various drafts of the text. The effects of their contributions are still strongly evident in the present edition. This fourth edition was also read by a series of reviewers, to whom we are deeply grateful. Several remained anonymous and so we cannot thank them by name, but we are delighted to be able to acknowledge the help of Jonathan Anderson, Mike Bonsall, Angela Douglas, Chris Elphick, Valerie Eviner, Andy Foggo, Jerry Franklin, Kevin Gaston, Charles Godfray, Sue Hartley, Marcel Holyoak, Jim Hone, Peter Hudson, Johannes Knops, Xavier Lambin, Svata Louda, Peter Morin, Steve Ormerod, Richard Sibly, Andrew Watkinson, Jacob Weiner, and David Wharton. At Blackwell, and in the production stage, we were particularly helped and encouraged by Jane Andrew, Elizabeth Frank, Rosie Hayden, Delia Sandford and Nancy Whilton.

This book is dedicated to our families – by Mike to Linda, Jessica and Robert, and by Colin to Laurel, Dominic, Jenny and Brennan, and especially to the memory of his mother, Jean Evelyn Townsend.

Mike Begon
Colin Townsend

Introduction: Ecology and its Domain

Definition and scope of ecology

The word 'ecology' was first used by Ernest Haeckel in 1869. Paraphrasing Haeckel we can describe ecology as the scientific study of the interactions between organisms and their environment. The word is derived from the Greek *oikos*, meaning 'home'. Ecology might therefore be thought of as the study of the 'home life' of living organisms. A less vague definition was suggested by Krebs (1972): 'Ecology is the scientific study of the interactions that determine the distribution and abundance of organisms'. Notice that Krebs' definition does not use the word 'environment'; to see why, it is necessary to define the word. The environment of an organism consists of all those factors and phenomena outside the organism that influence it, whether these are physical and chemical (abiotic) or other organisms (biotic). The 'interactions' in Krebs' definition are, of course, interactions with these very factors. The environment therefore retains the central position that Haeckel gave it. Krebs' definition has the merit of pinpointing the ultimate subject matter of ecology: the distribution and abundance of organisms – *where* organisms occur, *how many* occur there, and *why*. This being so, it might be better still to define ecology as:

> the scientific study of the distribution and abundance of organisms and the interactions that determine distribution and abundance.

As far as the subject matter of ecology is concerned, 'the distribution and abundance of organisms' is pleasantly succinct. But we need to expand it. The living world can be viewed as a biological hierarchy that starts with subcellular particles, and continues up through cells, tissues and organs. Ecology deals with the next three levels: the individual *organism*, the *population* (consisting of individuals of the same species) and the *community* (consisting of a greater or lesser number of species populations). At the level of the organism, ecology deals with how individuals are affected by (and how they affect) their environment. At the level of the population, ecology is concerned with the presence or absence of particular species, their abundance or rarity, and with the trends and fluctuations in their numbers. Community ecology then deals with the composition and organization of ecological communities. Ecologists also focus on the pathways followed by energy and matter as these move among living and nonliving elements of a further category of organization: the *ecosystem*, comprising the community together with its physical environment. With this in mind, Likens (1992) would extend our preferred definition of ecology to include 'the interactions between organisms and the transformation and flux of energy and matter'. However, we take energy/matter transformations as being subsumed in the 'interactions' of our definition.

There are two broad approaches that ecologists can take at each level of ecological organization. First, much can be gained by building from properties at the level below: physiology when studying organismal ecology; individual clutch size and survival probabilities when investigating the dynamics of individual species populations; food consumption rates when dealing with interactions between predator and prey populations; limits to the similarity of coexisting species when researching communities, and so on. An alternative approach deals directly with properties of the level of interest – for example, niche breadth at the organismal level; relative importance of density-dependent processes at the population level; species diversity at the level of community; rate of biomass production at the ecosystem level – and tries to relate these to abiotic or biotic aspects of the environment. Both approaches have their uses, and both will be used in each of the three parts of this book: Organisms; Species Interactions; and Communities and Ecosystems.

Explanation, description, prediction and control

At all levels of ecological organization we can try to do a number of different things. In the first place we can try to *explain* or *understand*. This is a search for knowledge in the pure scientific tradition. In order to do this, however, it is necessary first to *describe*. This, too, adds to our knowledge of the living world. Obviously, in order to understand something, we must first have a description of whatever it is that we wish to understand. Equally, but less obviously, the most valuable descriptions are those carried out with a particular problem or 'need for understanding' in mind. All descriptions are selective: but undirected description, carried out for its own sake, is often found afterwards to have selected the wrong things.

Ecologists also often try to *predict* what will happen to an organism, a population, a community or an ecosystem under a particular set of circumstances: and on the basis of these predictions we try to *control* the situation. We try to minimize the effects of locust plagues by predicting when they are likely to occur and taking appropriate action. We try to protect crops by predicting when conditions will be favorable to the crop and unfavorable to its enemies. We try to maintain endangered species by predicting the conservation policy that will enable them to persist. We try to conserve biodiversity to maintain ecosystem 'services' such as the protection of chemical quality of natural waters. Some prediction and control can be carried out without explanation or understanding. But confident predictions, precise predictions and predictions of what will happen in unusual circumstances can be made only when we can explain what is going on. Mathematical modeling has played, and will continue to play, a crucial role in the development of ecology, particularly in our ability to predict outcomes. But it is the real world we are interested in, and the worth of models must always be judged in terms of the light they shed on the working of natural systems.

It is important to realize that there are two different classes of explanation in biology: proximal and ultimate explanations. For example, the present distribution and abundance of a particular species of bird may be 'explained' in terms of the physical environment that the bird tolerates, the food that it eats and the parasites and predators that attack it. This is a *proximal* explanation. However, we may also ask how this species of bird comes to have these properties that now appear to govern its life. This question has to be answered by an explanation in evolutionary terms. The *ultimate* explanation of the present distribution and abundance of this bird lies in the ecological experiences of its ancestors. There are many problems in ecology that demand evolutionary, ultimate explanations: 'How have organisms come to possess particular combinations of size, developmental rate, reproductive output and so on?' (Chapter 4), 'What causes predators to adopt particular patterns of foraging behavior?' (Chapter 9) and 'How does it come about that coexisting species are often similar but rarely the same?' (Chapter 19). These problems are as much part of modern ecology as are the prevention of plagues, the protection of crops and the preservation of rare species. Our ability to control and exploit ecosystems cannot fail to be improved by an ability to explain and understand. And in the search for understanding, we must combine both proximal and ultimate explanations.

Pure and applied ecology

Ecologists are concerned not only with communities, populations and organisms *in nature*, but also with manmade or human-influenced environments (plantation forests, wheat fields, grain stores, nature reserves and so on), and with the consequences of human influence *on* nature (pollution, overharvesting, global climate change). In fact, our influence is so pervasive that we would be hard pressed to find an environment that was totally unaffected by human activity. Environmental problems are now high on the political agenda and ecologists clearly have a central role to play: a sustainable future depends fundamentally on ecological understanding and our ability to predict or produce outcomes under different scenarios.

When the first edition of this text was published in 1986, the majority of ecologists would have classed themselves as pure scientists, defending their right to pursue ecology for its own sake and not wishing to be deflected into narrowly applied projects. The situation has changed dramatically in 20 years, partly because governments have shifted the focus of grant-awarding bodies towards ecological applications, but also, and more fundamentally, because ecologists have themselves responded to the need to direct much of their research to the many environmental problems that have become ever more pressing. This is recognized in this new edition by a systematic treatment of ecological applications – each of the three sections of the book concludes with an applied chapter. We believe strongly that the application of ecological theory must be based on a sophisticated understanding of the pure science. Thus, our ecological application chapters are organized around the ecological understanding presented in the earlier chapters of each section.

Part 1
Organisms

Introduction

We have chosen to start this book with chapters about organisms, then to consider the ways in which they interact with each other, and lastly to consider the properties of the communities that they form. One could call this a 'constructive' approach. We could though, quite sensibly, have treated the subject the other way round – starting with a discussion of the complex communities of both natural and manmade habitats, proceeding to *de*construct them at ever finer scales, and ending with chapters on the characteristics of the individual organisms – a more analytical approach. Neither is 'correct'. Our approach avoids having to describe community patterns before discussing the populations that comprise them. But when we start with individual organisms, we have to accept that many of the environmental forces acting on them, especially the species with which they coexist, will only be dealt with fully later in the book.

This first section covers individual organisms and populations composed of just a single species. We consider initially the sorts of correspondences that we can detect between organisms and the environments in which they live. It would be facile to start with the view that every organism is in some way ideally fitted to live where it does. Rather, we emphasize in Chapter 1 that organisms frequently are as they are, and live where they do, because of the constraints imposed by their evolutionary history. All species are absent from almost everywhere, and we consider next, in Chapter 2, the ways in which environmental conditions vary from place to place and from time to time, and how these

put limits on the distribution of particular species. Then, in Chapter 3, we look at the resources that different types of organisms consume, and the nature of their interactions with these resources.

The particular species present in a community, and their abundance, give that community much of its ecological interest. Abundance and distribution (variation in abundance from place to place) are determined by the balance between birth, death, immigration and emigration. In Chapter 4 we consider some of the variety in the schedules of birth and death, how these may be quantified, and the resultant patterns in 'life histories': lifetime profiles of growth, differentiation, storage and reproduction. In Chapter 5 we examine perhaps the most pervasive interaction acting within single-species populations: intraspecific competition for shared resources in short supply. In Chapter 6 we turn to movement: immigration and emigration. Every species of plant and animal has a characteristic ability to disperse. This determines the rate at which individuals escape from environments that are or become unfavorable, and the rate at which they discover sites that are ripe for colonization and exploitation. The abundance or rarity of a species may be determined by its ability to disperse (or migrate) to unoccupied patches, islands or continents. Finally in this section, in Chapter 7, we consider the application of the principles that have been discussed in the preceding chapters, including niche theory, life history theory, patterns of movement, and the dynamics of small populations, paying particular attention to restoration after environmental damage, biosecurity (resisting the invasion of alien species) and species conservation.

Chapter 1
Organisms in their Environments: the Evolutionary Backdrop
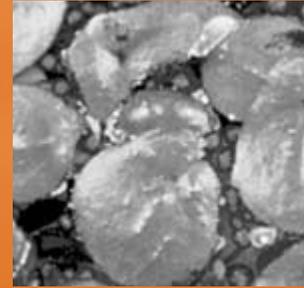
1.1 Introduction: natural selection and adaptation

From our definition of ecology in the Preface, and even from a layman's understanding of the term, it is clear that at the heart of ecology lies the relationship between organisms and their environments. In this opening chapter we explain how, fundamentally, this is an evolutionary relationship. The great Russian–American biologist Theodosius Dobzhansky famously said: 'Nothing in biology makes sense, except in the light of evolution'. This is as true of ecology as of any other aspect of biology. Thus, we try here to explain the processes by which the properties of different sorts of species make their life possible in particular environments, and also to explain their failure to live in other environments. In mapping out this evolutionary backdrop to the subject, we will also be introducing many of the questions that are taken up in detail in later chapters.

The phrase that, in everyday speech, is most commonly used to describe the match between organisms and environment is: 'organism X is adapted to' followed by a description of where the organism is found. Thus, we often hear that 'fish are adapted to live in water', or 'cacti are adapted to live in conditions of drought'. In everyday speech, this may mean very little: simply that fish have characteristics that allow them to live in water (and perhaps exclude them from other environments) or that cacti have characteristics that allow them to live where water is scarce. The word 'adapted' here says nothing about how the characteristics were acquired.

For an ecologist or evolutionary biologist, however, 'X is adapted to live in Y' means that environment Y has provided forces of natural selection that have affected the life of X's ancestors and so have molded and specialized the evolution of X. 'Adaptation' means that genetic change has occurred.

the meaning of adaptation

Regrettably, though, the word '*ad*aptation' implies that organisms are matched *to* their present environments, suggest-ing 'design' or even 'prediction'. But organisms have not been designed for, or fitted to the present: they have been molded (by *natural selection*) by past environments. Their characteristics reflect the successes and failures of ancestors. They appear to be apt for the environments that they live in at present only because present environments tend to be similar to those of the past.

The theory of evolution by natural selection is an ecological theory. It was first elaborated by Charles Darwin (1859), though its essence was also appreciated by a contemporary and corres-pondent of Darwin's, Alfred Russell Wallace (Figure 1.1). It rests on a series of propositions.

evolution by natural selection

1 The individuals that make up a population of a species are *not identical*: they vary, although sometimes only slightly, in size, rate of development, response to temperature, and so on.

2 Some, at least, of this variation is *heritable*. In other words, the characteristics of an individual are determined to some extent by its genetic make-up. Individuals receive their genes from their ancestors and therefore tend to share their characteristics.

3 All populations have the *potential* to populate the whole earth, and they would do so if each individual survived and each individual produced its maximum number of descendants. But they do not: many individuals die prior to reproduction, and most (if not all) reproduce at a less than maximal rate.

4 Different ancestors leave *different numbers of descendants*. This means much more than saying that different individuals produce different numbers of offspring. It includes also the chances of survival of offspring to reproductive age, the survival and reproduction of the progeny of these offspring, the survival and reproduction of their offspring in turn, and so on.

5 Finally, the number of descendants that an individual leaves depends, not entirely but crucially, on *the interaction between the characteristics of the individual and its environment*.

Figure 1.1   (a) Charles Darwin, 1849 (lithograph by Thomas H. Maguire; courtesy of The Royal Institution, London, UK/Bridgeman Art Library). (b) Alfred Russell Wallace, 1862 (courtesy of the Natural History Museum, London).

In any environment, some individuals will tend to survive and reproduce better, and leave more descendants, than others. If, because of this, the heritable characteristics of a population change from generation to generation, then evolution by natural selection is said to have occurred. This is the sense in which nature may loosely be thought of as *selecting*. But nature does not select in the way that plant and animal breeders select. Breeders have a defined end in view – bigger seeds or a faster racehorse. But nature does not *actively* select in this way: it simply sets the scene within which the evolutionary play of differential survival and reproduction is played out.

fitness: it's all relative

The fittest individuals in a population are those that leave the greatest number of descendants. In practice, the term is often applied not to a single individual, but to a typical individual or a type. For example, we may say that in sand dunes, yellow-shelled snails are fitter than brown-shelled snails. *Fitness*, then, is a relative not an absolute term. The fittest individuals in a population are those that leave the greatest number of descendants *relative* to the number of descendants left by other individuals in the population.

When we marvel at the diversity of complex specializations, there is a temptation to regard each case as an

evolved perfection? no

example of evolved perfection. But this would be wrong. The evolutionary process works on the genetic variation that is available. It follows that natural selection is unlikely to lead to the evolution of perfect, 'maximally fit' individuals. Rather, organisms

come to match their environments by being 'the fittest available' or 'the fittest yet': they are not 'the best imaginable'. Part of the lack of fit arises because the present properties of an organism have not all originated in an environment similar in every respect to the one in which it now lives. Over the course of its evolutionary history (its phylogeny), an organism's remote ancestors may have evolved a set of characteristics – evolutionary 'baggage' – that subsequently constrain future evolution. For many millions of years, the evolution of vertebrates has been limited to what can be achieved by organisms with a vertebral column. Moreover, much of what we now see as precise matches between an organism and its environment may equally be seen as constraints: koala bears live successfully on *Eucalyptus* foliage, but, from another perspective, koala bears cannot live without *Eucalyptus* foliage.

1.2   Specialization within species

The natural world is not composed of a continuum of types of organism each grading into the next: we recognize boundaries between one type of organism and another. Nevertheless, within what we recognize as *species* (defined below), there is often considerable variation, and some of this is heritable. It is on such intraspecific variation, after all, that plant and animal breeders (and natural selection) work.

Since the environments experienced by a species in different parts of its range are themselves different (to at least some extent), we might expect natural selection to have favored different variants of the species at different sites. The word '*ecotype*' was first coined for plant populations (Turesson, 1922a, 1922b) to describe genetically determined differences between populations within a species that reflect local matches between the organisms and their environments. But evolution forces the characteristics of populations to diverge from each other only if: (i) there is sufficient heritable variation on which selection can act; and (ii) the forces favoring divergence are strong enough to counteract the mixing and hybridization of individuals from different sites. Two populations will not diverge completely if their members (or, in the case of plants, their pollen) are continually migrating between them and mixing their genes.

Local, specialized populations become differentiated most conspicuously amongst organisms that are immobile for most of their lives. Motile organisms have a large measure of control over the environment in which they live; they can recoil or retreat from a lethal or unfavorable environment and actively seek another. Sessile, immobile organisms have no such freedom. They must live, or die, in the conditions where they settle. Populations of sessile organisms are therefore exposed to forces of natural selection in a peculiarly intense form.

This contrast is highlighted on the seashore, where the intertidal environment continually oscillates between the terrestrial and the aquatic. The fixed algae, sponges, mussels and barnacles all meet and tolerate life at the two extremes. But the mobile shrimps, crabs and fish track their aquatic habitat as it moves; whilst the shore-feeding birds track their terrestrial habitat. The mobility of such organisms enables them to match their environments to themselves. The immobile organism must match itself to its environment.

1.2.1   Geographic variation within species: ecotypes

The sapphire rockcress, *Arabis fecunda*, is a rare perennial herb restricted to calcareous soil outcrops in western Montana (USA) – so rare, in fact, that there are just 19 existing populations separated into two groups ('high elevation' and 'low elevation') by a distance of around 100 km. Whether there is local adaptation is of practical importance for conservation: four of the low elevation populations are under threat from spreading urban areas and may require reintroduction from elsewhere if they are to be sustained. Reintroduction may fail if local adaptation is too marked. Observing plants in their own habitats and checking for differences between them would not tell us if there was local adaptation in the evolutionary sense. Differences may simply be the result of immediate responses to contrasting environments made by plants that are essentially the same. Hence, high and low elevation plants were grown together in a 'common garden', eliminating any influence of contrasting immediate environments (McKay *et al.*, 2001). The low elevation sites were more prone to drought; both the air and the soil were warmer and drier. The low elevation plants in the common garden were indeed significantly more drought tolerant (Figure 1.2).

On the other hand, local selection by no means always overrides hybridization. For example, in a study of *Chamaecrista fasciculata*, an annual legume from disturbed habitats in eastern North America, plants were grown in a common garden that were derived from the 'home' site or were transplanted from distances of 0.1, 1, 10, 100, 1000 and 2000 km (Galloway & Fenster, 2000). The study was replicated three times: in Kansas, Maryland and northern Illinois. Five characteristics were measured: germination, survival, vegetative biomass, fruit production and the number of fruit produced per seed planted. But for all characters in all replicates there was little or no evidence for local adaptation except at the very furthest spatial scales (e.g. Figure 1.3). There is 'local adaptation' – but it's clearly not *that* local.

the balance between local adaptation and hybridization

We can also test whether organisms have evolved to become specialized to life in their local environment in *reciprocal transplant* experiments: comparing their performance when they are grown 'at home' (i.e. in their original habitat) with their performance 'away' (i.e. in the habitat of others). One such experiment (concerning white clover) is described in the next section.
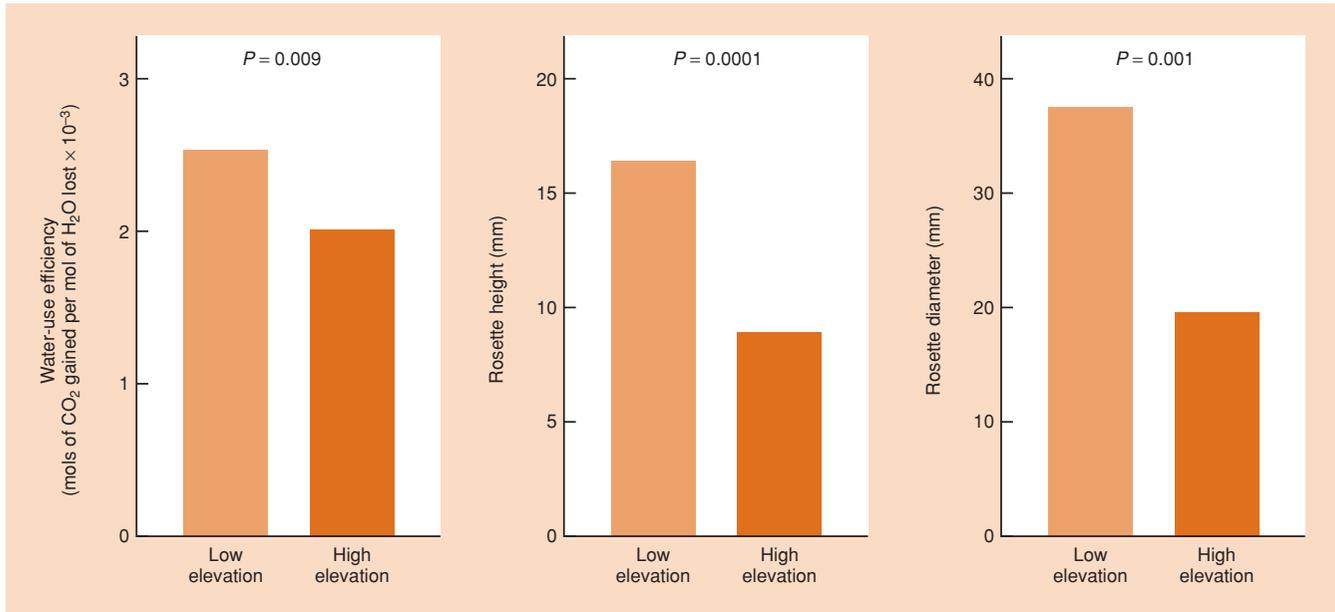
Figure 1.2 When plants of the rare sapphire rockcress from low elevation (drought-prone) and high elevation sites were grown together in a common garden, there was local adaptation: those from the low elevation site had significantly better water-use efficiency as well as having both taller and broader rosettes. (From McKay *et al.*, 2001.)
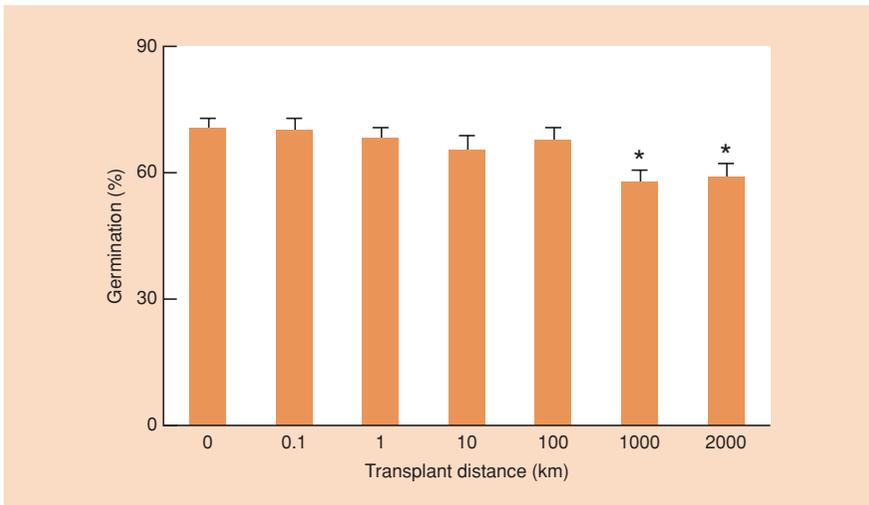
Figure 1.3 Percentage germination of local and transplanted *Chamaecrista fasciculata* populations to test for local adaptation along a transect in Kansas. Data for 1995 and 1996 have been combined because they do not differ significantly. Populations that differ from the home population at $P < 0.05$ are indicated by an asterisk. Local adaptation occurs at only the largest spatial scales. (From Galloway & Fenster, 2000.)

1.2.2 Genetic polymorphism

transient polymorphisms

On a finer scale than ecotypes, it may also be possible to detect levels of variation *within* populations. Such variation is known as polymorphism. Specifically, genetic polymorphism is 'the occurrence together in the same habitat of two or more discontinuous forms of a species

in such proportions that the rarest of them cannot merely be maintained by recurrent mutation or immigration' (Ford, 1940). Not all such variation represents a match between organism and environment. Indeed, some of it may represent a mismatch, if, for example, conditions in a habitat change so that one form is being replaced by another. Such polymorphisms are called transient. As all communities are always changing, much polymorphism that we observe in nature may be transient, representing

the extent to which the genetic response of populations to environmental change will always be out of step with the environment and unable to anticipate changing circumstances – this is illustrated in the peppered moth example below.

the maintenance of polymorphisms

Many polymorphisms, however, are actively maintained in a population by natural selection, and there are a number of ways in which this may occur.

1 *Heterozygotes may be of superior fitness*, but because of the mechanics of Mendelian genetics they continually generate less fit homozygotes within the population. Such 'heterosis' is seen in human sickle-cell anaemia where malaria is prevalent. The malaria parasite attacks red blood cells. The sickle-cell mutation gives rise to red cells that are physiologically imperfect and misshapen. However, sickle-cell heterozygotes are fittest because they suffer only slightly from anemia and are little affected by malaria; but they continually generate homozygotes that are either dangerously anemic (two sickle-cell genes) or susceptible to malaria (no sickle-cell genes). None the less, the superior fitness of the heterozygote maintains both types of gene in the population (that is, a polymorphism).

2 *There may be gradients of selective forces* favoring one form (morph) at one end of the gradient, and another form at the other. This can produce polymorphic populations at intermediate positions in the gradient – this, too, is illustrated below in the peppered moth study.

3 *There may be frequency-dependent selection* in which each of the morphs of a species is fittest when it is rarest (Clarke & Partridge, 1988). This is believed to be the case when rare color forms of prey are fit because they go unrecognized and are therefore ignored by their predators.

4 *Selective forces may operate in different directions within different patches* in the population. A striking example of this is provided by a reciprocal transplant study of white clover (*Trifolium repens*) in a field in North Wales (UK). To determine whether the characteristics of individuals matched local features of their environment, Turkington and Harper (1979) removed plants from marked positions in the field and multiplied them into clones in the common environment of a greenhouse. They then transplanted samples from each clone into the place in the sward of vegetation from which it had originally been taken (as a control), and also to the places from where all the others had been taken (a transplant). The plants were allowed to grow for a year before they were removed, dried and weighed. The mean weight of clover plants transplanted back into their home sites was 0.89 g but at away sites it was only 0.52 g, a statistically highly significant difference. This provides strong, direct evidence that clover clones in the pasture had evolved to become specialized such that they performed best in their local environment. But all this was going on within a single population, which was therefore polymorphic.

In fact, the distinction between local ecotypes and polymorphic populations is not always a clear one. This is illustrated by another study in North Wales, where there was a gradation in

no clear distinction between local ecotypes and a polymorphism

habitats at the margin between maritime cliffs and grazed pasture, and a common species, creeping bent grass (*Agrostis stolonifera*), was present in many of the habitats. Figure 1.4 shows a map of the site and one of the transects from which plants were sampled. It also shows the results when plants from the sampling points along this transect were grown in a common garden. The

Figure 1.4 (a) Map of Abraham's Bosom, the site chosen for a study of evolution over very short distances. The darker colored area is grazed pasture; the lighter areas are the cliffs falling to the sea. The numbers indicate the sites from which the grass *Agrostis stolonifera* was sampled. Note that the whole area is only 200 m long. (b) A vertical transect across the study area showing the gradual change from pasture to cliff conditions. (c) The mean length of stolons produced in the experimental garden from samples taken from the transect. (From Aston & Bradshaw, 1966.)
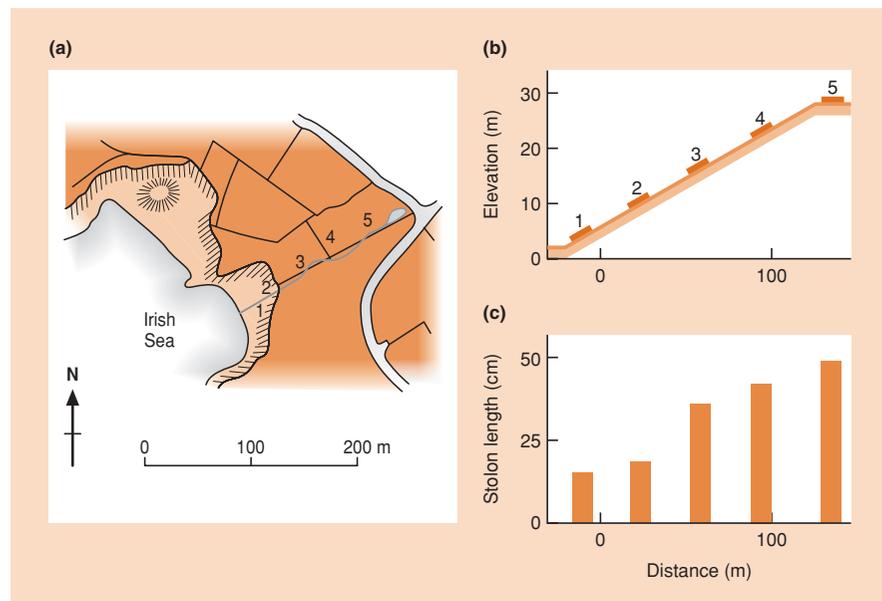
plants spread by sending out shoots along the ground surface (stolons), and the growth of plants was compared by measuring the lengths of these. In the field, cliff plants formed only short stolons, whereas those of the pasture plants were long. In the experimental garden, these differences were maintained, even though the sampling points were typically only around 30 m apart – certainly within the range of pollen dispersal between plants. Indeed, the gradually changing environment along the transect was matched by a gradually changing stolon length, presumably with a genetic basis, since it was apparent in the common garden. Thus, even though the spatial scale was so small, the forces of selection seem to outweigh the mixing forces of hybridization – but it is a moot point whether we should describe this as a small-scale series of local ecotypes or a polymorphic population maintained by a gradient of selection.

1.2.3   Variation within a species with manmade selection pressures

It is, perhaps, not surprising that some of the most dramatic examples of local specialization within species (indeed of natural selection in action) have been driven by manmade ecological forces, especially those of environmental pollution. These can provide rapid change under the influence of powerful selection pressures. *Industrial melanism*, for example, is the phenomenon in which black or blackish forms of species have come to dominate populations in industrial areas. In the dark individuals, a dominant gene is typically responsible for producing an excess of the black pigment melanin. Industrial melanism is known in most industrialized countries and more than 100 species of moth have evolved forms of industrial melanism.
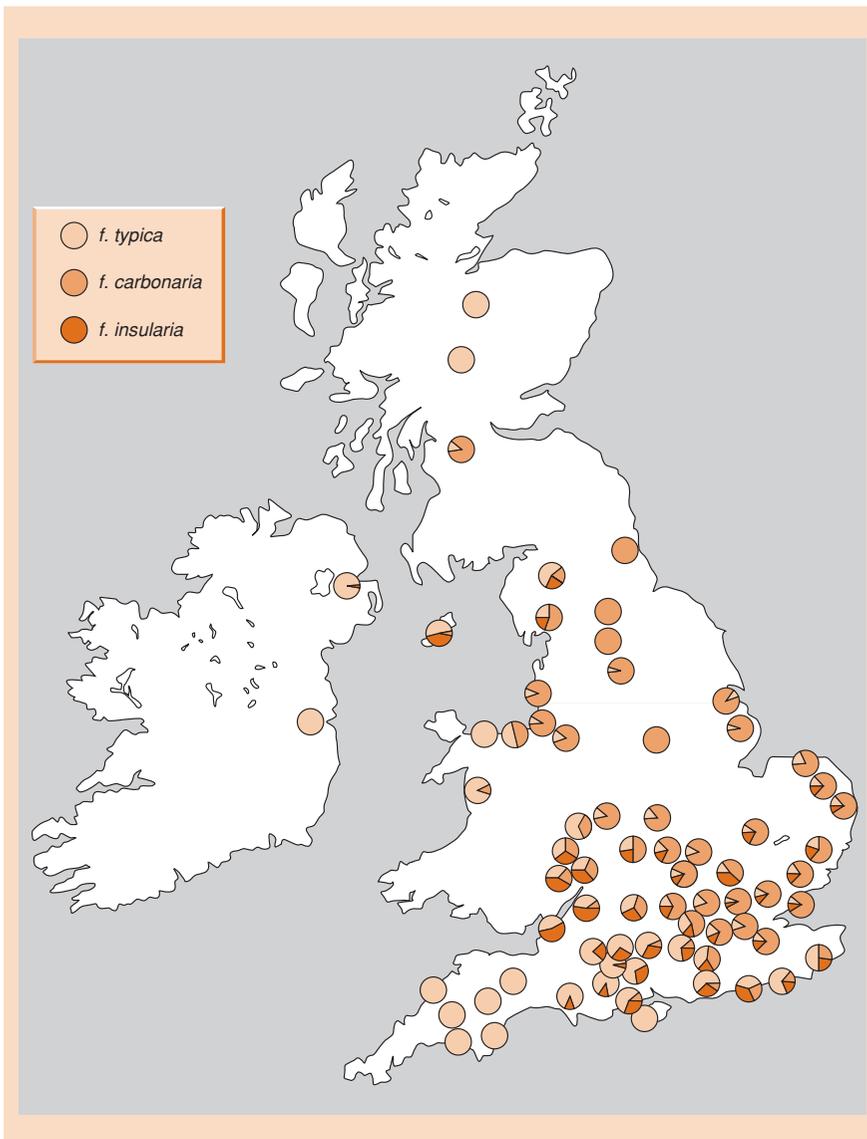
Figure 1.5   Sites in Britain where the frequencies of the pale (*forma typica*) and melanic forms of *Biston betularia* were recorded by Kettlewell and his colleagues. In all more than 20,000 specimens were examined. The principal melanic form (*forma carbonaria*) was abundant near industrial areas and where the prevailing westerly winds carry atmospheric pollution to the east. A further melanic form (*forma insularia*, which looks like an intermediate form but is due to several different genes controlling darkening) was also present but was hidden where the genes for *forma carbonaria* were present. (From Ford, 1975.)

The earliest recorded species to evolve in this way was the peppered moth (*Biston betularia*); the first black specimen in an otherwise pale population was caught in Manchester (UK) in 1848. By 1895, about 98% of the Manchester peppered moth population was melanic. Following many more years of pollution, a large-scale survey of pale and melanic forms of the peppered moth in Britain recorded more than 20,000 specimens between 1952 and 1970 (Figure 1.5). The winds in Britain are predominantly westerlies, spreading industrial pollutants (especially smoke and sulfur dioxide) toward the east. Melanic forms were concentrated toward the east and were completely absent from the unpolluted western parts of England and Wales, northern Scotland and Ireland. Notice from the figure, though, that many populations were polymorphic: melanic and nonmelanic forms coexisted. Thus, the polymorphism seems to be a result both of environments changing (becoming more polluted) – to this extent the polymorphism is transient – and of there being a gradient of selective pressures from the less polluted west to the more polluted east.

The main selective pressure appears to be applied by birds that prey on the moths. In field experiments, large numbers of melanic and pale ('typical') moths were reared and released in equal numbers. In a rural and largely unpolluted area of southern England, most of those captured by birds were melanic. In an industrial area near the city of Birmingham, most were typicals (Kettlewell, 1955). Any idea, however, that melanic forms were favored simply because they were camouflaged against smoke-stained backgrounds in the polluted areas (and typicals were favored in unpolluted areas because they were camouflaged against pale backgrounds) may be only part of the story. The moths rest on tree trunks during the day, and nonmelanic moths are well hidden against a background of mosses and lichens. Industrial pollution has not just blackened the moths' background; sulfur dioxide, especially, has also destroyed most of the moss and lichen on the tree trunks. Thus, sulfur dioxide pollution may have been as important as smoke in selecting melanic moths.

In the 1960s, industrialized environments in Western Europe and the United States started to change again, as oil and electricity began to replace coal, and legislation was passed to impose smoke-free zones and to reduce industrial emissions of sulfur dioxide. The frequency of melanic forms then fell back to near pre-Industrial levels with remarkable speed (Figure 1.6). Again, there was transient polymorphism – but this time while populations were *en route* in the other direction.

1.3 Speciation

It is clear, then, that natural selection can force populations of plants and animals to change their character – to evolve. But none of the examples we have considered has involved the evolution of
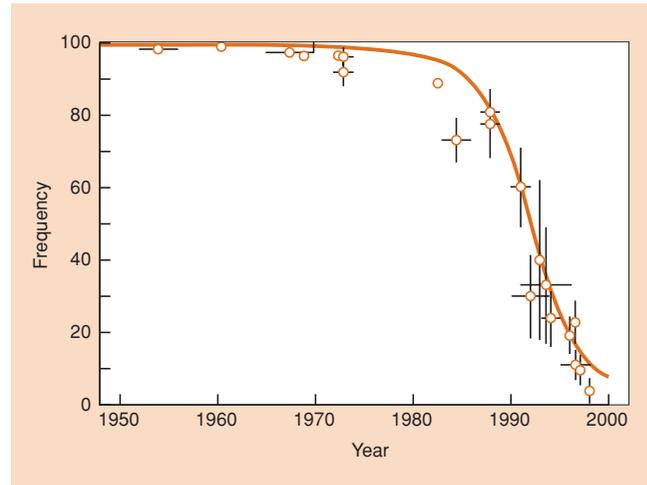

Figure 1.6   Change in the frequency of the *carbonaria* form of the peppered moth *Biston betularia* in the Manchester area since 1950. Vertical lines show the standard error and the horizontal lines show the range of years included. (After Cook *et al.*, 1999.)

a new species. What, then, justifies naming two populations as different species? And what is the process – 'speciation' – by which two or more new species are formed from one original species?

1.3.1   What do we mean by a 'species'?

Cynics have said, with some truth, that a species is what a competent taxonomist regards as a species. On the other hand, back in the 1930s two American biologists, Mayr and Dobzhansky, proposed an empirical test that could be used to decide whether two populations were part of the same species or of two different species. They recognized organisms as being members of a single species if they could, at least potentially, breed together in nature to produce fertile offspring. They called a species tested and defined in this way a *biological species* or *biospecies*. In the examples that we have used earlier in this chapter we know that melanic and normal peppered moths can mate and that the offspring are fully fertile; this is also true of plants from the different types of *Agrostis*. They are all variations within species – not separate species.

In practice, however, biologists do not apply the Mayr–Dobzhansky test before they recognize every species: there is simply not enough time or resources, and in any case, there are vast portions of the living world – most microorganisms, for example – where an absence of sexual reproduction makes a strict interbreeding criterion inappropriate. What is more important is that the test recognizes a crucial element in the evolutionary process that we have met already in considering specialization
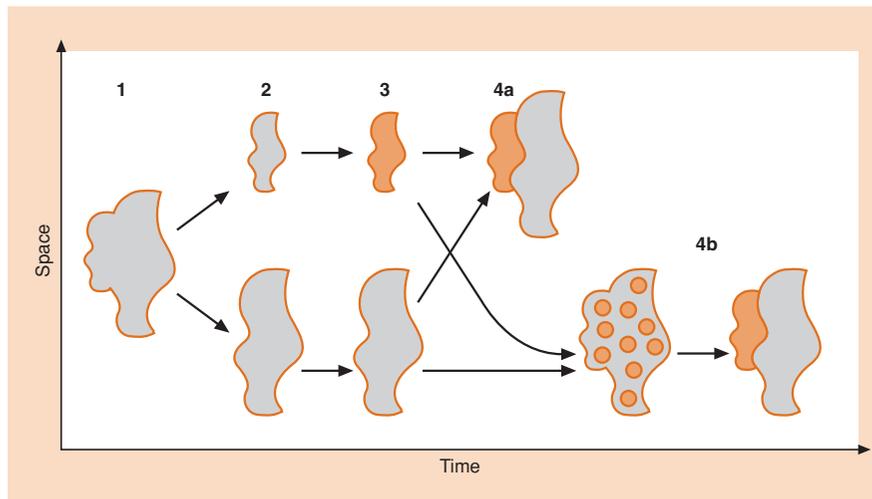
Figure 1.7 The orthodox picture of ecological speciation. A uniform species with a large range (1) differentiates (2) into subpopulations (for example, separated by geographic barriers or dispersed onto different islands), which become genetically isolated from each other (3). After evolution in isolation they may meet again, when they are either already unable to hybridize (4a) and have become true biospecies, or they produce hybrids of lower fitness (4b), in which case evolution may favor features that prevent interbreeding between the 'emerging species' until they are true biospecies.

within species. If the members of two populations are able to hybridize, and their genes are combined and reassorted in their progeny, then natural selection can never make them truly distinct. Although natural selection may tend to force a population to evolve into two or more distinct forms, sexual reproduction and hybridization mix them up again.

orthodox ecological speciation

'Ecological' speciation is speciation driven by divergent natural selection in distinct subpopulations (Schluter, 2001). The most orthodox scenario for this comprises a number of stages (Figure 1.7). First, two subpopulations become geographically isolated and natural selection drives genetic adaptation to their local environments. Next, as a *by-product* of this genetic differentiation, a degree of reproductive isolation builds up between the two. This may be 'pre-zygotic', tending to prevent mating in the first place (e.g. differences in courtship ritual), or 'post-zygotic': reduced viability, perhaps inviability, of the offspring themselves. Then, in a phase of '*secondary contact*', the two subpopulations re-meet. The hybrids between individuals from the different subpopulations are now of low fitness, because they are literally neither one thing nor the other. Natural selection will then favor any feature in either subpopulation that *reinforces* reproductive isolation, especially pre-zygotic characteristics, preventing the production of low-fitness hybrid offspring. These breeding barriers then cement the distinction between what have now become separate species.

allopatric and sympatric speciation

It would be wrong, however, to imagine that all examples of speciation conform fully to this orthodox picture (Schluter, 2001). First, there may never be secondary contact. This would be pure 'allopatric' speciation (that is, with all divergence occurring in subpopulations in *different* places). Second, there is clearly room for considerable variation in the relative importances of pre-zygotic and post-zygotic

mechanisms in both the allopatric and the secondary-contact phases.

Most fundamentally, perhaps, there has been increasing support for the view that an allopatric phase is not necessary: that is, 'sympatric' speciation is possible, with subpopulations diverging despite not being geographically separated from one another. Probably the most studied circumstance in which this seems likely to occur (see Drès & Mallet, 2002) is where insects feed on more than one species of host plant, and where each requires specialization by the insects to overcome the plant's defenses. (Consumer resource defense and specialization are examined more fully in Chapters 3 and 9.) Particularly persuasive in this is the existence of a continuum identified by Drès and Mallet: from populations of insects feeding on more than one host plant, through populations differentiated into 'host races' (defined by Drès and Mallet as sympatric subpopulations exchanging genes at a rate of more than around 1% per generation), to coexisting, closely related species. This reminds us, too, that the origin of a species, whether allopatric or sympatric, is a process, not an event. For the formation of a new species, like the boiling of an egg, there is some freedom to argue about when it is completed.

The evolution of species and the balance between natural selection and hybridization are illustrated by the extraordinary case of two species of sea gull. The lesser black-backed gull (*Larus fuscus*) originated in Siberia and colonized progressively to the west, forming a chain or *cline* of different forms, spreading from Siberia to Britain and Iceland (Figure 1.8). The neighboring forms along the cline are distinctive, but they hybridize readily in nature. Neighboring populations are therefore regarded as part of the same species and taxonomists give them only 'subspecific' status (e.g. *L. fuscus graellsii, L. fuscus fuscus*). Populations of the gull have, however, also spread east from Siberia, again forming a cline of freely hybridizing forms. Together, the populations spreading east and west encircle the northern hemisphere. They meet and overlap
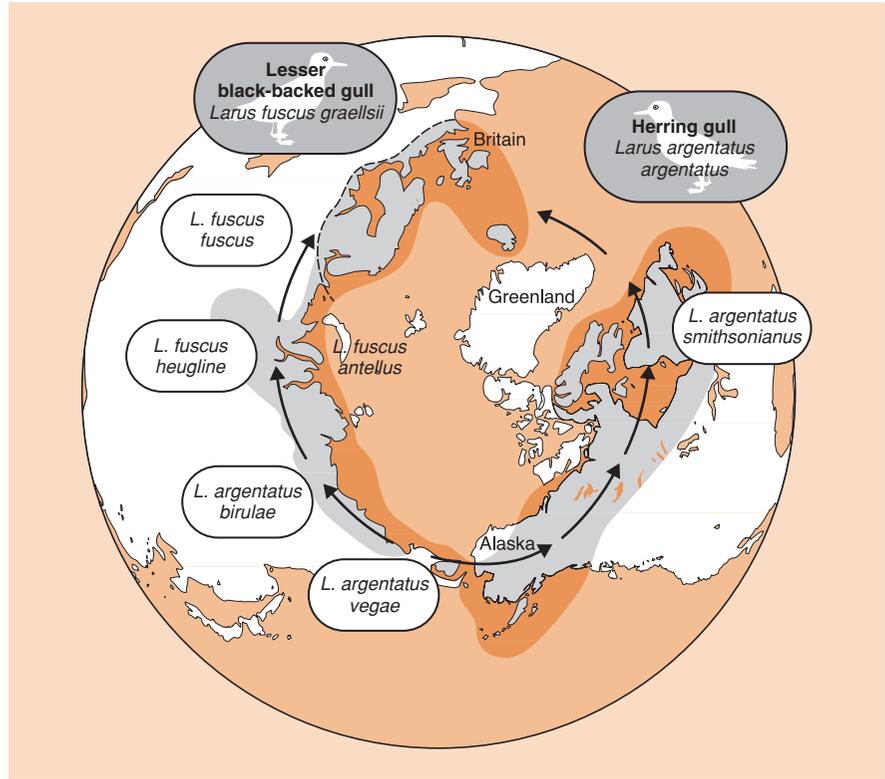
Figure 1.8 Two species of gull, the herring gull and the lesser black-backed gull, have diverged from a common ancestry as they have colonized and encircled the northern hemisphere. Where they occur together in northern Europe they fail to interbreed and are clearly recognized as two distinct species. However, they are linked along their ranges by a series of freely interbreeding races or subspecies. (After Brookes, 1998.)

in northern Europe. There, the eastward and westward clines have diverged so far that it is easy to tell them apart, and they are recognized as two different *species*, the lesser black-backed gull (*L. fuscus*) and the herring gull (*L. argentatus*). Moreover, the two species do not hybridize: they have become true biospecies. In this remarkable example, then, we can see how two distinct species have evolved from one primal stock, and that the stages of their divergence remain frozen in the cline that connects them.

1.3.2 Islands and speciation

Darwin's finches We will see repeatedly later in the book (and especially in Chapter 21) that the isolation of islands – and not just land islands in a sea of water – can have a profound effect on the ecology of the populations and communities living there. Such isolation also provides arguably the most favorable environment for populations to diverge into distinct species. The most celebrated example of evolution and speciation on islands is the case of Darwin's finches in the Galápagos archipelago. The Galápagos are volcanic islands isolated in the Pacific Ocean about 1000 km west of Ecuador and 750 km from the island of Cocos, which is itself 500 km from Central America. At more than 500 m above sea level the vegetation is open grassland. Below this

is a humid zone of forest that grades into a coastal strip of desert vegetation with some endemic species of prickly pear cactus (*Opuntia*). Fourteen species of finch are found on the islands. The evolutionary relationships amongst them have been traced by molecular techniques (analyzing variation in 'microsatellite' DNA) (Figure 1.9) (Petren *et al.*, 1999). These accurate modern tests confirm the long-held view that the family tree of the Galápagos finches radiated from a single trunk: a single ancestral species that invaded the islands from the mainland of Central America. The molecular data also provide strong evidence that the warbler finch (*Certhidea olivacea*) was the first to split off from the founding group and is likely to be the most similar to the original colonist ancestors. The entire process of evolutionary divergence of these species appears to have happened in less than 3 million years.

Now, in their remote island isolation, the Galápagos finches, despite being closely related, have radiated into a variety of species with contrasting ecologies (Figure 1.9), occupying ecological niches that elsewhere are filled by quite unrelated species. Members of one group, including *Geospiza fuliginosa* and *G. fortis*, have strong bills and hop and scratch for seeds on the ground. *G. scandens* has a narrower and slightly longer bill and feeds on the flowers and pulp of the prickly pears as well as on seeds. Finches of a third group have parrot-like bills and feed on leaves, buds, flowers and fruits, and a fourth group with a parrot-like bill (*Camarhynchus*
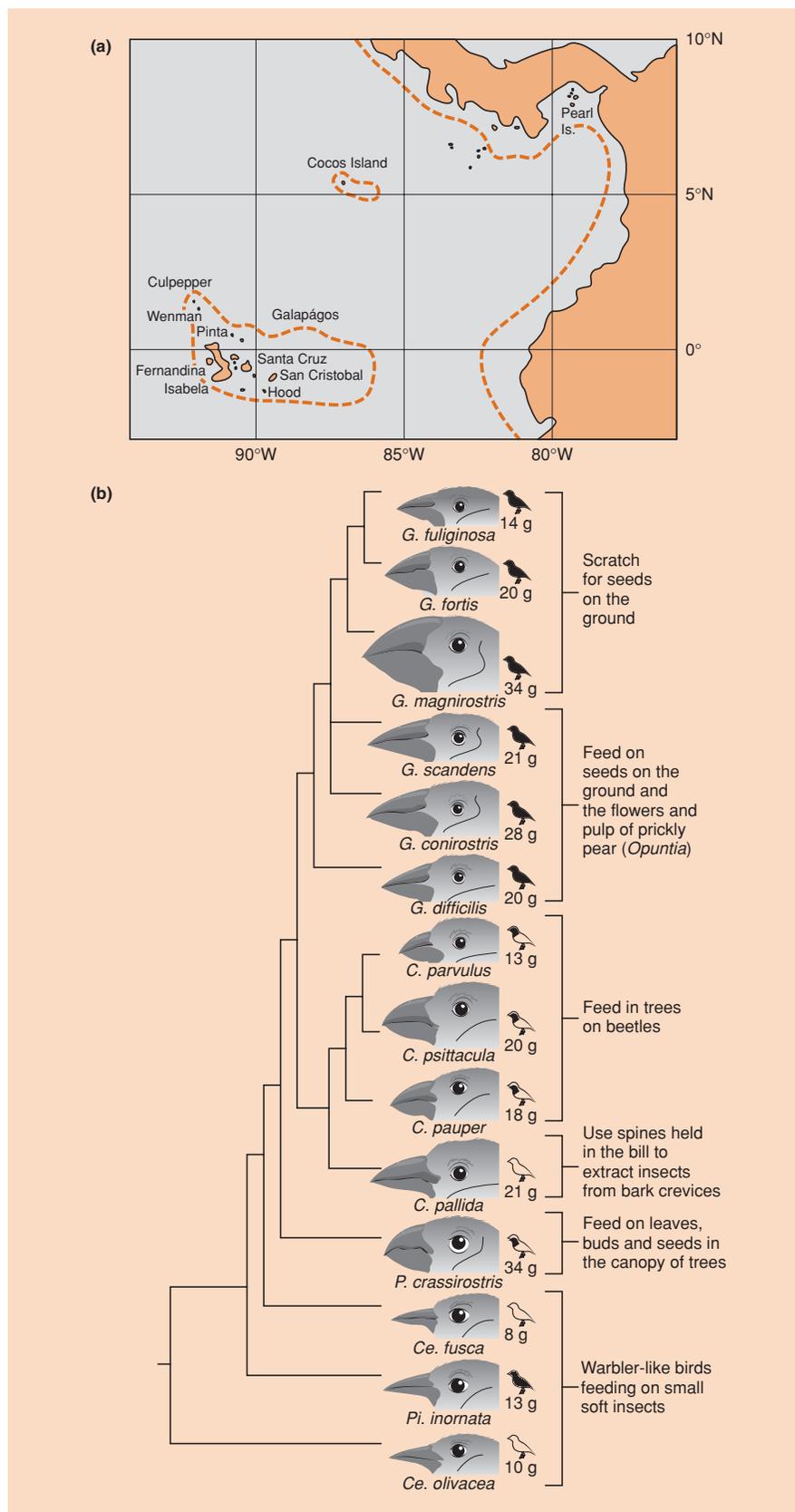
Figure 1.9 (a) Map of the Galápagos Islands showing their position relative to Central America; on the equator 5° equals approximately 560 km. (b) A reconstruction of the evolutionary history of the Galápagos finches based on variation in the length of microsatellite deoxyribonucleic acid (DNA). The feeding habits of the various species are also shown. Drawings of the birds are proportional to actual body size. The maximum amount of black coloring in male plumage and the average body mass are shown for each species. The *genetic distance* (a measure of the genetic difference) between species is shown by the length of the horizontal lines. Notice the great and early separation of the warbler finch (*Certhidea olivacea*) from the others, suggesting that it may closely resemble the founders that colonized the islands. C, *Camarhynchus*; Ce, *Certhidea*; G, *Geospiza*; P, *Platyspiza*; Pi, *Pinaroloxias*. (After Petren et al., 1999.)

psittacula) has become insectivorous, feeding on beetles and other insects in the canopy of trees. A so-called woodpecker finch, *Camarhynchus* (*Cactospiza*) *pallida*, extracts insects from crevices by holding a spine or a twig in its bill, while yet a further group includes the warbler finch, which flits around actively and collects small insects in the forest canopy and in the air. Isolation – both of the archipelago itself and of individual islands within it – has led to an original evolutionary line radiating into a series of species, each matching its own environment.

1.4 Historical factors

Our world has not been constructed by someone taking each species in turn, testing it against each environment, and molding it so that every species finds its perfect place. It is a world in which species live where they do for reasons that are often, at least in part, accidents of history. We illustrate this first by continuing our examination of islands.

1.4.1 Island patterns

Many of the species on islands are either subtly or profoundly different from those on the nearest comparable area of mainland. Put simply, there are two main reasons for this.

1 The animals and plants on an island are limited to those types having ancestors that managed to disperse there, although the extent of this limitation depends on the isolation of the island and the intrinsic dispersal ability of the animal or plant in question.
2 Because of this isolation, as we saw in the previous section, the rate of evolutionary change on an island may often be fast enough to outweigh the effects of the exchange of genetic material between the island population and related populations elsewhere.

Thus, islands contain many species unique to themselves ('*endemics*' – species found in only one area), as well as many differentiated 'races' or 'subspecies' that are distinguishable from mainland forms. A few individuals that disperse by chance to a habitable island can form the nucleus of an expanding new species. Its character will have been colored by the particular genes that were represented among the colonists – which are unlikely to be a perfect sample of the parent population. What natural selection can do with this *founder population* is limited by what is in its limited sample of genes (plus occasional rare mutations). Indeed much of the deviation among populations isolated on islands appears to be due to a *founder effect* – the chance composition of the pool of founder genes puts limits and constraints on what variation there is for natural selection to act upon.

The *Drosophila* fruit-flies of Hawaii provide a further spectacular example of species formation on islands. The Hawaiian chain of islands (Figure 1.10) is volcanic in origin, having been formed gradually over the last 40 million years, as the center of the Pacific tectonic plate moved steadily over a 'hot spot' in a southeasterly direction (Niihau is the most ancient of the islands, Hawaii itself the most recent). The richness of the Hawaiian *Drosophila* is spectacular: there are probably about 1500 *Drosophila* spp. worldwide, but at least 500 of these are found only in the Hawaiian islands.

Of particular interest are the 100 or so species of 'picture-winged' *Drosophila*. The lineages through which these species have evolved can be traced by analyzing the banding patterns on the giant chromosomes in the salivary glands of their larvae. The evolutionary tree that emerges is shown in Figure 1.10, with each species lined up above the island on which it is found (there are only two species found on more than one island). The historical element in 'what lives where' is plainly apparent: the more ancient species live on the more ancient islands, and, as new islands have been formed, rare dispersers have reached them and eventually evolved in to new species. At least some of these species appear to match the same environment as others on different islands. Of the closely related species, for example, *D. adiastola* (species 8) is only found on Maui and *D. setosimentum* (species 11) only on Hawaii, but the environments that they live in are apparently indistinguishable (Heed, 1968). What is most noteworthy, of course, is the power and importance of isolation (coupled with natural selection) in generating new species. Thus, island biotas illustrate two important, related points: (i) that there is a historical element in the match between organisms and environments; and (ii) that there is not just one perfect organism for each type of environment.

1.4.2 Movements of land masses

Long ago, the curious distributions of species between continents, seemingly inexplicable in terms of dispersal over vast distances, led biologists, especially Wegener (1915), to suggest that the continents themselves must have moved. This was vigorously denied by geologists, until geomagnetic measurements required the same, apparently wildly improbable explanation. The discovery that the tectonic plates of the earth's crust move and carry with them the migrating continents, reconciles geologist and biologist (Figure 1.11b–e). Thus, whilst major evolutionary developments were occurring in the plant and animal kingdoms, populations were being split and separated, and land areas were moving across climatic zones.

Figure 1.12 shows just one example of a major group of organisms (the large flightless birds), whose distributions begin to make sense only in the light of the movement of land masses. It would be

Hawaiian Drosophila

large flightless birds
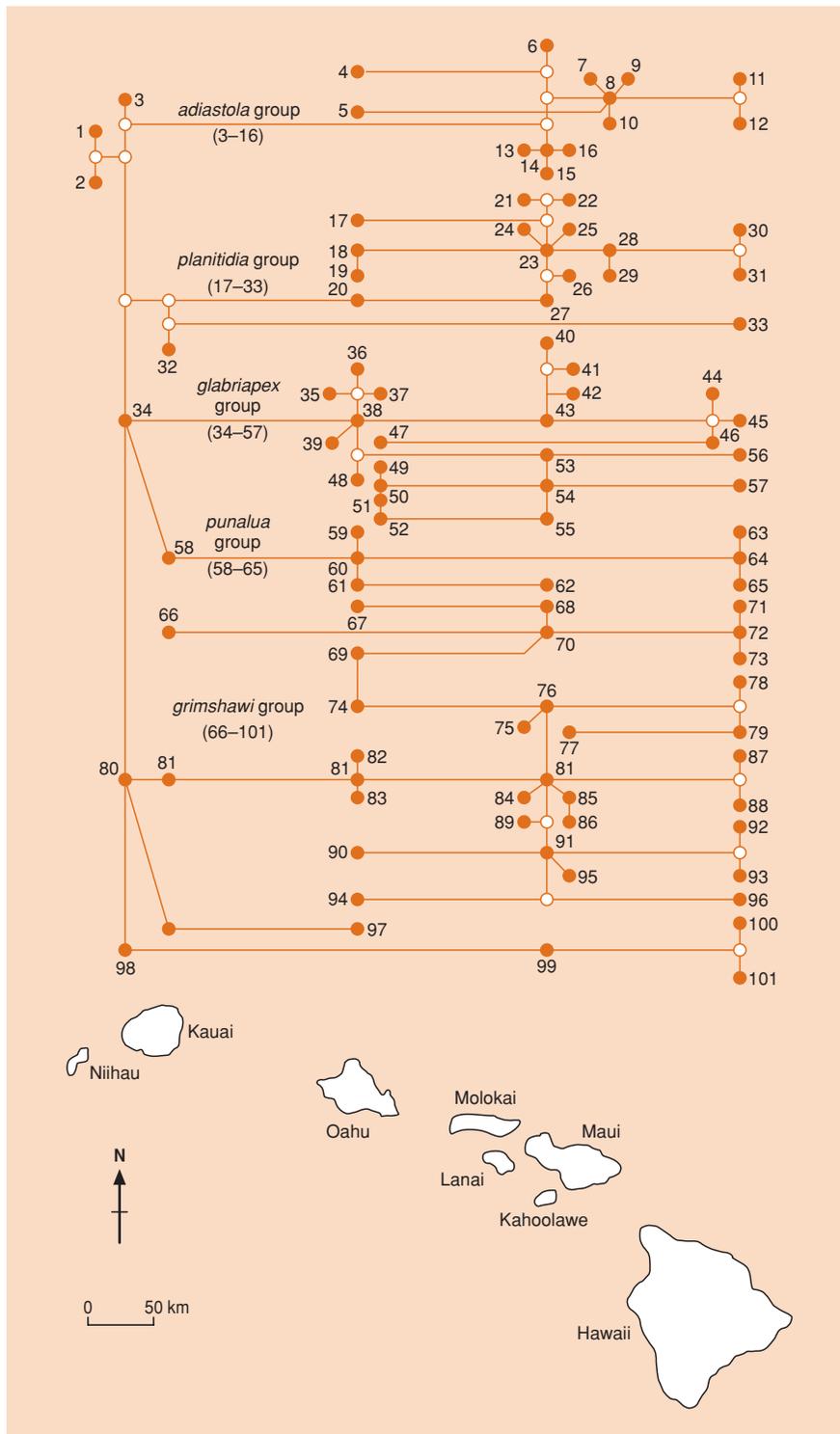
Figure 1.10 An evolutionary tree linking the picture-winged *Drosophila* of Hawaii, traced by the analysis of chromosomal banding patterns. The most ancient species are *D. primaeva* (species 1) and *D. attigua* (species 2), found only on the island of Kauai. Other species are represented by solid circles; hypothetical species, needed to link the present day ones, are represented by open circles. Each species has been placed above the island or islands on which it is found (although Molokai, Lanai and Maui are grouped together). Niihau and Kahoolawe support no *Drosophila*. (After Carson & Kaneshiro, 1976; Williamson, 1981.)
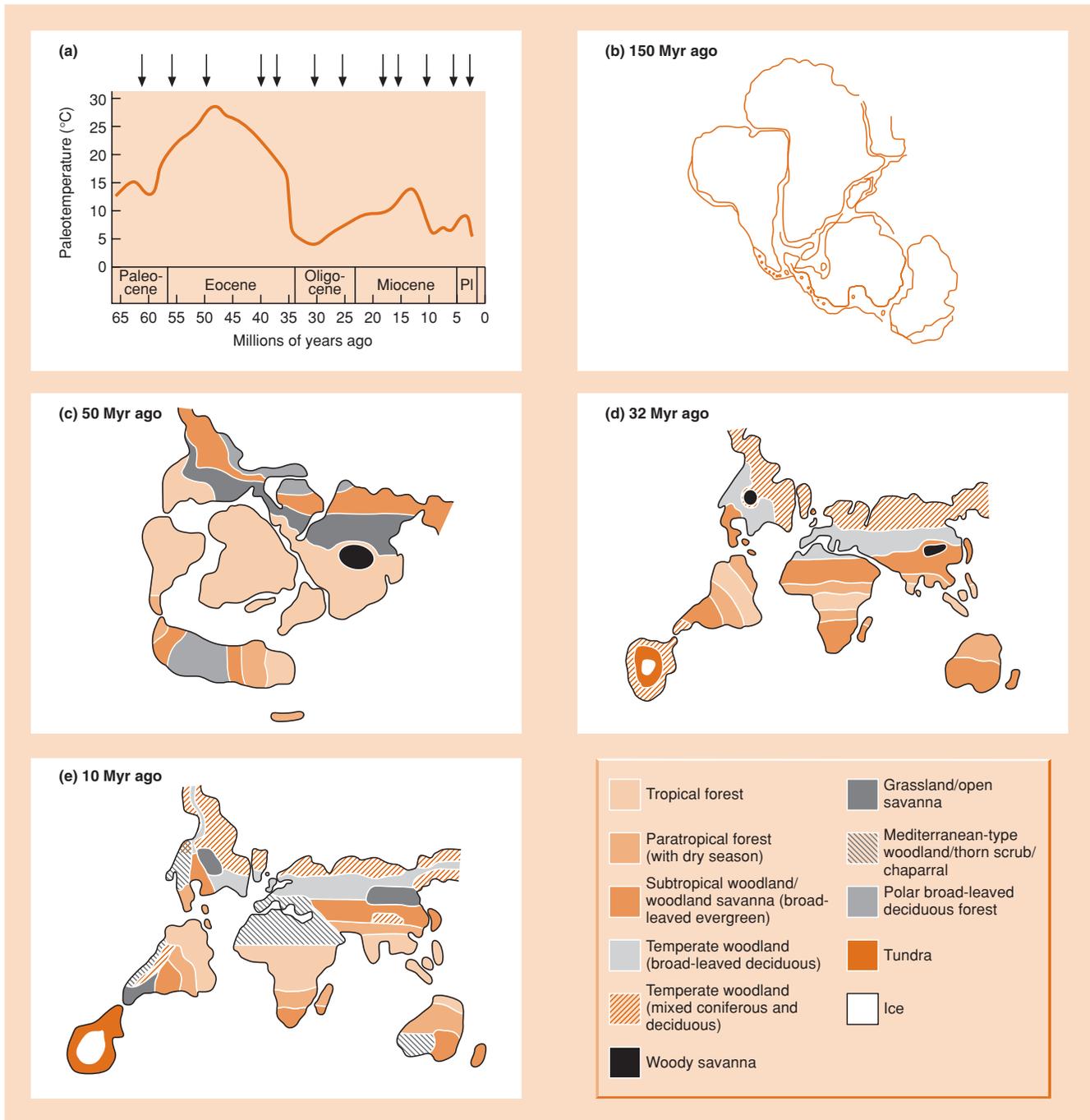
Figure 1.11 (a) Changes in temperature in the North Sea over the past 60 million years. During this period there were large changes in sea level (arrows) that allowed dispersal of both plants and animals between land masses. (b–e) Continental drift. (b) The ancient supercontinent of Gondwanaland began to break up about 150 million years ago. (c) About 50 million years ago (early Middle Eocene) recognizable bands of distinctive vegetation had developed, and (d) by 32 million years ago (early Oligocene) these had become more sharply defined. (e) By 10 million years ago (early Miocene) much of the present geography of the continents had become established but with dramatically different climates and vegetation from today; the position of the Antarctic ice cap is highly schematic. (Adapted from Norton & Sclater, 1979; Janis, 1993; and other sources).
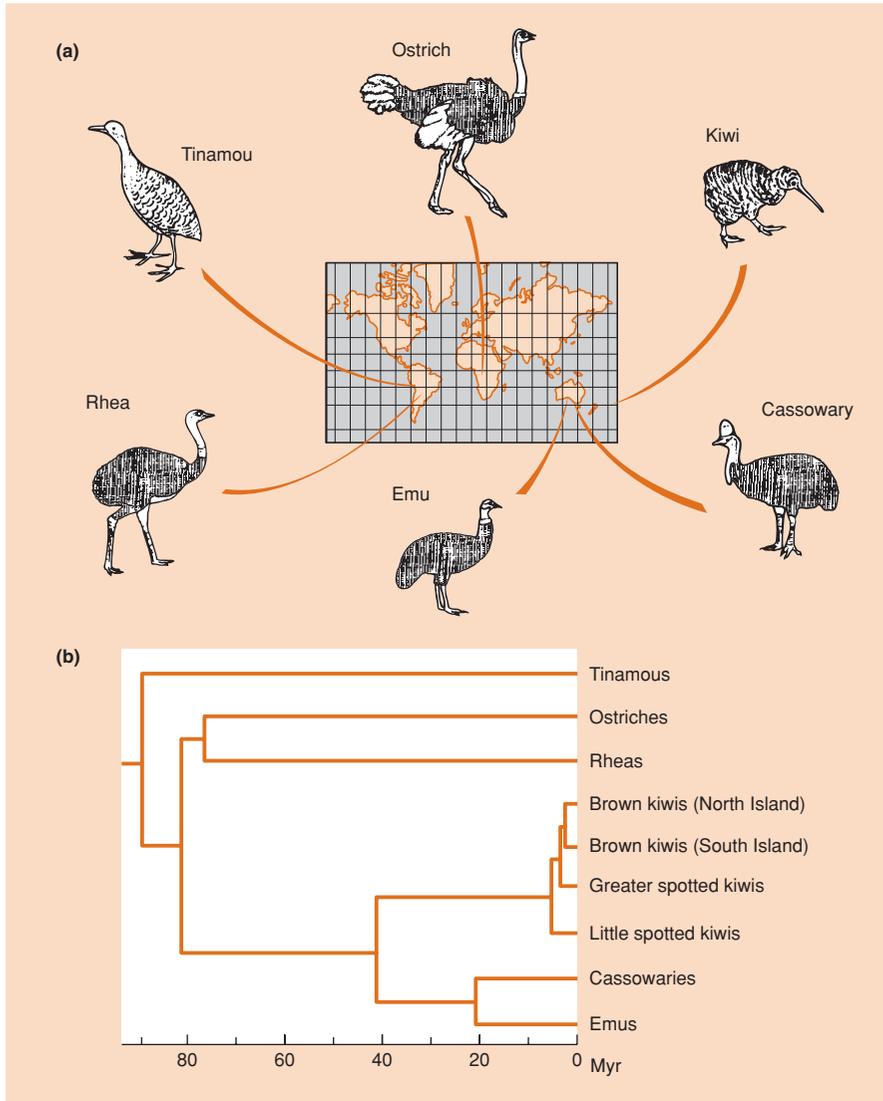
Figure 1.12 (a) The distribution of terrestrial flightless birds. (b) The phylogenetic tree of the flightless birds and the estimated times (million years, Myr) of their divergence. (After Diamond, 1983; from data of Sibley & Ahlquist.)

unwarranted to say that the emus and cassowaries are where they are because they represent the best match to Australian environments, whereas the rheas and tinamous are where *they* are because they represent the best match to South American environments. Rather, their disparate distributions are essentially determined by the prehistoric movements of the continents, and the subsequent impossibility of geographically isolated evolutionary lines reaching into each others' environment. Indeed, molecular techniques make it possible to analyze the time at which the various flightless birds started their evolutionary divergence (Figure 1.12). The tinamous seem to have been the first to diverge and became evolutionarily separate from the rest, the *ratites*. Australasia next split away from the other southern continents, and from the latter, the ancestral stocks of ostriches and rheas were subsequently separated when the Atlantic opened up between Africa

and South America. Back in Australasia, the Tasman Sea opened up about 80 million years ago and ancestors of the kiwi are thought to have made their way, by island hopping, about 40 million years ago across to New Zealand, where divergence into the present species happened relatively recently. An account of the evolutionary trends amongst mammals over much the same period is given by Janis (1993).

1.4.3 Climatic changes

Changes in climate have occurred on shorter timescales than the movements of land masses (Boden *et al.*, 1990; IGBP, 1990). Much of what we see in the present distribution of species represents phases in a recovery from past climatic shifts. Changes in

climate during the Pleistocene ice ages, in particular, bear a lot of the responsibility for the present patterns of distribution of plants and animals. The extent of these climatic and biotic changes is only beginning to be unraveled as the technology for discovering, analyzing and dating biological remains becomes more sophisticated (particularly by the analysis of buried pollen samples). These methods increasingly allow us to determine just how much of the present distribution of organisms represents a precise local match to present environments, and how much is a fingerprint left by the hand of history.

the Pleistocene glacial cycles . . . Techniques for the measurement of oxygen isotopes in ocean cores indicate that there may have been as many as 16 glacial cycles in the Pleistocene, each lasting for about 125,000 years (Figure 1.13a). It seems that each glacial phase may have lasted for as long as 50,000–100,000 years, with brief intervals of 10,000–20,000 years when the temperatures rose close to those we experience today. This suggests that it is present floras and faunas that are unusual, because they have developed towards the end of one of a series of unusual catastrophic warm events!

During the 20,000 years since the peak of the last glaciation, global temperatures have risen by about 8°C, and the rate at which vegetation has changed over much of this period has been detected by examining pollen records. The woody species that dominate pollen profiles at Rogers Lake in Connecticut (Figure 1.13b) have arrived in turn: spruce first and chestnut most recently. Each new arrival has added to the number of the species present, which has increased continually over the past 14,000-year period. The same picture is repeated in European profiles.

As the number of pollen records has increased, it has become possible not only to plot the changes in vegetation **. . . from which trees are still recovering**
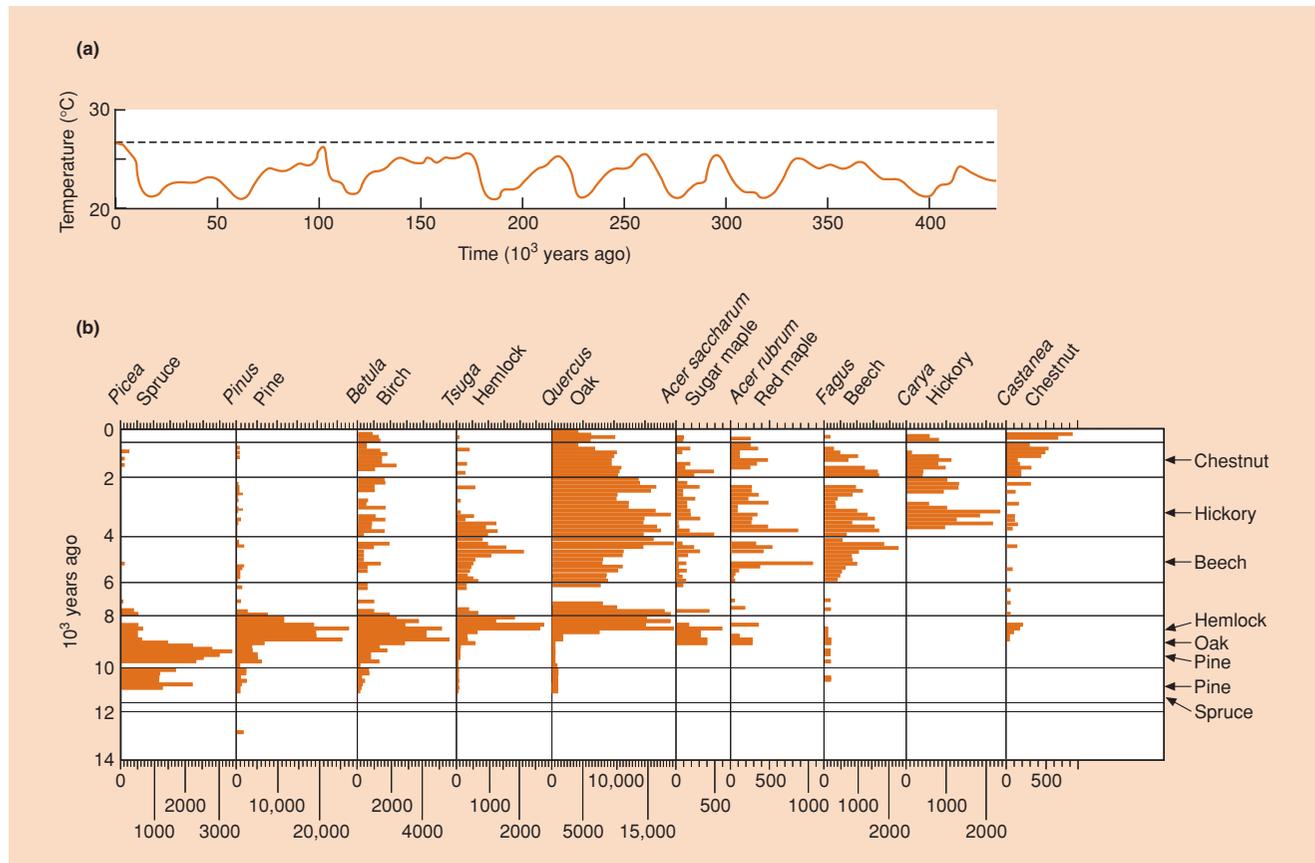
Figure 1.13 (a) An estimate of the temperature variations with time during glacial cycles over the past 400,000 years. The estimates were obtained by comparing oxygen isotope ratios in fossils taken from ocean cores in the Caribbean. The dashed line corresponds to the ratio 10,000 years ago, at the start of the present warming period. Periods as warm as the present have been rare events, and the climate during most of the past 400,000 years has been glacial. (After Emiliani, 1966; Davis, 1976.) (b) The profiles of pollen accumulated from late glacial times to the present in the sediments of Rogers Lake, Connecticut. The estimated date of arrival of each species in Connecticut is shown by arrows at the right of the figure. The horizontal scales represent pollen influx: 10^3 grains cm^{-2} year^{-1}. (After Davis et al., 1973.)
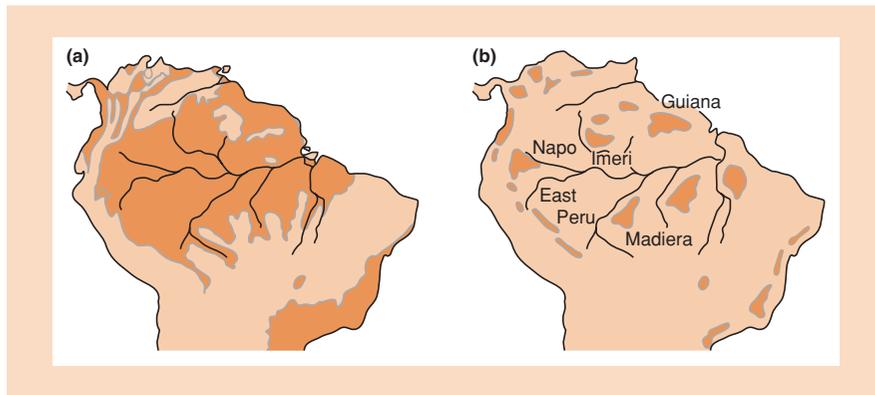
Figure 1.14 (a) The present-day distribution of tropical forest in South America. (b) The possible distribution of tropical forest refuges at the time when the last glaciation was at its peak, as judged by present-day hot spots of species diversity within the forest. (After Ridley, 1993.)

at a point in space, but to begin to map the movements of the various species as they have spread across the continents (see Bennet, 1986). In the invasions that followed the retreat of the ice in eastern North America, spruce was followed by jack pine or red pine, which spread northwards at a rate of 350–500 m year^{-1} for several thousands of years. White pine started its migration about 1000 years later, at the same time as oak. Hemlock was also one of the rapid invaders (200–300 m year^{-1}), and arrived at most sites about 1000 years after white pine. Chestnut moved slowly (100 m year^{-1}), but became a dominant species once it had arrived. Forest trees are still migrating into deglaciated areas, even now. This clearly implies that the timespan of an average interglacial period is too short for the attainment of floristic equilibrium (Davis, 1976). Such historical factors will have to be borne in mind when we consider the various patterns in species richness and biodiversity in Chapter 21.

'history' on a smaller scale

'History' may also have an impact on much smaller space and time scales. Disturbances to the benthic (bottom dwelling) community of a stream occurs when high discharge events (associated with storms or snow melt) result in a very small-scale mosaic of patches of scour (substrate loss), fill (addition of substrate) and no change (Matthaei *et al.*, 1999). The invertebrate communities associated with the different patch histories are distinctive for a period of months, within which time another high discharge event is likely to occur. As with the distribution of trees in relation to repeating ice ages, the stream fauna may rarely achieve an equilibrium between flow disturbances (Matthaei & Townsend, 2000).

changes in the tropics

The records of climatic change in the tropics are far less complete than those for temperate regions. There is therefore the temptation to imagine that whilst dramatic climatic shifts and ice invasions were dominating temperate regions, the tropics persisted in the state we know today. This is almost certainly wrong. Data from a variety of sources indicate that there were abrupt fluctuations in postglacial climates in Asia and Africa. In continental monsoon areas (e.g. Tibet, Ethiopia, western Sahara and subequatorial Africa) the postglacial period started with an extensive phase of high humidity followed by a series of phases of intense aridity (Zahn, 1994). In South America, a picture is emerging of vegetational changes that parallel those occurring in temperate regions, as the extent of tropical forest increased in warmer, wetter periods, and contracted, during cooler, drier glacial periods, to smaller patches surrounded by a sea of savanna. Support for this comes from the present-day distribution of species in the tropical forests of South America (Figure 1.14). There, particular 'hot spots' of species diversity are apparent, and these are thought to be likely sites of forest refuges during the glacial periods, and sites too, therefore, of increased rates of speciation (Prance, 1987; Ridley, 1993). On this interpretation, the present distributions of species may again be seen as largely accidents of history (where the refuges were) rather than precise matches between species and their differing environments.

how will global warming compare?

Evidence of changes in vegetation that followed the last retreat of the ice hint at the consequence of the global warming (maybe 3°C in the next 100 years) that is predicted to result from continuing increases in atmospheric carbon dioxide (discussed in detail in Sections 2.9.1 and 18.4.6). But the scales are quite different. Postglacial warming of about 8°C occurred over 20,000 years, and changes in the vegetation failed to keep pace even with this. But current projections for the 21st century require range shifts for trees at rates of 300–500 km per century compared to typical rates in the past of 20–40 km per century (and exceptional rates of 100–150 km). It is striking that the only precisely dated extinction of a tree species in the Quaternary, that of *Picea critchfeldii*, occurred around 15,000 years ago at a time of especially rapid postglacial warming (Jackson & Weng, 1999). Clearly, even more rapid change in the future could result in extinctions of many additional species (Davis & Shaw, 2001).

1.4.4 Convergents and parallels

analogous and
homologous structures

A match between the nature of organisms and their environment can often be seen as a similarity in form and behavior between organisms living in a similar environment, but belonging to different phyletic lines (i.e. different branches of the evolutionary tree). Such similarities also undermine further the idea that for every environment there is one, and only one, perfect organism. The evidence is particularly persuasive when the phyletic lines are far removed from each other, and when similar roles are played by structures that have quite different evolutionary origins, i.e. when the structures are *analogous* (similar in superficial form or function) but not *homologous* (derived from an equivalent structure in a common ancestry). When this is seen to occur, we speak of *convergent evolution.* Many flowering plants and some ferns, for example, use the support of others to climb high in the canopies of vegetation, and so gain access to more light than if they depended on their own supporting tissues. The ability to climb has evolved in many different families, and quite different organs have become modified into climbing structures (Figure 1.15a): they are analogous structures but not homologous. In other plant species the same organ has been modified into quite different structures with quite different roles: they are therefore homologous, although they may not be analogous (Figure 1.15b).

Other examples can be used to show the *parallels* in evolutionary pathways within separate groups that have radiated after they were isolated from each other. The classic example of such parallel evolution is the radiation amongst the placental and marsupial mammals. Marsupials arrived on the Australian continent in the Cretaceous period (around 90 million years ago), when the only other mammals present were the curious egg-laying monotremes (now represented only by the spiny anteaters (*Tachyglossus aculeatus*) and the duckbill platypus (*Ornithorynchus anatinus*)). An evolutionary process of radiation then occurred that in many

(a)

Dioscorea (Dioscoreaceae), twiner

Calamus (Arecaceae), hooks

Clematis (Ranunculaceae), twining petiole

Cobaea (Cobaeaceae), tendril

Ficus (Moraceae), adventitious roots

*Parthenocissus* (Vitaceae), sticky pads
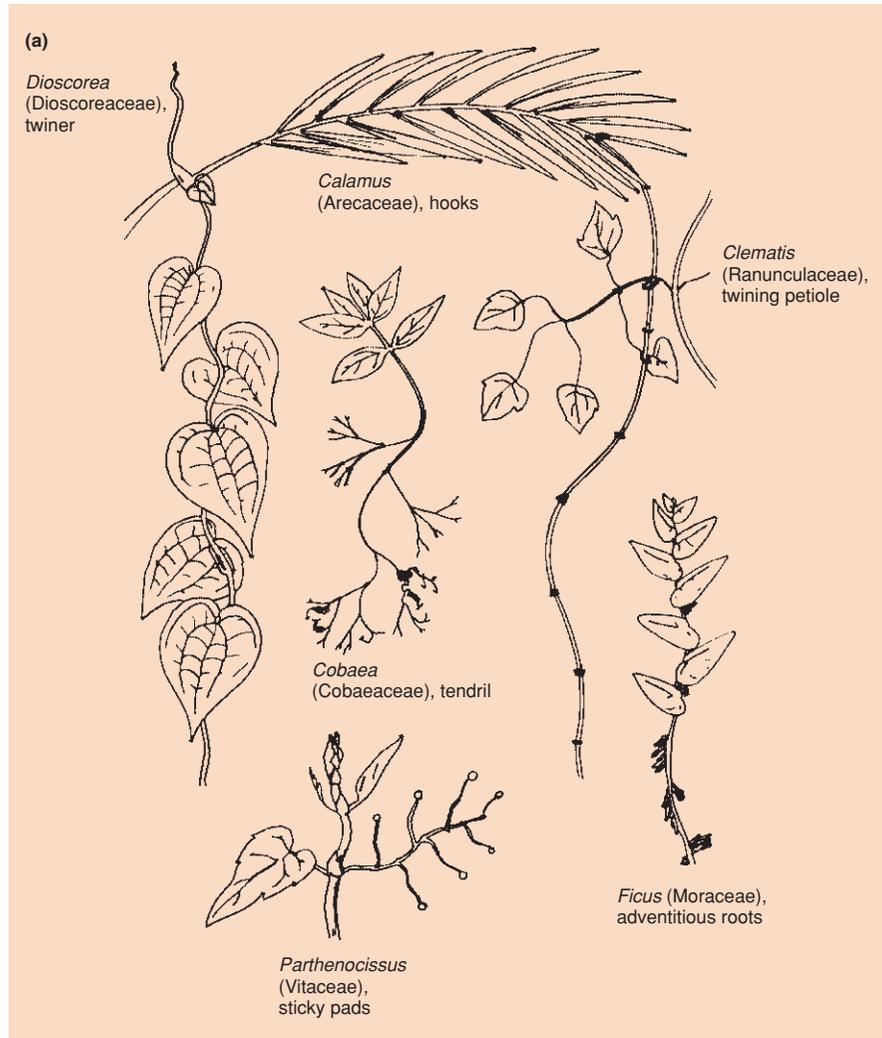
Figure 1.15 A variety of morphological features that allow flowering plants to climb. (a) Structural features that are analogous, i.e. derived from modifications of quite different organs, e.g. leaves, petioles, stems, roots and tendrils.
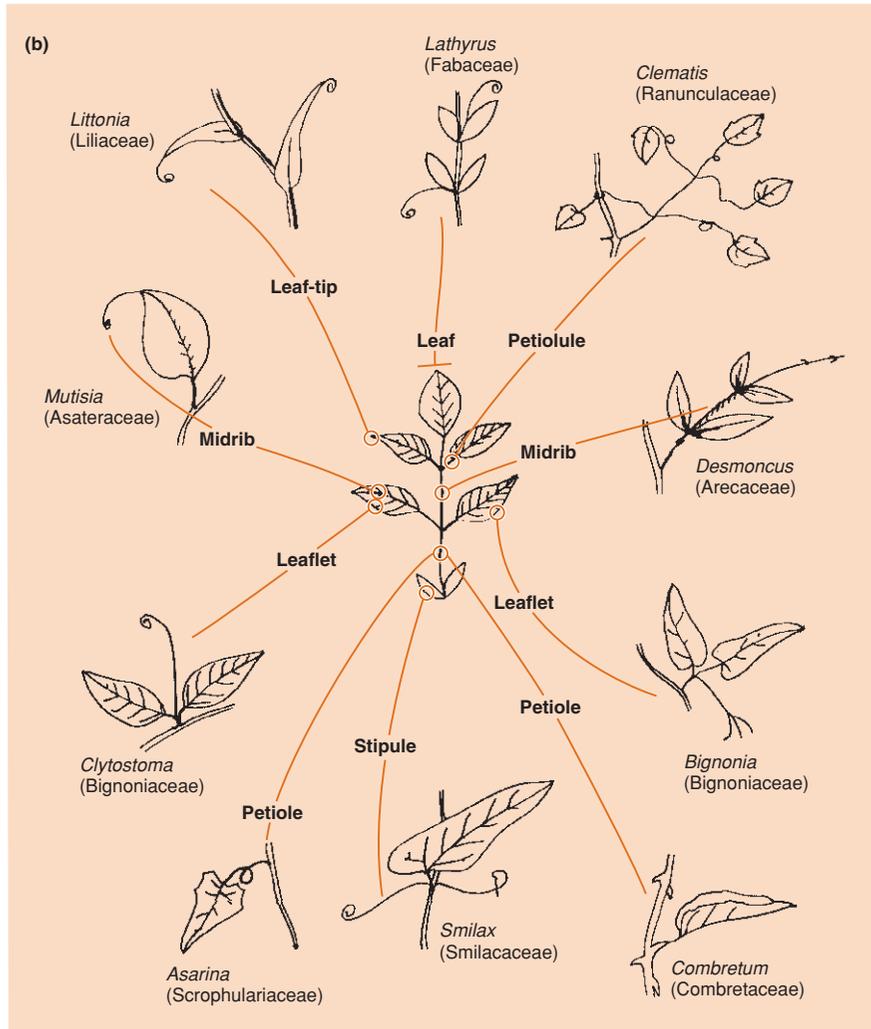
Figure 1.15   (*continued*) (b) Structural features that are homologous, i.e. derived from modifications of a single organ, the leaf, shown by reference to an idealized leaf in the center of the figure. (Courtesy of Alan Bryant.)

ways accurately paralleled what occurred in the placental mammals on other continents (Figure 1.16). The subtlety of the parallels in both the form of the organisms and their lifestyle is so striking that it is hard to escape the view that the environments of placentals and marsupials provided similar opportunities to which the evolutionary processes of the two groups responded in similar ways.

1.5   The match between communities and their environments

1.5.1   Terrestrial biomes of the earth

Before we examine the differences and similarities between communities, we need to consider the larger groupings, 'biomes', in which biogeographers recognize marked differences in the flora and fauna of different parts of the world. The number of biomes that are distinguished is a matter of taste. They certainly grade into one another, and sharp boundaries are a convenience for cartographers rather than a reality of nature. We describe eight terrestrial biomes and illustrate their global distribution in Figure 1.17, and show how they may be related to annual temperature and precipitation (Figure 1.18) (see Woodward, 1987 for a more detailed account). Apart from anything else, understanding the terminology that describes and distinguishes these biomes is necessary when we come to consider key questions later in the book (especially in Chapters 20 and 21). Why are there more species in some communities than in others? Are some communities more stable in their composition than others, and if so why? Do more productive environments support more diverse communities? Or do more diverse communities make more productive use of the resources available to them?

Placentals	Marsupials
Dog-like carnivore	
Wolf (*Canis*)	Tasmanian wolf (*Thylacinus*)
Cat-like carnivore	
Ocelot (*Felis*)	Native cat (*Dasyurus*)
Arboreal glider	
Flying squirrel (*Glaucomys*)	Flying phalanger (*Petaurus*)
Fossorial herbivore	
Ground hog (*Marmota*)	Wombat (*Vombatus*)
Digging ant feeder	
Anteater (*Myrmecophaga*)	Anteater (*Myrmecabius*)
Subterranean insectivore	
Common mole (*Talpa*)	Marsupial mole (*Notoryctes*)

Figure 1.16   Parallel evolution of marsupial and placental mammals. The pairs of species are similar in both appearance and habit, and usually (but not always) in lifestyle.

tundra

Tundra (see Plate 1.1, facing p. XX) occurs around the Arctic Circle, beyond the tree line. Small areas also occur on sub-Antarctic islands in the southern hemisphere. 'Alpine' tundra is found under similar conditions but at high altitude. The environment is characterized by the presence of permafrost – water permanently frozen in the soil – while liquid water is present for only short periods of the year. The typical flora includes lichens, mosses, grasses, sedges and dwarf trees. Insects are extremely seasonal in their activity, and the native bird and mammal fauna is enriched by species that migrate from warmer latitudes in the summer. In the colder areas, grasses and sedges disappear, leaving nothing rooted in the permafrost. Ultimately, vegetation that consists only of lichens and mosses gives way, in its turn, to the polar desert. The number of species of higher plants (i.e. excluding mosses and lichens) decreases
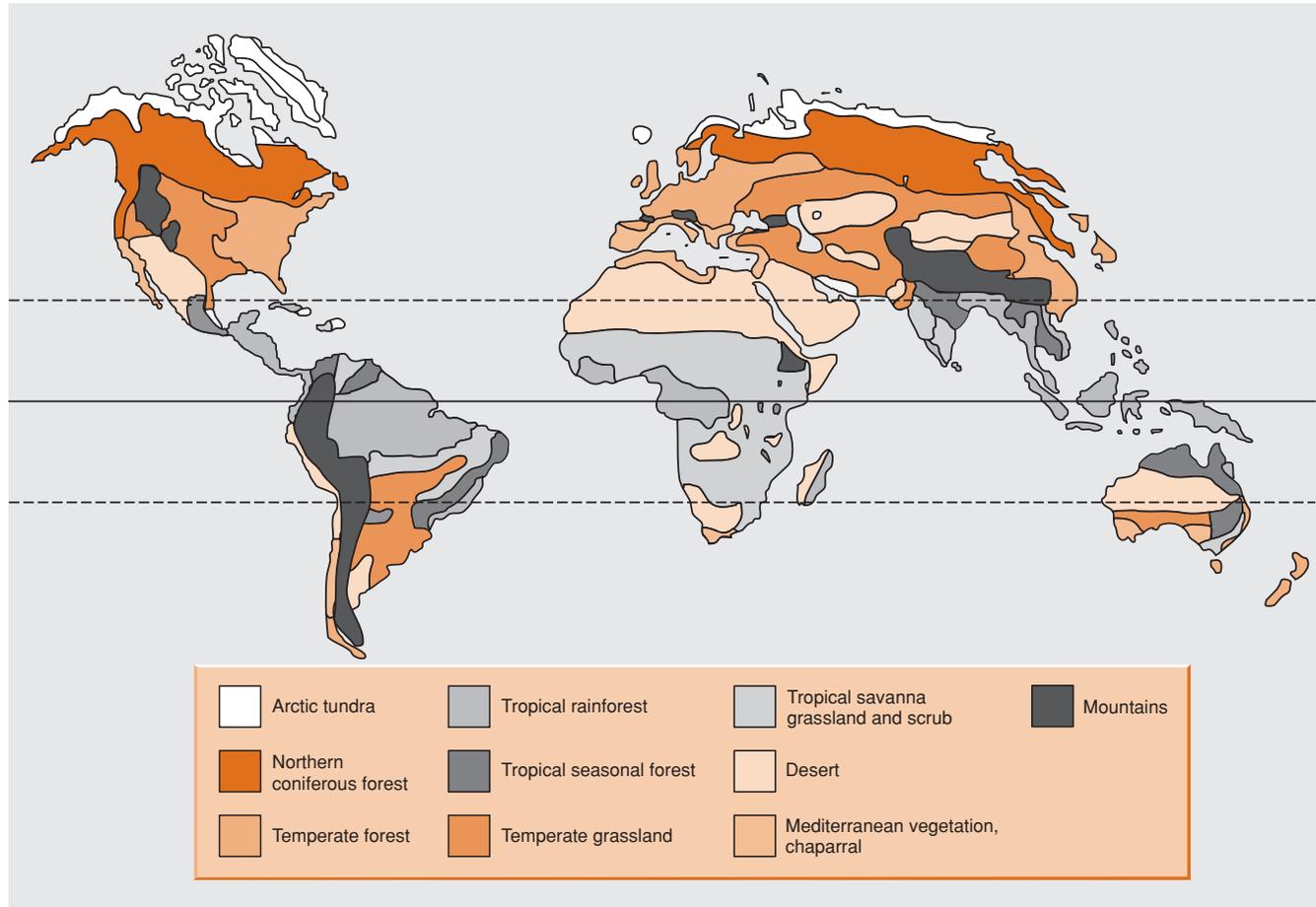
Figure 1.17   World distribution of the major biomes of vegetation. (After Audesirk & Audesirk, 1996.)

from the Low Arctic (around 600 species in North America) to the High Arctic (north of 83°, e.g. around 100 species in Greenland and Ellesmere Island). In contrast, the flora of Antarctica contains only two native species of vascular plant and some lichens and mosses that support a few small invertebrates. The biological productivity and diversity of Antarctica are concentrated at the coast and depend almost entirely on resources harvested from the sea.

taiga

Taiga or northern coniferous forest (see Plate 1.2, facing p. XX) occupies a broad belt across North America and Eurasia. Liquid water is unavailable for much of the winter, and plants and many of the animals have a conspicuous winter dormancy in which metabolism is very slow. Generally, the tree flora is very limited. In areas with less severe winters, the forests may be dominated by pines (*Pinus* species, which are all evergreens) and deciduous trees such as larch (*Larix*), birch (*Betula*) or aspens (*Populus*), often as mixtures of species. Farther north, these species give way to single-species forests of spruce (*Picea*) covering immense areas. The overriding environmental constraint in

northern spruce forests is the presence of permafrost, creating drought except when the sun warms the surface. The root system of spruce can develop in the superficial soil layer, from which the trees derive all their water during the short growing season.

Temperate forests (see Plate 1.3, between pp. XX and XX) range from the mixed conifer and broad-leaved forests

temperate forests

of much of North America and northern central Europe (where there may be 6 months of freezing temperatures), to the moist dripping forests of broad-leaved evergreen trees found at the biome's low latitude limits in, for example, Florida and New Zealand. In most temperate forests, however, there are periods of the year when liquid water is in short supply, because potential evaporation exceeds the sum of precipitation and water available from the soil. Deciduous trees, which dominate in most temperate forests, lose their leaves in the fall and become dormant. On the forest floor, diverse floras of perennial herbs often occur, particularly those that grow quickly in the spring before the new tree foliage has developed. Temperate forests also
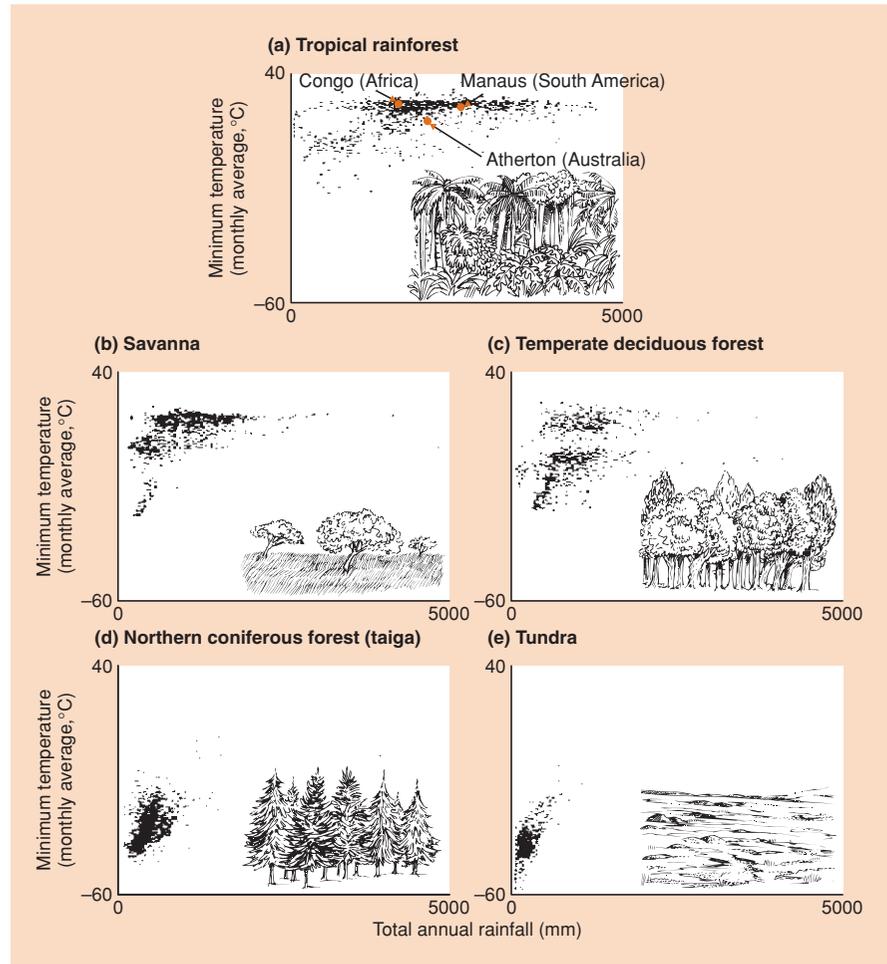
Figure 1.18 The variety of environmental conditions experienced in terrestrial environments can be described in terms of their annual rainfall and mean monthly minimum temperatures. The range of conditions experienced in: (a) tropical rainforest, (b) savanna, (c) temperate deciduous forest, (d) northern coniferous forest (taiga), and (e) tundra. (After Heal *et al.*, 1993; © UNESCO.)

provide food resources for animals that are usually very seasonal in their occurrence. Many of the birds of temperate forests are migrants that return in spring but spend the remainder of the year in warmer biomes.

grassland

Grassland occupies the drier parts of temperate and tropical regions. Temperate grassland has many local names: the steppes of Asia, the prairies of North America, the pampas of South America and the veldt of South Africa. Tropical grassland or savanna (see Plate 1.4, between pp. XX and XX) is the name applied to tropical vegetation ranging from pure grassland to some trees with much grass. Almost all of these temperate and tropical grasslands experience seasonal drought, but the role of climate in determining their vegetation is almost completely overridden by the effects of grazing animals that limit the species present to those that can recover from frequent defoliation. In the savanna, fire is also a common hazard in the dry season and, like grazing animals, it tips the balance in the vegetation against trees and towards grassland. None the less, there is typically a seasonal glut of food, alternating with shortage, and as a consequence

the larger grazing animals suffer extreme famine (and mortality) in drier years. A seasonal abundance of seeds and insects supports large populations of migrating birds, but only a few species can find sufficiently reliable resources to be resident year-round.

Many of these natural grasslands have been cultivated and replaced by arable annual 'grasslands' of wheat, oats, barley, rye and corn. Such annual grasses of temperate regions, together with rice in the tropics, provide the staple food of human populations worldwide. At the drier margins of the biome, many of the grasslands are 'managed' for meat or milk production, sometimes requiring a nomadic human lifestyle. The natural populations of grazing animals have been driven back in favor of cattle, sheep and goats. Of all the biomes, this is the one most coveted, used and transformed by humans.

Chaparral or *maquis* occurs in Mediterranean-type climates (mild, wet winters and summer drought) in Europe, California and northwest Mexico, and in a few small areas in Australia, Chile and South Africa. Chaparral develops in regions with less rainfall than temperate grasslands and is dominated mainly by a

chaparral

drought-resistant, hard-leaved scrub of low-growing woody plants. Annual plants are also common in chaparral regions during the winter and early spring, when rainfall is more abundant. Chaparral is subject to periodic fires; many plants produce seeds that will only germinate after fire while others can quickly resprout because of food reserves in their fire-resistant roots.

desert

Deserts (see Plate 1.5, between pp. XX and XX) are found in areas that experience extreme water shortage: rainfall is usually less than about 25 cm year^{-1}, is usually very unpredictable and is considerably less than potential evaporation. The desert biome spans a very wide range of temperatures, from hot deserts, such as the Sahara, to very cold deserts, such as the Gobi in Mongolia. In their most extreme form, the hot deserts are too arid to bear any vegetation; they are as bare as the cold deserts of Antarctica. Where there is sufficient rainfall to allow plants to grow in arid deserts, its timing is always unpredictable. Desert vegetation falls into two sharply contrasted patterns of behavior. Many species have an opportunistic lifestyle, stimulated into germination by the unpredictable rains. They grow fast and complete their life history by starting to set new seed after a few weeks. These are the species that can occasionally make a desert bloom. A different pattern of behavior is to be long-lived with sluggish physiological processes. Cacti and other succulents, and small shrubby species with small, thick and often hairy leaves, can close their stomata (pores through which gas exchange takes place) and tolerate long periods of physiological inactivity. The relative poverty of animal life in arid deserts reflects the low productivity of the vegetation and the indigestibility of much of it.

tropical rainforest

Tropical rainforest (see Plate 1.6, between pp. XX and XX) is the most productive of the earth's biomes – a result of the coincidence of high solar radiation received throughout the year and regular and reliable rainfall. The productivity is achieved, overwhelmingly, high in the dense forest canopy of evergreen foliage. It is dark at ground level except where fallen trees create gaps. Often, many tree seedlings and saplings remain in a suppressed state from year to year and only leap into action if a gap forms in the canopy above them. Apart from the trees, the vegetation is largely composed of plant forms that reach up into the canopy vicariously; they either climb and then scramble in the tree canopy (vines and lianas, including many species of fig) or grow as epiphytes, rooted on the damp upper branches. Most species of both animals and plants in tropical rain forest are active throughout the year, though the plants may flower and ripen fruit in sequence. Dramatically high species richness is the norm for tropical rainforest, and communities rarely if ever become dominated by one or a few species. The diversity of rainforest trees provides for a corresponding diversity of resources for herbivores, and so on up the food chain. Erwin (1982) estimated that there are 18,000 species of beetle in 1 ha of Panamanian rainforest (compared with only 24,000 in the whole of the United States and Canada!).

aquatic biomes?

All of these biomes are terrestrial. Aquatic ecologists could also come up with a set of biomes, although the tradition has largely been a terrestrial one. We might distinguish springs, rivers, ponds, lakes, estuaries, coastal zones, coral reefs and deep oceans, among other distinctive kinds of aquatic community. For present purposes, we recognize just two aquatic biomes, *marine* and *freshwater*. The oceans cover about 71% of the earth's surface and reach depths of more than 10,000 m. They extend from regions where precipitation exceeds evaporation to regions where the opposite is true. There are massive movements within this body of water that prevent major differences in salt concentrations developing (the average concentration is about 3%). Two main factors influence the biological activity of the oceans. Photosynthetically active radiation is absorbed in its passage through water, so photosynthesis is confined to the surface region. Mineral nutrients, especially nitrogen and phosphorus, are commonly so dilute that they limit the biomass that can develop. Shallow waters (e.g. coastal regions and estuaries) tend to have high biological activity because they receive mineral input from the land and less incident radiation is lost than in passage through deep waters. Intense biological activity also occurs where nutrient-rich waters from the ocean depths come to the surface; this accounts for the concentration of many of the world's fisheries in Arctic and Antarctic waters.

Freshwater biomes occur mainly on the route from land drainage to the sea. The chemical composition of the water varies enormously, depending on its source, its rate of flow and the inputs of organic matter from vegetation that is rooted in or around the aquatic environment. In water catchments where the rate of evaporation is high, salts leached from the land may accumulate and the concentrations may far exceed those present in the oceans; brine lakes or even salt pans may be formed in which little life is possible. Even in aquatic situations liquid water may be unavailable, as is the case in the polar regions.

Differentiating between biomes allows only a very crude recognition of the sorts of differences and similarities that occur between communities of organisms. Within biomes there are both small- and large-scale patterns of variation in the structure of communities and in the organisms that inhabit them. Moreover, as we see next, what characterizes a biome is not necessarily the particular species that live there.

1.5.2 The 'life form spectra' of communities

We pointed out earlier the crucial importance of geographic isolation in allowing populations to diverge under selection. The geographic distributions of species, genera, families and even higher taxonomic categories of plants and animals often reflect this geographic divergence. All species of lemurs, for example, are found on the island of Madagascar and nowhere else. Similarly,

230 species in the genus *Eucalyptus* (gum tree) occur naturally in Australia (and two or three in Indonesia and Malaysia). The lemurs and the gum trees occur where they do because they evolved there – not because these are the only places where they could survive and prosper. Indeed, many *Eucalyptus* species grow with great success and spread rapidly when they have been introduced to California or Kenya. A map of the natural world distribution of lemurs tells us quite a lot about the evolutionary history of this group. But as far as its relationship with a biome is concerned, the most we can say is that lemurs happen to be one of the constituents of the tropical rainforest biome in Madagascar.

Similarly, particular biomes in Australia include certain marsupial mammals, while the *same* biomes in other parts of the world are home to their placental counterparts. A map of biomes, then, is not usually a map of the distribution of species. Instead, we recognize different biomes and different types of aquatic community from the *types* of organisms that live in them. How can we describe their similarities so that we can classify, compare and map them? In addressing this question, the Danish biogeographer Raunkiaer developed, in 1934, his idea of 'life forms', a deep insight into the ecological significance of plant forms (Figure 1.19). He then used the spectrum of life forms present in different types of vegetation as a means of describing their ecological character.

Raunkiaer's classification

Plants grow by developing new shoots from the buds that lie at the apices (tips) of existing shoots and in the leaf axils. Within the buds, the meristematic cells are the most sensitive part of the whole shoot – the 'Achilles' heel' of plants. Raunkiaer argued that the ways in which these buds are protected in different plants are powerful indicators of the hazards in their environments and may be used to define the different plant forms (Figure 1.19). Thus, trees expose their buds high in the air, fully exposed to the wind, cold and drought; Raunkiaer called them *phanerophytes* (Greek *phanero*, 'visible'; *phyte*, 'plant'). By contrast, many perennial herbs form cushions or tussocks in which buds are borne above ground but are protected from drought and cold in the dense mass of old leaves and shoots (*chamaephytes*: 'on the ground plants'). Buds are even better protected when they are formed at or in the soil surface (*hemicryptophytes*: 'half hidden plants') or on buried dormant storage organs (bulbs, corms and rhizomes – *cryptophytes*: 'hidden plants'; or *geophytes*: 'earth plants'). These allow the plants to make rapid growth and to flower before they die back to a dormant state. A final major category consists of annual plants that depend wholly on dormant seeds to carry their populations through seasons of drought and cold (*therophytes*: 'summer plants'). Therophytes are the plants of deserts (they make up nearly 50% of the flora of Death Valley, USA), sand dunes and repeatedly disturbed habitats. They also include the annual weeds of arable lands, gardens and urban wastelands.

But there is, of course, no vegetation that consists entirely of one growth form. All vegetation contains a mixture, a spectrum, of Raunkiaer's life forms. The composition of the spectrum in any particular habitat is as good a shorthand description of its vegetation as ecologists have yet managed to devise. Raunkiaer compared these with a 'global spectrum' obtained by sampling from a compendium of all species known and described in his time (the *Index Kewensis*), biased by the fact that the tropics were, and still are, relatively unexplored. Thus, for example, we recognize a chaparral type of vegetation when we see it in Chile, Australia, California or Crete because the life form spectrums are similar. Their detailed taxonomies would only emphasize how different they are.

Faunas are bound to be closely tied to floras – if only because most herbivores are choosy about their diet. Terrestrial carnivores range more widely than their herbivore prey, but the distribution of herbivores still gives the carnivores a broad vegetational allegiance. Plant scientists have tended to be keener on classifying floras than animal scientists on classifying faunas, but one interesting attempt to classify faunas compared the mammals of forests in Malaya, Panama, Australia and Zaire (Andrews *et al.*, 1979). They were classified into carnivores, herbivores, insectivores and mixed feeders, and these categories were subdivided into those that were aerial (mainly bats and flying foxes), arboreal (tree dwellers), scansorial (climbers) or small ground mammals (Figure 1.20). The comparison reveals some strong contrasts and similarities. For example, the ecological diversity spectra for the Australian and Malayan forests were very similar despite the fact that their faunas are taxonomically very distinct – the Australian mammals are marsupials and the Malaysian mammals are placentals.

1.6 The diversity of matches within communities

Although a particular type of organism is often characteristic of a particular ecological situation, it will almost inevitably be only part of a diverse community of species. A satisfactory account, therefore, must do more than identify the similarities between organisms that allow them to live in the same environment – it must also try to explain why species that live in the same environment are often profoundly different. To some extent, this 'explanation' of diversity is a trivial exercise. It comes as no surprise that a plant utilizing sunlight, a fungus living on the plant, a herbivore eating the plant and a parasitic worm living in the herbivore should all coexist in the same community. On the other hand, most communities also contain a variety of different species that are all constructed in a fairly similar way and all living (at least superficially) a fairly similar life. There are several elements in an explanation of this diversity.

1.6.1 Environments are heterogeneous

There are no homogeneous environments in nature. Even a continuously stirred culture of microorganisms is heterogeneous
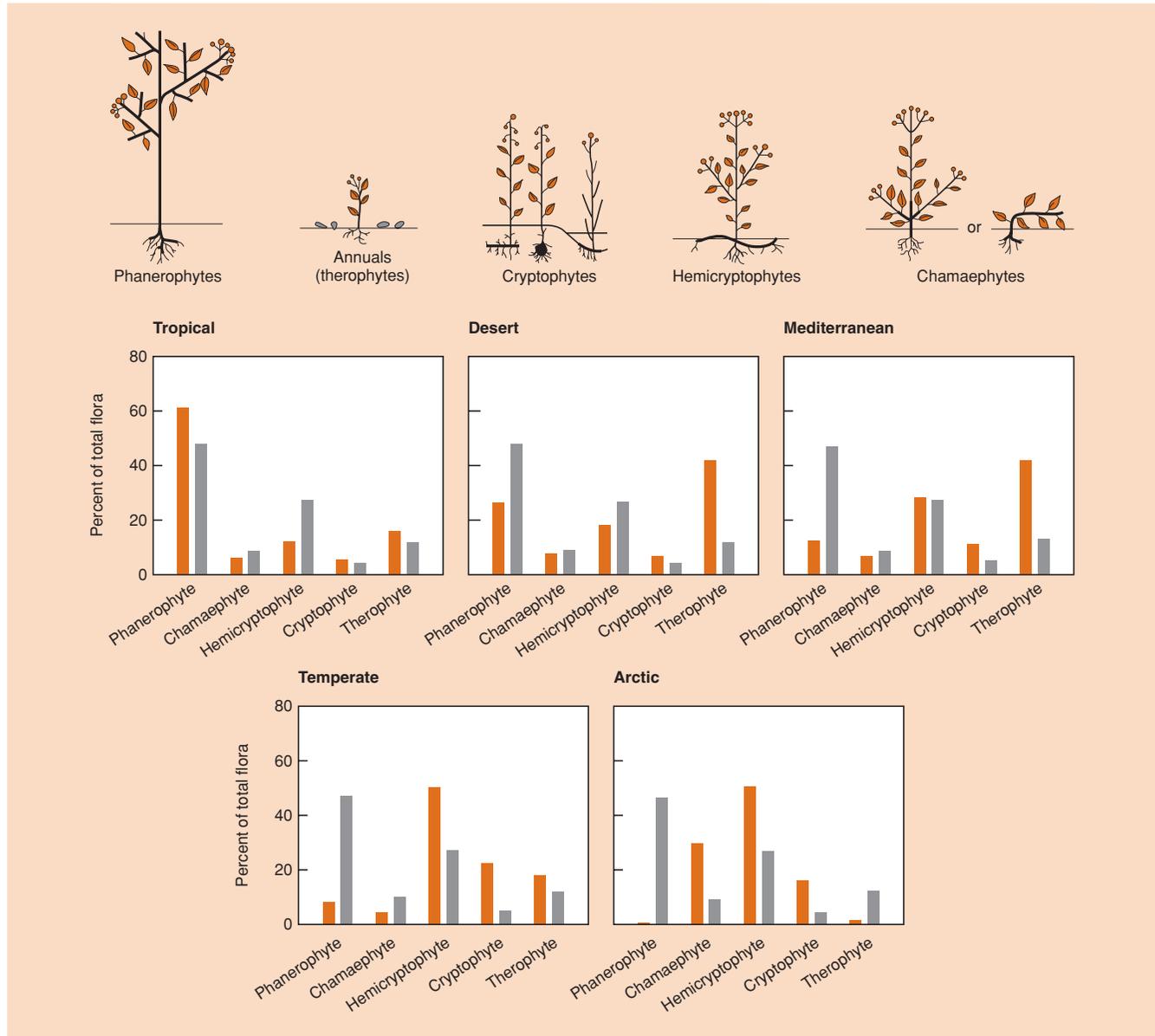
Figure 1.19 The drawings above depict the variety of plant forms distinguished by Raunkiaer on the basis of where they bear their buds (shown in color). Below are life form spectrums for five different biomes. The colored bars show the percentage of the total flora that is composed of species with each of the five different life forms. The gray bars are the proportions of the various life forms in the world flora for comparison. (From Crawley, 1986.)

because it has a boundary – the walls of the culture vessel – and cultured microorganisms often subdivide into two forms: one that sticks to the walls and the other that remains free in the medium.

The extent to which an environment is heterogeneous depends on the scale of the organism that senses it. To a mustard seed, a grain of soil is a mountain; and to a caterpillar, a single leaf may represent a lifetime's diet. A seed lying in the shadow of a leaf may be inhibited in its germination while a seed lying outside that shadow germinates freely. What appears to the human observer as a homogeneous environment may, to an organism within it, be a mosaic of the intolerable and the adequate.

There may also be gradients in space (e.g. altitude) or gradients in time, and the latter, in their turn, may be rhythmic (like
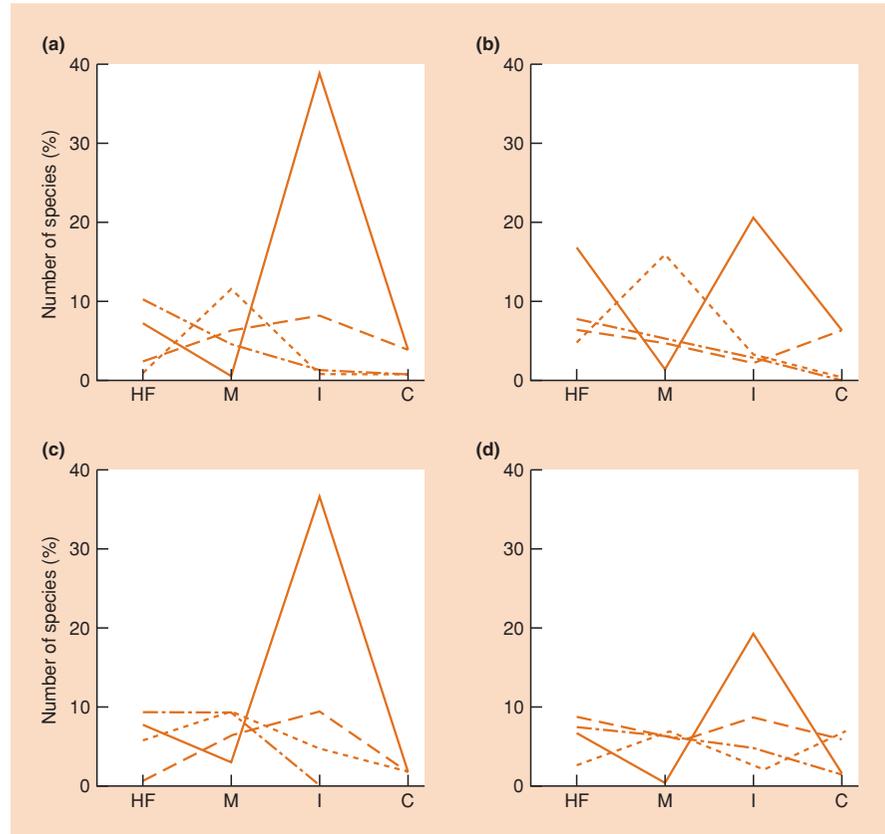
Figure 1.20 The percentages of forest mammals in various locomotory and feeding habitat categories in communities in: (a) Malaya, all forested areas (161 species), (b) Panama dry forest (70 species), (c) Australia, Cape York forest (50 species), and (d) Zaire, Irangi forest (96 species). C, carnivores; HF, herbivores and fructivores; I, insectivores; M, mixed feeders; (——) aerial; (·····) arboreal; (– – –) scansorial; (—·—) small ground mammals. (After Andrews *et al.*, 1979.)

daily and seasonal cycles), directional (like the accumulation of a pollutant in a lake) or erratic (like fires, hailstorms and typhoons).

Heterogeneity crops up again and again in later chapters – in part because of the challenges it poses to organisms in moving from patch to patch (Chapter 6), in part because of the variety of opportunities it provides for different species (Chapters 8 and 19), and in part because heterogeneity can alter communities by interrupting what would otherwise be a steady march to an equilibrium state (Chapters 10 and 19).

1.6.2 Pairs of species

As we have already noted, the existence of one type of organism in an area immediately diversifies it for others. Over its lifetime, an organism may increase the diversity of its environment by contributing dung, urine, dead leaves and ultimately its dead body. During its life, its body may serve as a place in which other species find homes. Indeed, some of the most strongly developed matches between organisms and their environment are those in which one species has developed a dependence upon another. This is the case in many relationships between consumers and their foods. Whole syndromes of form, behavior and metabolism constrain the animal within its narrow food niche, and deny it access to what might otherwise appear suitable alternative foods. Similar tight matches are characteristic of the relationships between parasites and their hosts. The various interactions in which one species is consumed by another are the subject matter of Chapters 9–12.

Where two species have evolved a mutual dependence, the fit may be even tighter. We examine such 'mutualisms' in detail in Chapter 13. The association of nitrogen-fixing bacteria with the roots of leguminous plants, and the often extremely precise relationships between insect pollinators and their flowers, are two good examples.

When a population has been exposed to variations in the physical factors of the environment, for example a short growing season or a high risk of frost or drought, a once-and-for-all tolerance may ultimately evolve. The physical factor cannot itself change or evolve as a result of the evolution of the organisms. By contrast, when members of two species interact, the change in each produces alterations in the life of the other, and each may generate selective forces that direct the evolution of the other. In such a coevolutionary process the interaction between two species may continually escalate. What we then see in nature may be pairs of species that have driven each other into ever narrowing ruts of specialization – an ever closer match.
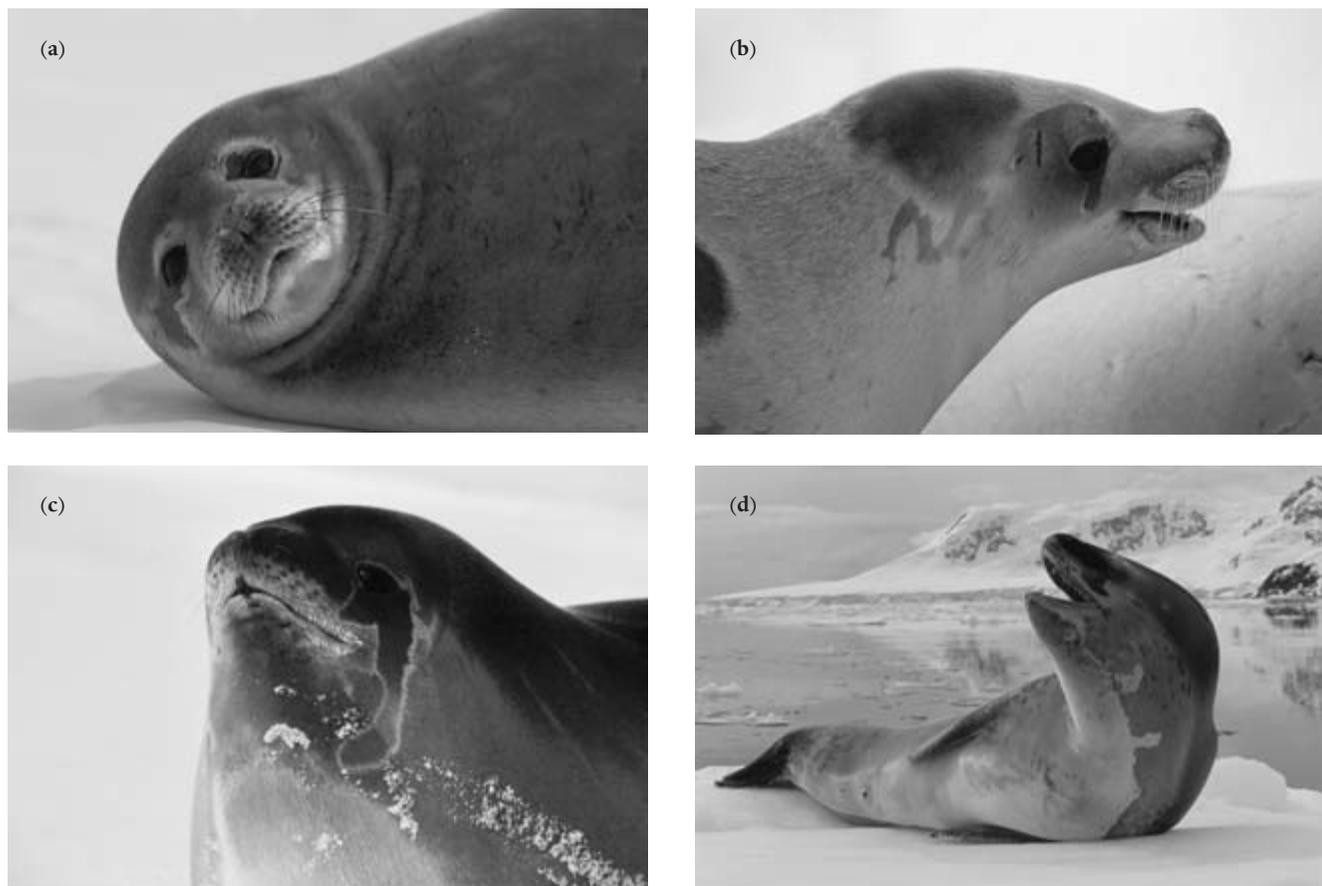
Figure 1.21   Antarctic seals, similar species that coexist: (a) the Weddell seal, *Leptonychotes weddellii* (© Imageshop – zefa visual media uk ltd / Alamy), (b) the crab-eater seal *Lobodon carcinophagus* (© Bryan & Cherry Alexander Photography / Alamy), (c) the Ross seal, *Omatophoca rossii* (© Chris Sattlberger / Science Photo Library), and (d) the leopard seal, *Hydrurga leptonyx* (© Kevin Schafer / Alamy).

1.6.3   Coexistence of similar species

While it is no surprise that species with rather different roles coexist within the same community, it is also generally the case that communities support a variety of species performing apparently rather similar roles. The Antarctic seals are an example. It is thought that the ancestral seals evolved in the northern hemisphere, where they are present as Miocene fossils, but one group of seals moved south into warmer waters and probably colonized the Antarctic in the Late Miocene or Early Pliocene (about 5 million years ago). When they entered the Antarctic, the Southern Ocean was probably rich in food and free from major predators, as it is today. It was within this environment that the group appears to have undergone radiative evolution (Figure 1.21). For example, the Weddell seal feeds primarily on fish and has unspecialized dentition; the crab-eater seal feeds almost exclusively on krill and its teeth are suited to filtering these from the sea water; the Ross seal has small, sharp teeth and feeds mainly on pelagic squid; and the leopard seal has large, cusped, grasping teeth and feeds on a wide variety of foods, including other seals and, in some seasons, penguins.

Do these species compete with one another? Do competing species need to be different if they are to coexist? If so, how different do they need to be: is there some limit to their similarity? Do species like the seals interact with one another at the present time, or has evolution in the past led to the absence of such interactions in contemporary communities? We return to these questions about coexisting, similar species in Chapter 8.

Even at this stage, though, we may note that coexisting species, even when apparently very similar, commonly differ in subtle ways – not simply in their morphology or physiology but also in their responses to their environment and the role they play within the community of which they are part. The 'ecological niches' of such species are said to be differentiated from one another. The concept of the ecological niche is itself explained in the next two chapters.

Summary

'Nothing in biology makes sense, except in the light of evolution'. We try in this chapter to illustrate the processes by which the properties of different sorts of species make their life possible in particular environments.

We explain what is meant by evolutionary adaptation and by the theory of evolution by natural selection, an ecological theory first elaborated by Charles Darwin in 1859. Through natural selection, organisms come to match their environments by being 'the fittest available' or 'the fittest yet': they are not 'the best imaginable'.

Adaptive variation within species can occur at a range of levels: all represent a balance between local adaptation and hybridization. Ecotypes are genetically determined variants between populations within a species that reflect local matches between the organisms and their environments. Genetic polymorphism is the occurrence together in the same habitat of two or more distinct forms. Dramatic examples of local specialization have been driven by manmade ecological forces, especially those of environmental pollution.

We describe the process of speciation by which two or more new species are formed from one original species and explain what we mean by a 'species', especially a biospecies. Islands provide arguably the most favorable environment for populations to diverge into distinct species.

Species live where they do for reasons that are often accidents of history. We illustrate this by examining island patterns, the movements of land masses over geological time, climatic changes especially during the Pleistocene ice ages (and we compare this with predicted changes consequent on current global warming) and the concepts of convergent and parallel evolution.

The various terrestrial biomes of the earth are reviewed and their aquatic equivalents touched on briefly. Raunkiaer's concept of life form spectra, in particular, emphasizes that ecological communities may be fundamentally very similar even when taxonomically quite distinct.

All communities comprise a diversity of species: a diversity of matches to the local environment. Environmental heterogeneity, interactions between predators and prey, parasites and hosts and mutualists, and the coexistence of similar species all contribute to this.
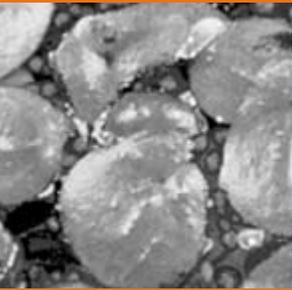
Chapter 2
Conditions

2.1 Introduction

In order to understand the distribution and abundance of a species we need to know its history (Chapter 1), the resources it requires (Chapter 3), the individuals' rates of birth, death and migration (Chapters 4 and 6), their interactions with their own and other species (Chapters 5 and 8–13) and the effects of environmental conditions. This chapter deals with the limits placed on organisms by environmental conditions.

conditions may be altered – but not consumed

A condition is as an abiotic environmental factor that influences the functioning of living organisms. Examples include temperature, relative humidity, pH, salinity and the concentration of pollutants. A condition may be modified by the presence of other organisms. For example, temperature, humidity and soil pH may be altered under a forest canopy. But unlike resources, conditions are not consumed or used up by organisms.

For some conditions we can recognize an optimum concentration or level at which an organism performs best, with its activity tailing off at both lower and higher levels (Figure 2.1a). But we need to define what we mean by 'performs best'. From an evolutionary point of view, 'optimal' conditions are those under which individuals leave most descendants (are fittest), but these are often impossible to determine in practice because measures of fitness should be made over several generations. Instead, we more often measure the effect of conditions on some key property like the activity of an enzyme, the respiration rate of a tissue, the growth rate of individuals or their rate of reproduction. However, the effect of variation in conditions on these various properties will often not be the same; organisms can usually survive over a wider range of conditions than permit them to grow or reproduce (Figure 2.1a).

The precise shape of a species' response will vary from condition to condition. The generalized form of response, shown in Figure 2.1a, is appropriate for conditions like temperature and pH
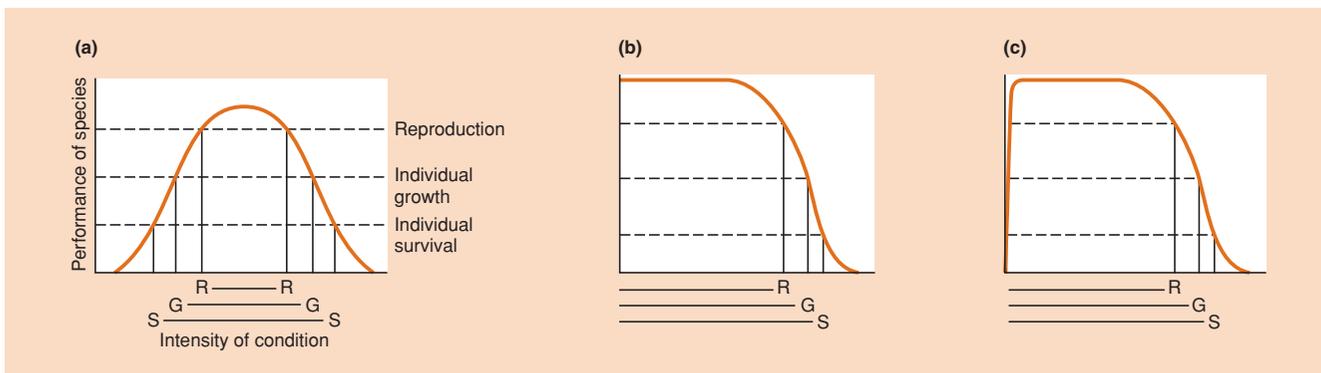
Figure 2.1 Response curves illustrating the effects of a range of environmental conditions on individual survival (S), growth (G) and reproduction (R). (a) Extreme conditions are lethal; less extreme conditions prevent growth; only optimal conditions allow reproduction. (b) The condition is lethal only at high intensities; the reproduction–growth–survival sequence still applies. (c) Similar to (b), but the condition is required by organisms, as a resource, at low concentrations.

in which there is a continuum from an adverse or lethal level (e.g. freezing or very acid conditions), through favorable levels of the condition to a further adverse or lethal level (heat damage or very alkaline conditions). There are, though, many environmental conditions for which Figure 2.1b is a more appropriate response curve: for instance, most toxins, radioactive emissions and chemical pollutants, where a low-level intensity or concentration of the condition has no detectable effect, but an increase begins to cause damage and a further increase may be lethal. There is also a different form of response to conditions that are toxic at high levels but essential for growth at low levels (Figure 2.1c). This is the case for sodium chloride – an essential resource for animals but lethal at high concentrations – and for the many elements that are essential micronutrients in the growth of plants and animals (e.g. copper, zinc and manganese), but that can become lethal at the higher concentrations sometimes caused by industrial pollution.

In this chapter, we consider responses to temperature in much more detail than other conditions, because it is the single most important condition that affects the lives of organisms, and many of the generalizations that we make have widespread relevance. We move on to consider a range of other conditions, before returning, full circle, to temperature because of the effects of other conditions, notably pollutants, on global warming. We begin, though, by explaining the framework within which each of these conditions should be understood here: the ecological niche.

2.2 Ecological niches

The term *ecological niche* is frequently misunderstood and misused. It is often used loosely to describe the sort of place in which an organism lives, as in the sentence: 'Woodlands are the niche of woodpeckers'. Strictly, however, where an organism lives is its *habitat*. A niche is not a place but an idea: a summary of the organism's tolerances and requirements. The habitat of a gut microorganism would be an animal's alimentary canal; the habitat of an aphid might be a garden; and the habitat of a fish could be a whole lake. Each habitat, however, provides many different niches: many other organisms also live in the gut, the garden or the lake – and with quite different lifestyles. The word *niche* began to gain its present scientific meaning when Elton wrote in 1933 that the niche of an organism is its mode of life 'in the sense that we speak of trades or jobs or professions in a human community'. The niche of an organism started to be used to describe how, rather than just where, an organism lives.

niche dimensions

The modern concept of the niche was proposed by Hutchinson in 1957 to address the ways in which tolerances and requirements interact to define the conditions (this chapter) and resources (Chapter 3) needed by an individual or a species in order to practice its way of life. Temperature, for instance, limits the growth and reproduction of all organisms, but different organisms tolerate different ranges of temperature. This range is one *dimension* of an organism's ecological niche. Figure 2.2a shows how species of plants vary in this dimension of their niche: how they vary in the range of temperatures at which they can survive. But there are many such dimensions of a species' niche – its tolerance of various other conditions (relative humidity, pH, wind speed, water flow and so on) and its need for various resources. Clearly the real niche of a species must be *multi*dimensional.

the *n*-dimensional hypervolume

It is easy to visualize the early stages of building such a multidimensional niche. Figure 2.2b illustrates the way in which two niche dimensions (temperature and salinity) together define a two-dimensional area that is part of the niche of a sand shrimp. Three dimensions, such as temperature, pH and the availability of a particular food, may define a three-dimensional niche volume (Figure 2.2c). In fact, we consider a niche to be an *n-dimensional hypervolume*, where *n* is the number of dimensions that make up the niche. It is hard to imagine (and impossible to draw) this more realistic picture. None the less, the simplified three-dimensional version captures the idea of the ecological niche of a species. It is defined by the boundaries that limit where it can live, grow and reproduce, and it is very clearly a concept rather than a place. The concept has become a cornerstone of ecological thought.

Provided that a location is characterized by conditions within acceptable limits for a given species, and provided also that it contains all the necessary resources, then the species can, potentially, occur and persist there. Whether or not it does so depends on two further factors. First, it must be able to reach the location, and this depends in turn on its powers of colonization and the remoteness of the site. Second, its occurrence may be precluded by the action of individuals of other species that compete with it or prey on it.

fundamental and realized niches

Usually, a species has a larger ecological niche in the absence of competitors and predators than it has in their presence. In other words, there are certain combinations of conditions and resources that can allow a species to maintain a viable population, but only if it is not being adversely affected by enemies. This led Hutchinson to distinguish between the *fundamental* and the *realized* niche. The former describes the overall potentialities of a species; the latter describes the more limited spectrum of conditions and resources that allow it to persist, even in the presence of competitors and predators. Fundamental and realized niches will receive more attention in Chapter 8, when we look at interspecific competition.

The remainder of this chapter looks at some of the most important condition dimensions of species' niches, starting with temperature; the following chapter examines resources, which add further dimensions of their own.
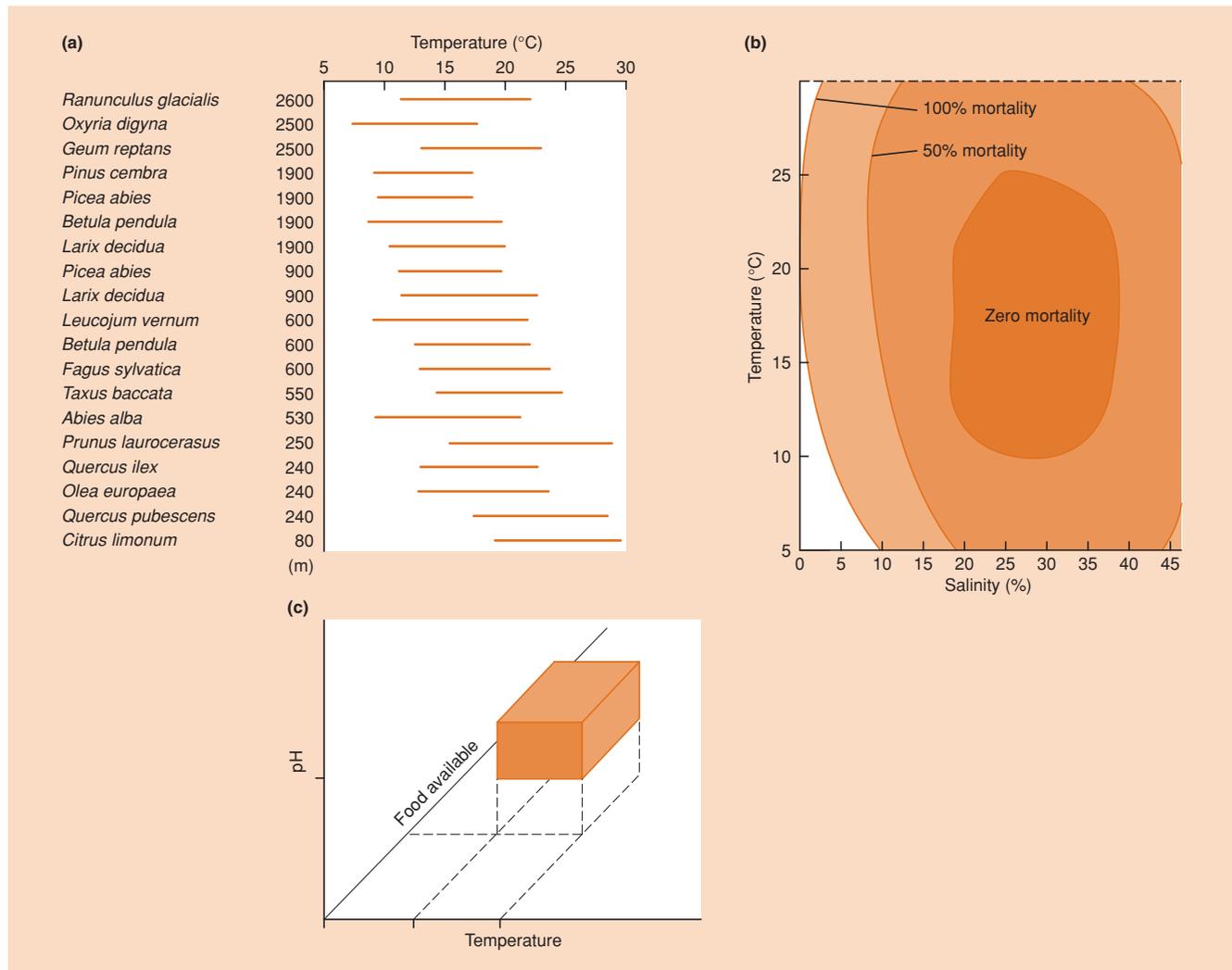
Figure 2.2 (a) A niche in one dimension. The range of temperatures at which a variety of plant species from the European Alps can achieve net photosynthesis of low intensities of radiation (70 W m^{-2}). (After Pisek *et al.*, 1973.) (b) A niche in two dimensions for the sand shrimp (*Crangon septemspinosa*) showing the fate of egg-bearing females in aerated water at a range of temperatures and salinities. (After Haefner, 1970.) (c) A diagrammatic niche in three dimensions for an aquatic organism showing a volume defined by the temperature, pH and availability of food.

2.3 Responses of individuals to temperature

2.3.1 What do we mean by 'extreme'?

It seems natural to describe certain environmental conditions as 'extreme', 'harsh', 'benign' or 'stressful'. It may seem obvious when conditions are 'extreme': the midday heat of a desert, the cold of an Antarctic winter, the salinity of the Great Salt Lake. But this only means that these conditions are extreme *for us*, given our particular physiological characteristics and tolerances.

To a cactus there is nothing extreme about the desert conditions in which cacti have evolved; nor are the icy fastnesses of Antarctica an extreme environment for penguins (Wharton, 2002). It is too easy and dangerous for the ecologist to assume that all other organisms sense the environment in the way we do. Rather, the ecologist should try to gain a worm's-eye or plant's-eye view of the environment: to see the world as others see it. Emotive words like harsh and benign, even relativities such as hot and cold, should be used by ecologists only with care.
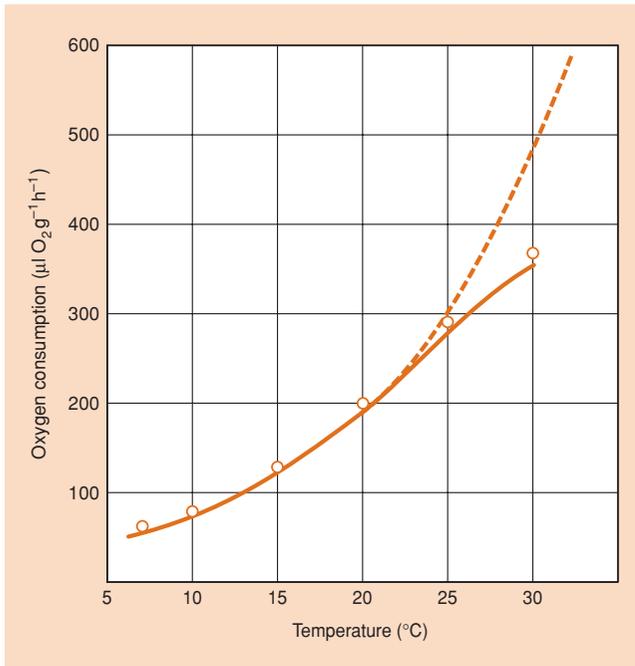
Figure 2.3 The rate of oxygen consumption of the Colorado beetle (*Leptinotarsa decemineata*), which doubles for every 10°C rise in temperature up to 20°C, but increases less fast at higher temperatures. (After Marzusch, 1952.)

2.3.2 Metabolism, growth, development and size

exponential effects of temperature on metabolic reactions

Individuals respond to temperature essentially in the manner shown in Figure 2.1a: impaired function and ultimately death at the upper and lower extremes (discussed in Sections 2.3.4 and 2.3.6), with a functional range between the extremes, within which there is an optimum. This is accounted for, in part, simply by changes in metabolic effectiveness. For each 10°C rise in temperature, for example, the rate of biological enzymatic processes often roughly doubles, and thus appears as an exponential curve on a plot of rate against temperature (Figure 2.3). The increase is brought about because high temperature increases the speed of molecular movement and speeds up chemical reactions. The factor by which a reaction changes over a 10°C range is referred to as a Q_{10}: a rough doubling means that $Q_{10} \approx 2$.

effectively linear effects on rates of growth and development

For an ecologist, however, effects on individual chemical reactions are likely to be less important than effects on rates of growth (increases in mass), on rates of development (progression through lifecycle stages) and on final body size, since, as we shall discuss much more fully in Chapter 4, these tend

to drive the core ecological activities of survival, reproduction and movement. And when we plot rates of growth and development of whole organisms against temperature, there is quite commonly an extended range over which there are, at most, only slight deviations from linearity (Figure 2.4).

When the relationship between growth or development *is* effectively

day-degree concept

linear, the temperatures experienced by an organism can be summarized in a single very useful value, the number of 'day-degrees'. For instance, Figure 2.4c shows that at 15°C (5.1°C above a development threshold of 9.9°C) the predatory mite, *Amblyseius californicus*, took 24.22 days to develop (i.e. the proportion of its total development achieved each day was 0.041 (= 1/24.22)), but it took only 8.18 days to develop at 25°C (15.1°C above the same threshold). At both temperatures, therefore, development required 123.5 day-degrees (or, more properly, 'day-degrees above threshold'), i.e. $24.22 \times 5.1 = 123.5$, and $8.18 \times 15.1 = 123.5$. This is also the requirement for development in the mite at other temperatures within the nonlethal range. Such organisms cannot be said to require a certain length of time for development. What they require is a combination of time and temperature, often referred to as 'physiological time'.

Together, the rates of growth and development determine the final size of an organism. For instance, for a given rate of growth, a faster rate of devel-

temperature–size rule

opment will lead to smaller final size. Hence, if the responses of growth and development to variations in temperature are not the same, temperature will also affect final size. In fact, development usually increases more rapidly with temperature than does growth, such that, for a very wide range of organisms, final size tends to decrease with rearing temperature: the 'temperature–size rule' (see Atkinson *et al.*, 2003). An example for single-celled protists (72 data sets from marine, brackish and freshwater habitats) is shown in Figure 2.5: for each 1°C increase in temperature, final cell volume decreased by roughly 2.5%.

These effects of temperature on growth, development and size may be of practical rather than simply scientific importance. Increasingly, ecologists are called upon to predict. We may wish to know what the consequences would be, say, of a 2°C rise in temperature resulting from global warming (see Section 2.9.2). Or we may wish to understand the role of temperature in seasonal, interannual and geographic variations in the productivity of, for example, marine ecosystems (Blackford *et al.*, 2004). We cannot afford to assume exponential relationships with temperature if they are really linear, nor to ignore the effects of changes in organism size on their role in ecological communities.

Motivated, perhaps, by this need to be able to extrapolate from the known to the unknown, and also simply by a wish to discover fundamental organizing principles governing the world

'universal temperature dependence'?

Figure 2.5 The temperature–size rule (final size decreases with increasing temperature) illustrated in protists (65 data sets combined). The horizontal scale measures temperature as a deviation from 15°C. The vertical scale measures standardized size: the difference between the cell volume observed and the cell volume at 15°C, divided by cell volume at 15°C. The slope of the mean regression line, which must pass through the point (0,0), was −0.025 (SE, 0.004); the cell volume decreased by 2.5% for every 1°C rise in rearing temperature. (After Atkinson *et al.*, 2003.)

Figure 2.4 Effectively linear relationships between rates of growth and development and temperature. (a) Growth of the protist *Strombidinopsis multiauris*. (After Montagnes *et al.*, 2003.) (b) Egg development in the beetle *Oulema duftschmidi*. (After Severini *et al.*, 2003.) (c) Egg to adult development in the mite *Amblyseius californicus*. (After Hart *et al.*, 2002.) The vertical scales in (b) and (c) represent the proportion of total development achieved in 1 day at the temperature concerned.

around us, there have been attempts to uncover universal rules of temperature dependence, for metabolism itself and for development rates, linking all organisms by scaling such dependences with aspects of body size (Gillooly *et al.*, 2001, 2002). Others have suggested that such generalizations may be *over*simplified, stressing for example that characteristics of whole organisms, like growth and development rates, are determined not only by the temperature dependence of individual chemical reactions, but also by those of the availability of resources, their rate of diffusion from the environment to metabolizing tissues, and so on (Rombough, 2003; Clarke, 2004). It may be that there is room for coexistence between broad-sweep generalizations at the grand scale and the more complex relationships at the level of individual species that these generalizations subsume.

2.3.3   Ectotherms and endotherms

Many organisms have a body temperature that differs little, if at all, from their environment. A parasitic worm in the gut of a mammal, a fungal mycelium in the soil and a sponge in the sea acquire the temperature of the medium in which they live. Terrestrial organisms, exposed to the sun and the air, are different because they may acquire heat directly by absorbing solar radiation or be cooled by the latent heat of evaporation of water (typical
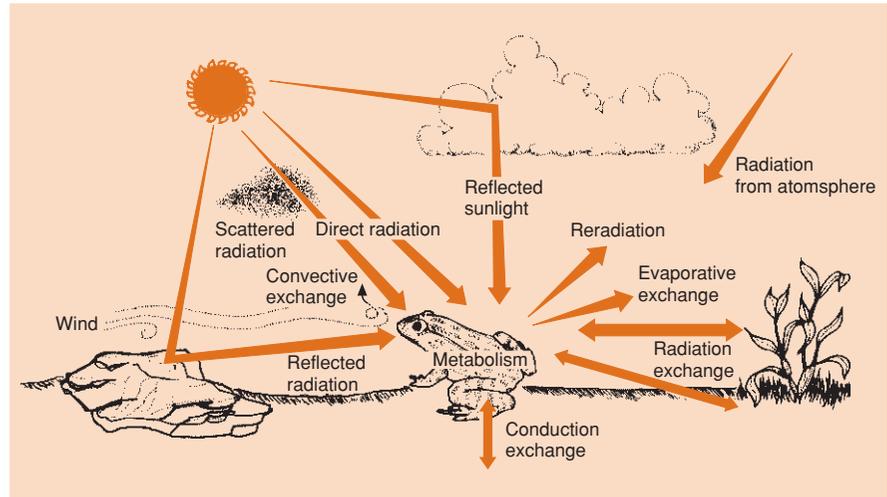
Figure 2.6 Schematic diagram of the avenues of heat exchange between an ectotherm and a variety of physical aspects of its environment. (After Tracy, 1976; from Hainsworth, 1981.)

pathways of heat exchange are shown in Figure 2.6). Various fixed properties may ensure that body temperatures are higher (or lower) than the ambient temperatures. For example, the reflective, shiny or silvery leaves of many desert plants reflect radiation that might otherwise heat the leaves. Organisms that can move have further control over their body temperature because they can seek out warmer or cooler environments, as when a lizard chooses to warm itself by basking on a hot sunlit rock or escapes from the heat by finding shade.

Amongst insects there are examples of body temperatures raised by controlled muscular work, as when bumblebees raise their body temperature by shivering their flight muscles. Social insects such as bees and termites may combine to control the temperature of their colonies and regulate them with remarkable thermostatic precision. Even some plants (e.g. *Philodendron*) use metabolic heat to maintain a relatively constant temperature in their flowers; and, of course, birds and mammals use metabolic heat almost all of the time to maintain an almost perfectly constant body temperature.

An important distinction, therefore, is between *endotherms* that regulate their temperature by the production of heat within their own bodies, and *ectotherms* that rely on external sources of heat. But this distinction is not entirely clear cut. As we have noted, apart from birds and mammals, there are also other taxa that use heat generated in their own bodies to regulate body temperature, but only for limited periods; and there are some birds and mammals that relax or suspend their endothermic abilities at the most extreme temperatures. In particular, many endothermic animals escape from some of the costs of endothermy by hibernating during the coldest seasons: at these times they behave almost like ectotherms.

Birds and mammals usually maintain a constant body temperature between

35 and 40°C, and they therefore tend to lose heat in most environments; but this loss is moderated by insulation in the form of fur, feathers and fat, and by controlling blood flow near the skin surface. When it is necessary to increase the rate of heat loss, this too can be achieved by the control of surface blood flow and by a number of other mechanisms shared with ectotherms like panting and the simple choice of an appropriate habitat. Together, all these mechanisms and properties give endotherms a powerful (but not perfect) capability for regulating their body temperature, and the benefit they obtain from this is a constancy of near-optimal performance. But the price they pay is a large expenditure of energy (Figure 2.7), and thus a correspondingly large requirement for food to provide that energy. Over a certain temperature range (the thermoneutral zone) an endotherm consumes energy at a basal rate. But at environmental temperatures further and further above or below that zone, the endotherm consumes more and more energy in maintaining a constant body temperature. Even in the thermoneutral zone, though, an endotherm typically consumes energy many times more rapidly than an ectotherm of comparable size.

The responses of endotherms and ectotherms to changing temperatures, then, are not so different as they may at first appear to be. Both are at risk of being killed by even short exposures to very low temperatures and by more prolonged exposure to moderately low temperatures. Both have an optimal environmental temperature and upper and lower lethal limits. There are also costs to both when they live at temperatures that are not optimal. For the ectotherm these may be slower growth and reproduction, slow movement, failure to escape predators and a sluggish rate of search for food. But for the endotherm, the maintenance of body temperature costs energy that might have been used to catch more prey, produce and nurture more offspring or escape more predators. There are also costs of insulation (e.g. blubber in whales, fur in mammals) and even costs of changing the insulation between
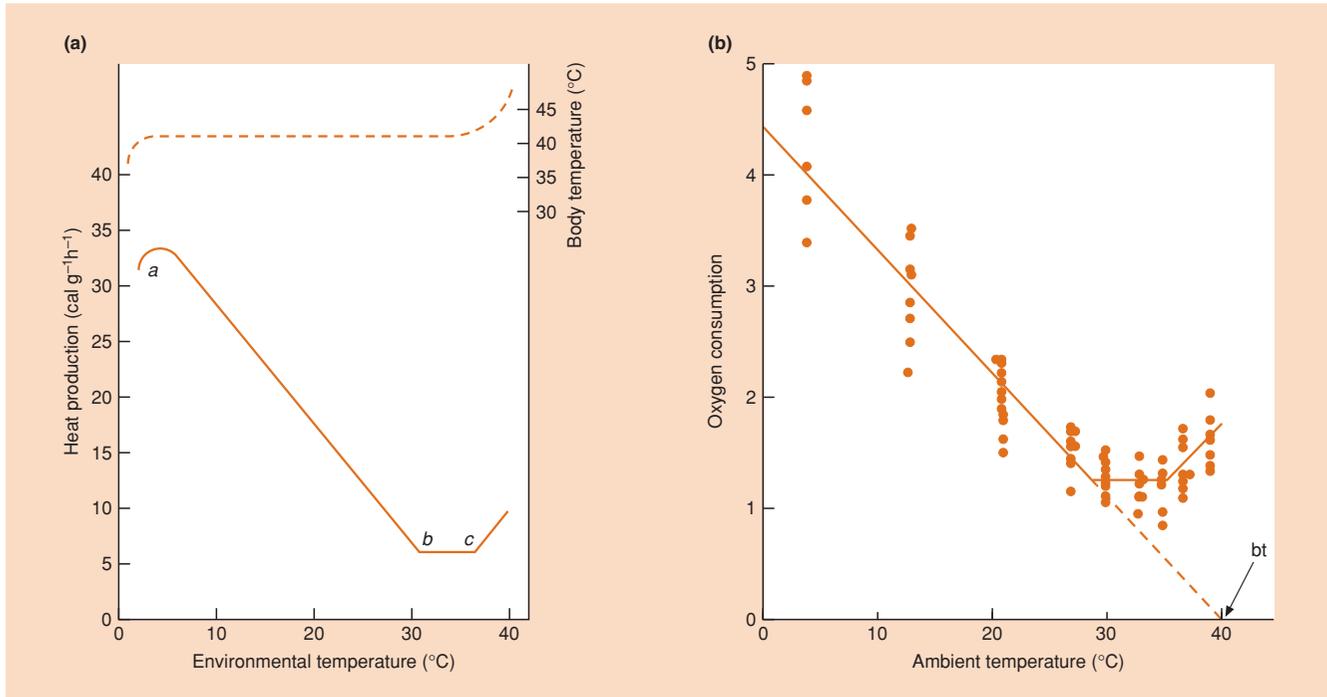
Figure 2.7 (a) Thermostatic heat production by an endotherm is constant in the thermoneutral zone, i.e. between *b*, the lower critical temperature, and *c*, the upper critical temperature. Heat production rises, but body temperature remains constant, as environmental temperature declines below *b*, until heat production reaches a maximum possible rate at a low environmental temperature. Below *a*, heat production and body temperature both fall. Above *c*, metabolic rate, heat production and body temperature all rise. Hence, body temperature is constant at environmental temperatures between *a* and *c*. (After Hainsworth, 1981.) (b) The effect of environmental temperature on the metabolic rate (rate of oxygen consumption) of the eastern chipmunk (*Tamias striatus*). bt, body temperature. Note that at temperatures between 0 and 30°C oxygen consumption decreases approximately linearly as the temperature increases. Above 30°C a further increase in temperature has little effect until near the animal's body temperature when oxygen consumption increases again. (After Neumann, 1967; Nedgergaard & Cannon, 1990.)

seasons. Temperatures only a few degrees higher than the metabolic optimum are liable to be lethal to endotherms as well as ectotherms (see Section 2.3.6).

ectotherms and endotherms coexist: both strategies 'work'

It is tempting to think of ectotherms as 'primitive' and endotherms as having gained 'advanced' control over their environment, but it is difficult to justify this view. Most environments on earth are inhabited by mixed communities of endothermic and ectothermic animals. This includes some of the hottest – e.g. desert rodents and lizards – and some of the coldest – penguins and whales together with fish and krill at the edge of the Antarctic ice sheet. Rather, the contrast, crudely, is between the high cost–high benefit strategy of endotherms and the low cost–low benefit strategy of ectotherms. But their coexistence tells us that both strategies, in their own ways, can 'work'.

2.3.4 Life at low temperatures

The greater part of our planet is below 5°C: 'cold is the fiercest and most widespread enemy of life on earth' (Franks *et al.*, 1990). More than 70% of the planet is covered with seawater: mostly deep ocean with a remarkably constant temperature of about 2°C. If we include the polar ice caps, more than 80% of earth's biosphere is permanently cold.

By definition, all temperatures below the optimum are harmful, but there is usually a wide range of such temperatures that cause no physical damage and over which any effects are fully reversible. There are, however, two quite distinct types of damage at low temperatures that can be lethal, either to tissues or to whole organisms: chilling and freezing. Many organisms are damaged by exposure to temperatures that are low but above freezing point – so-called

chilling injury

'chilling injury'. The fruits of the banana blacken and rot after exposure to chilling temperatures and many tropical rainforest species are sensitive to chilling. The nature of the injury is obscure, although it seems to be associated with the breakdown of membrane permeability and the leakage of specific ions such as calcium (Minorsky, 1985).

Temperatures below 0°C can have lethal physical and chemical consequences even though ice may not be formed. Water may 'supercool' to temperatures at least as low as −40°C, remaining in an unstable liquid form in which its physical properties change in ways that are bound to be biologically significant: its viscosity increases, its diffusion rate decreases and its degree of ionization of water decreases. In fact, ice seldom forms in an organism until the temperature has fallen several degrees below 0°C. Body fluids remain in a supercooled state until ice forms suddenly around particles that act as nuclei. The concentration of solutes in the remaining liquid phase rises as a consequence. It is very rare for ice to form within cells and it is then inevitably lethal, but the freezing of extracellular water is one of the factors that prevents ice forming within the cells themselves (Wharton, 2002), since water is withdrawn from the cell, and solutes in the cytoplasm (and vacuoles) become more concentrated. The effects of freezing are therefore mainly osmoregulatory: the water balance of the cells is upset and cell membranes are destabilized. The effects are essentially similar to those of drought and salinity.

freeze-avoidance and freeze-tolerance

Organisms have at least two different metabolic strategies that allow survival through the low temperatures of winter. A 'freeze-avoiding' strategy uses low-molecular-weight polyhydric alcohols (polyols, such as glycerol) that depress both the freezing and the supercooling point and also 'thermal hysteresis' proteins that prevent ice nuclei from forming (Figure 2.8a, b). A contrasting 'freeze-tolerant' strategy, which also involves the formation of polyols, encourages the formation of extracellular ice, but protects the cell membranes from damage when water is withdrawn from the cells (Storey, 1990). The tolerances of organisms to low temperatures are not fixed but are preconditioned by the experience of temperatures in their recent past. This process is called *acclimation* when it occurs in the laboratory and *acclimatization* when it occurs naturally. Acclimatization may start as the weather becomes colder in the fall, stimulating the conversion of almost the entire glycogen reserve of animals into polyols (Figure 2.8c), but this can be an energetically costly affair: about 16% of the carbohydrate reserve may be consumed in the conversion of the glycogen reserves to polyols.

acclimation and acclimatization

The exposure of an individual for several days to a relatively low temperature can shift its whole temperature response downwards along the temperature scale. Similarly, exposure to a high temperature can shift the temperature response upwards. Antarctic springtails (tiny arthropods), for instance, when taken from 'summer' temperatures in the field (around 5°C in the Antarctic) and subjected to a range of acclimation temperatures, responded to temperatures in the range +2°C to −2°C (indicative of winter) by showing a marked drop in the temperature at which they froze (Figure 2.9); but at lower acclimation temperatures still (−5°C, −7°C), they showed no such drop because the temperatures were themselves too low for the physiological processes required to make the acclimation response.

Acclimatization aside, individuals commonly vary in their temperature response depending on the stage of development they have reached. Probably the most extreme form of this is when an organism has a dormant stage in its life cycle. Dormant stages are typically dehydrated, metabolically slow and tolerant of extremes of temperature.

2.3.5 Genetic variation and the evolution of cold tolerance

Even within species there are often differences in temperature response between populations from different locations, and these differences have frequently been found to be the result of genetic differences rather than being attributable solely to acclimatization. Powerful evidence that cold tolerance varies between geographic races of a species comes from a study of the cactus, *Opuntia fragilis*. Cacti are generally species of hot dry habitats, but *O. fragilis* extends as far north as 56°N and at one site the lowest extreme minimum temperature recorded was −49.4°C. Twenty populations were sampled from diverse localities in northern USA and Canada, and were tested for freezing tolerance and ability to acclimate to cold. Individuals from the most freeze-tolerant population (from Manitoba) tolerated −49°C in laboratory tests and acclimated by 19.9°C, whereas plants from a population in the more equable climate of Hornby Island, British Columbia, tolerated only −19°C and acclimated by only 12.1°C (Loik & Nobel, 1993).

There are also striking cases where the geographic range of a crop species has been extended into colder regions by plant breeders. Programs of deliberate selection applied to corn (*Zea mays*) have expanded the area of the USA over which the crop can be profitably grown. From the 1920s to the 1940s, the production of corn in Iowa and Illinois increased by around 24%, whereas in the colder state of Wisconsin it increased by 54%.

If deliberate selection can change the tolerance and distribution of a domesticated plant we should expect natural selection to have done the same thing in nature. To test this, the plant *Umbilicus rupestris*, which lives in mild maritime areas of Great Britain, was deliberately grown outside its normal range (Woodward, 1990). A population of plants and seeds was taken from a donor population in the mild-wintered habitat of Cardiff in the west and introduced in a cooler environment at an altitude of
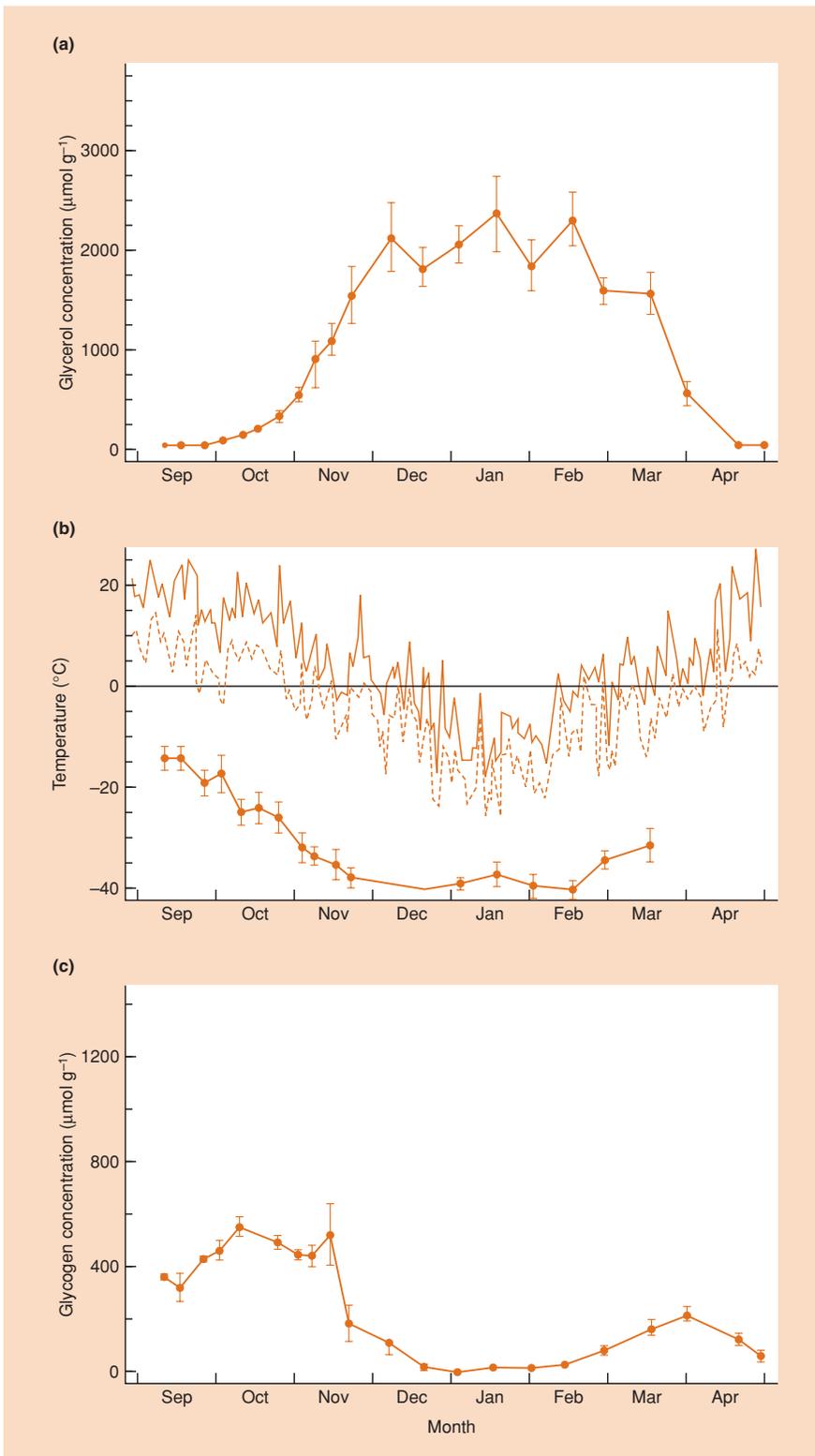
Figure 2.8 (a) Changes in the glycerol concentration per gram wet mass of the freeze-avoiding larvae of the goldenrod gall moth, *Epiblema scudderiana*. (b) The daily temperature maxima and minima (above) and whole larvae supercooling points (below) over the same period. (c) Changes in glycogen concentration over the same period. (After Rickards *et al.*, 1987.)

Figure 2.9 Acclimation to low temperatures. Samples of the Antarctic springtail *Cryptopygus antarcticus* were taken from field sites in the summer (*c.* 5°C) on a number of days and their supercooling point (at which they froze) was determined either immediately (●) or after a period of acclimation (●) at the temperatures shown. The supercooling points of the controls themselves varied because of temperature variations from day to day, but acclimation at temperatures in the range +2 to −2°C (indicative of winter) led to a drop in the supercooling point, whereas no such drop was observed at higher temperatures (indicative of summer) or lower temperatures (too low for a physiological acclimation response). Bars are standard errors. (After Worland & Convey, 2001.)
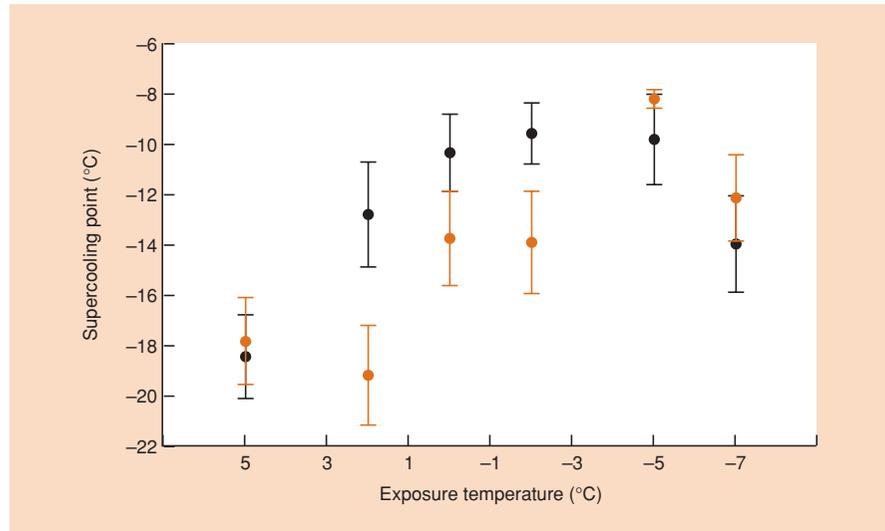
157 m in Sussex in the south. After 8 years, the temperature response of seeds from the donor and the introduced populations had diverged quite strikingly (Figure 2.10a), and subfreezing temperatures that kill in Cardiff (−12°C) were then tolerated

by 50% of the Sussex population (Figure 2.10b). This suggests that past climatic changes, for example ice ages, will have changed the temperature tolerance of species as well as forcing their migration.
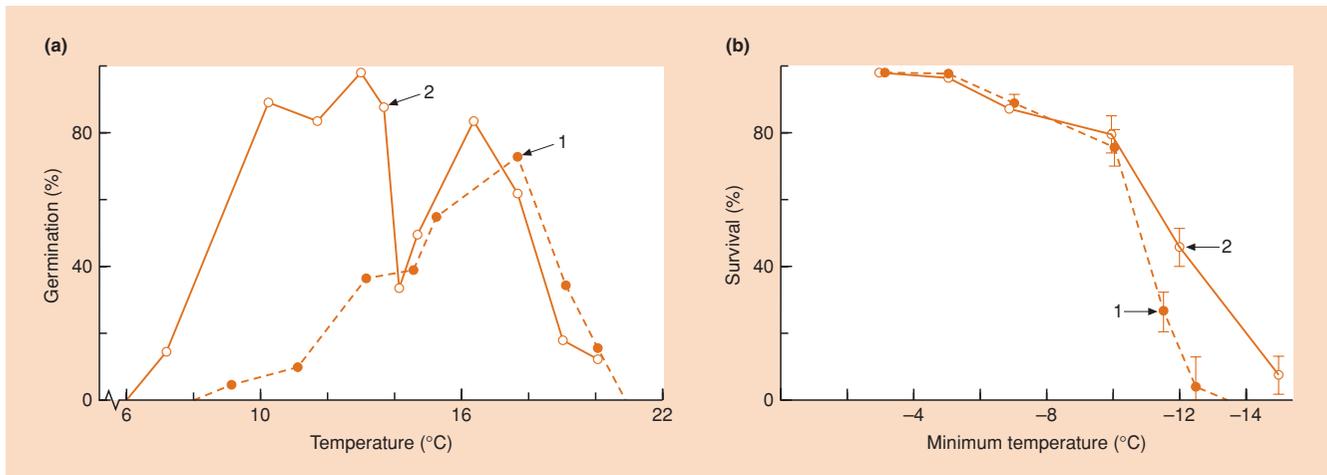
Figure 2.10 Changes in the behavior of populations of the plant *Umbilicus rupestris*, established for a period of 8 years in a cool environment in Sussex from a donor population in a mild-wintered area in South Wales (Cardiff, UK). (a) Temperature responses of seed germination: (1) responses of samples from the donor population (Cardiff) in 1978, and (2) responses from the Sussex population in 1987. (b) The low-temperature survival of the donor population at Cardiff, 1978 (1) and of the established population in Sussex, 1987 (2). (After Woodward, 1990.)

2.3.6 Life at high temperatures

Perhaps the most important thing about dangerously high temperatures is that, for a given organism, they usually lie only a few degrees above the metabolic optimum. This is largely an unavoidable consequence of the physicochemical properties of most enzymes (Wharton, 2002). High temperatures may be dangerous because they lead to the inactivation or even the denaturation of enzymes, but they may also have damaging indirect effects by leading to dehydration. All terrestrial organisms need to conserve water, and at high temperatures the rate of water loss by evaporation can be lethal, but they are caught between the devil and the deep blue sea because evaporation is an important means of reducing body temperature. If surfaces are protected from evaporation (e.g. by closing stomata in plants or spiracles in insects) the organisms may be killed by too high a body temperature, but if their surfaces are not protected they may die of desiccation.

high temperature and water loss

Death Valley, California, in the summer, is probably the hottest place on earth in which higher plants make active growth. Air temperatures during the daytime may approach 50°C and soil surface temperatures may be very much higher. The perennial plant, desert honeysweet (*Tidestromia oblongifolia*), grows vigorously in such an environment despite the fact that its leaves are killed if they reach the same temperature as the air. Very rapid transpiration keeps the temperature of the leaves at 40–45°C, and in this range they are capable of extremely rapid photosynthesis (Berry & Björkman, 1980).

Most of the plant species that live in very hot environments suffer severe shortage of water and are therefore unable to use the latent heat of evaporation of water to keep leaf temperatures down. This is especially the case in desert succulents in which water loss is minimized by a low surface to volume ratio and a low frequency of stomata. In such plants the risk of overheating may be reduced by spines (which shade the surface of a cactus) or hairs or waxes (which reflect a high proportion of the incident radiation). Nevertheless, such species experience and tolerate temperatures in their tissues of more than 60°C when the air temperature is above 40°C (Smith *et al.*, 1984).

fire

Fires are responsible for the highest temperatures that organisms face on earth and, before the fire-raising activities of humans, were caused mainly by lightning strikes. The recurrent risk of fire has shaped the species composition of arid and semiarid woodlands in many parts of the world. All plants are damaged by burning but it is the remarkable powers of regrowth from protected meristems on shoots and seeds that allow a specialized subset of species to recover from damage and form characteristic fire floras (see, for example, Hodgkinson, 1992).

Decomposing organic matter in heaps of farmyard manure, compost heaps and damp hay may reach very high temperatures. Stacks of damp hay are heated to temperatures of 50–60°C by the metabolism of fungi such as *Aspergillus fumigatus*, carried further to approximately 65°C by other thermophilic fungi such as *Mucor pusillus* and then a little further by bacteria and actinomycetes. Biological activity stops well short of 100°C but autocombustible products are formed that cause further heating, drive off water and may even result in fire. Another hot environment is that of natural hot springs and in these the microbe *Thermus aquaticus* grows at temperatures of 67°C and tolerates temperatures up to 79°C. This organism has also been isolated from domestic hot water systems. Many (perhaps all) of the extremely thermophilic species are prokaryotes. In environments with very high temperatures the communities contain few species. In general, animals and plants are the most sensitive to heat followed by fungi, and in turn by bacteria, actinomycetes and archaebacteria. This is essentially the same order as is found in response to many other extreme conditions, such as low temperature, salinity, metal toxicity and desiccation.

thermal vents and other hot environments

An ecologically very remarkable hot environment was first described only towards the end of the last century. In 1979, a deep oceanic site was discovered in the eastern Pacific at which fluids at high temperatures ('smokers') were vented from the sea floor forming thin-walled 'chimneys' of mineral materials. Since that time many more vent sites have been discovered at mid-ocean crests in both the Atlantic and Pacific Oceans. They lie 2000–4000 m below sea level at pressures of 200–400 bars (20–40 MPa). The boiling point of water is raised to 370°C at 200 bars and to 404°C at 400 bars. The superheated fluid emerges from the chimneys at temperatures as high as 350°C, and as it cools to the temperature of seawater at about 2°C it provides a continuum of environments at intermediate temperatures.

Environments at such extreme pressures and temperatures are obviously extraordinarily difficult to study *in situ* and in most respects impossible to maintain in the laboratory. Some thermophilic bacteria collected from vents have been cultured successfully at 100°C at only slightly above normal barometric pressures (Jannasch & Mottl, 1985), but there is much circumstantial evidence that some microbial activity occurs at much higher temperatures and may form the energy resource for the warm water communities outside the vents. For example, particulate DNA has been found in samples taken from within the 'smokers' at concentrations that point to intact bacteria being present at temperatures very much higher than those conventionally thought to place limits on life (Baross & Deming, 1995).

There is a rich eukaryotic fauna in the local neighborhood of vents that is quite atypical of the deep oceans in general. At one vent in Middle Valley, Northeast Pacific, surveyed photographically and by video, at least 55 taxa were documented of which 15 were new or probably new species (Juniper *et al.*, 1992). There can be few environments in which so complex and specialized a community depends on so localized a special condition. The

closest known vents with similar conditions are 2500 km distant. Such communities add a further list to the planet's record of species richness. They present tantalizing problems in evolution and daunting problems for the technology needed to observe, record and study them.

2.3.7 Temperature as a stimulus

We have seen that temperature as a condition affects the rate at which organisms develop. It may also act as a stimulus, determining whether or not the organism starts its development at all. For instance, for many species of temperate, arctic and alpine herbs, a period of chilling or freezing (or even of alternating high and low temperatures) is necessary before germination will occur. A cold experience (physiological evidence that winter has passed) is required before the plant can start on its cycle of growth and development. Temperature may also interact with other stimuli (e.g. photoperiod) to break dormancy and so time the onset of growth. The seeds of the birch (*Betula pubescens*) require a photoperiodic stimulus (i.e. experience of a particular regime of day length) before they will germinate, but if the seed has been chilled it starts growth without a light stimulus.

2.4 Correlations between temperature and the distribution of plants and animals

2.4.1 Spatial and temporal variations in temperature

Variations in temperature on and within the surface of the earth have a variety of causes: latitudinal, altitudinal, continental, seasonal, diurnal and microclimatic effects and, in soil and water, the effects of depth.

Latitudinal and seasonal variations cannot really be separated. The angle at which the earth is tilted relative to the sun changes with the seasons, and this drives some of the main temperature differentials on the earth's surface. Superimposed on these broad geographic trends are the influences of altitude and 'continentality'. There is a drop of 1°C for every 100 m increase in altitude in dry air, and a drop of 0.6°C in moist air. This is the result of the 'adiabatic' expansion of air as atmospheric pressure falls with increasing altitude. The effects of continentality are largely attributable to different rates of heating and cooling of the land and the sea. The land surface reflects less heat than the water, so the surface warms more quickly, but it also loses heat more quickly. The sea therefore has a moderating, 'maritime' effect on the temperatures of coastal regions and especially islands; both daily and seasonal variations in temperature are far less marked than at more inland, continental locations at the same latitude. Moreover, there are comparable effects within land masses: dry, bare areas like deserts suffer greater daily and seasonal extremes of temperature

than do wetter areas like forests. Thus, global maps of temperature zones hide a great deal of local variation.

It is much less widely appreciated that on a smaller scale still there can be a great deal of microclimatic variation. For example, the sinking of dense, cold air into the bottom of a valley at night can make it as much as 30°C colder than the side of the valley only 100 m higher; the winter sun, shining on a cold day, can heat the south-facing side of a tree (and the habitable cracks and crevices within it) to as high as 30°C; and the air temperature in a patch of vegetation can vary by 10°C over a vertical distance of 2.6 m from the soil surface to the top of the canopy (Geiger, 1955). Hence, we need not confine our attention to global or geographic patterns when seeking evidence for the influence of temperature on the distribution and abundance of organisms.

Long-term temporal variations in temperature, such as those associated with the ice ages, were discussed in the previous chapter. Between these, however, and the very obvious daily and seasonal changes that we are all aware of, a number of medium-term patterns have become increasingly apparent. Notable amongst these are the El Niño-Southern Oscillation (ENSO) and the North Atlantic Oscillation (NAO) (Figure 2.11) (see Stenseth *et al.*, 2003). The ENSO originates in the tropical Pacific Ocean off the coast of South America and is an alternation (Figure 2.11a) between a warm (El Niño) and a cold (La Niña) state of the water there, though it affects temperature, and the climate generally, in terrestrial and marine environments throughout the whole Pacific basin (Figure 2.11b; for color, see Plate 2.1, between pp. 000 and 000) and beyond. The NAO refers to a north–south alternation in atmospheric mass between the subtropical Atlantic and the Arctic (Figure 2.11c) and again affects climate in general rather than just temperature (Figure 2.11d; for color, see Plate 2.2, between pp. 000 and 000). Positive index values (Figure 2.11c) are associated, for example, with relatively warm conditions in North America and Europe and relatively cool conditions in North Africa and the Middle East. An example of the effect of NAO variation on species abundance, that of cod, *Gadus morhua*, in the Barents Sea, is shown in Figure 2.12.

2.4.2 Typical temperatures and distributions

There are very many examples of plant and animal distributions that are strikingly correlated with some aspect of environmental temperature even at gross taxonomic and systematic levels (Figure 2.13). At a finer scale, the distributions of many species closely match maps of some aspect of temperature. For example, the northern limit of the distribution of wild madder plants (*Rubia peregrina*) is closely correlated with the position of the January 4.5°C

microclimatic variation

ENSO and NAO

isotherms
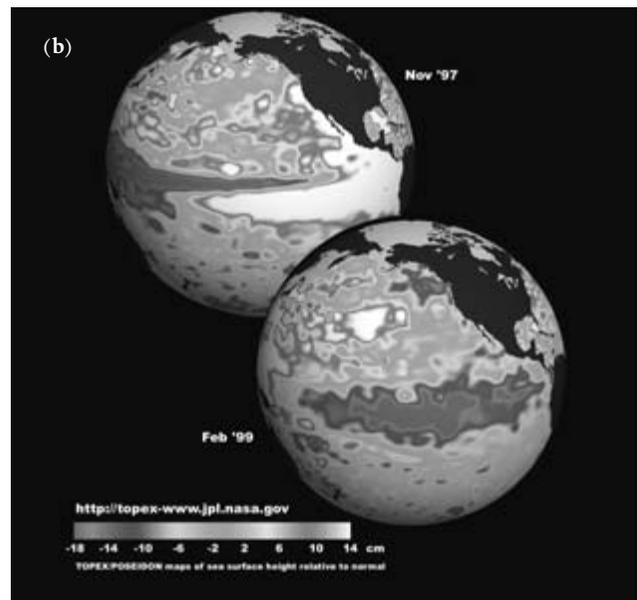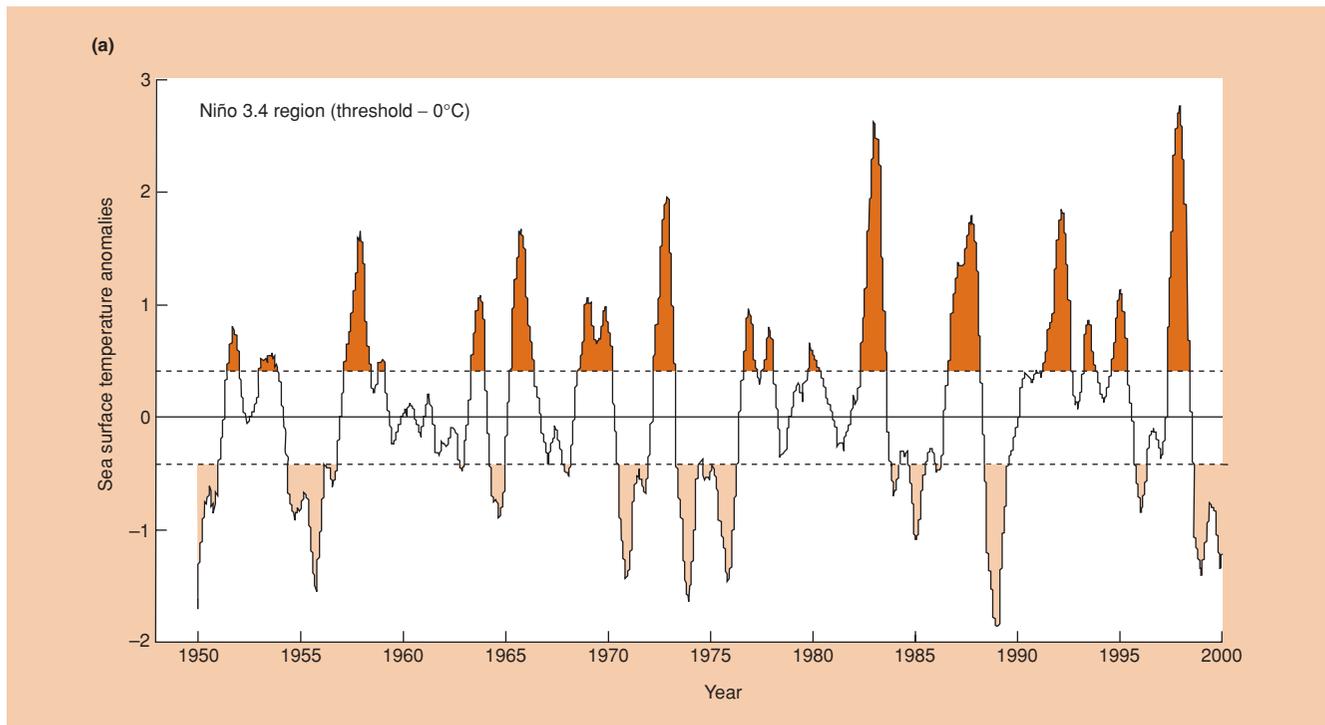
Figure 2.11   (a) The El Niño–Southern Oscillation (ENSO) from 1950 to 2000 as measured by sea surface temperature anomalies (differences from the mean) in the equatorial mid-Pacific. The El Niño events (> 0.4°C above the mean) are shown in dark color, and the La Niña events (> 0.4°C below the mean) are shown in pale color. (Image from http://www.cgd.ucar.edu/cas/catalog/ climind/Nino_3_3.4_indices.html.) (b) Maps of examples of El Niño (November 1997) and La Niña (February 1999) events in terms of sea height above average levels. Warmer seas are higher; for example, a sea height 15–20 cm below average equates to a temperature anomaly of approximately 2–3°C. (Image from http://topex-www.jpl.nasa.gov/science/images/el-nino-la-nina.jpg.) (For color, see Plate 2.1, between pp. 000 and 000.)
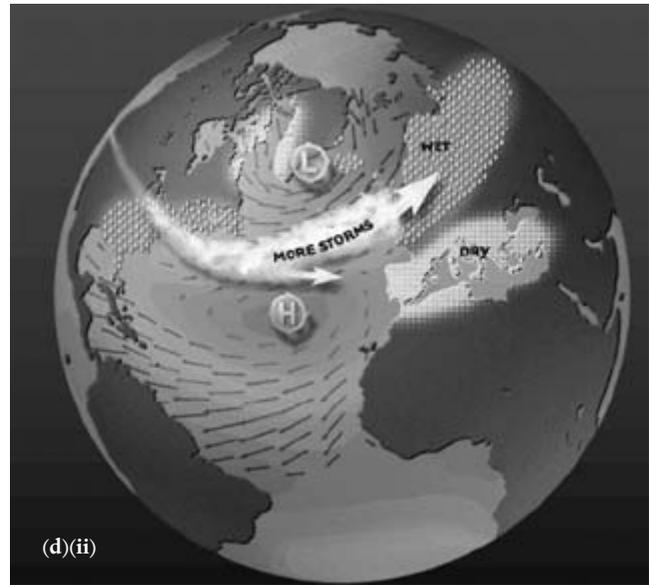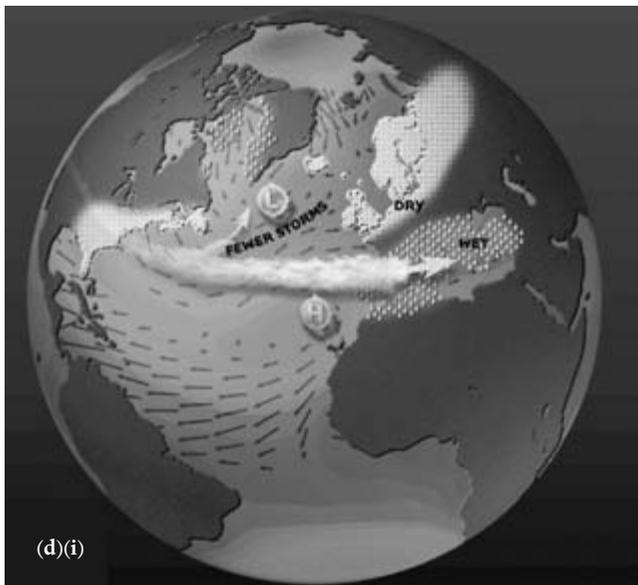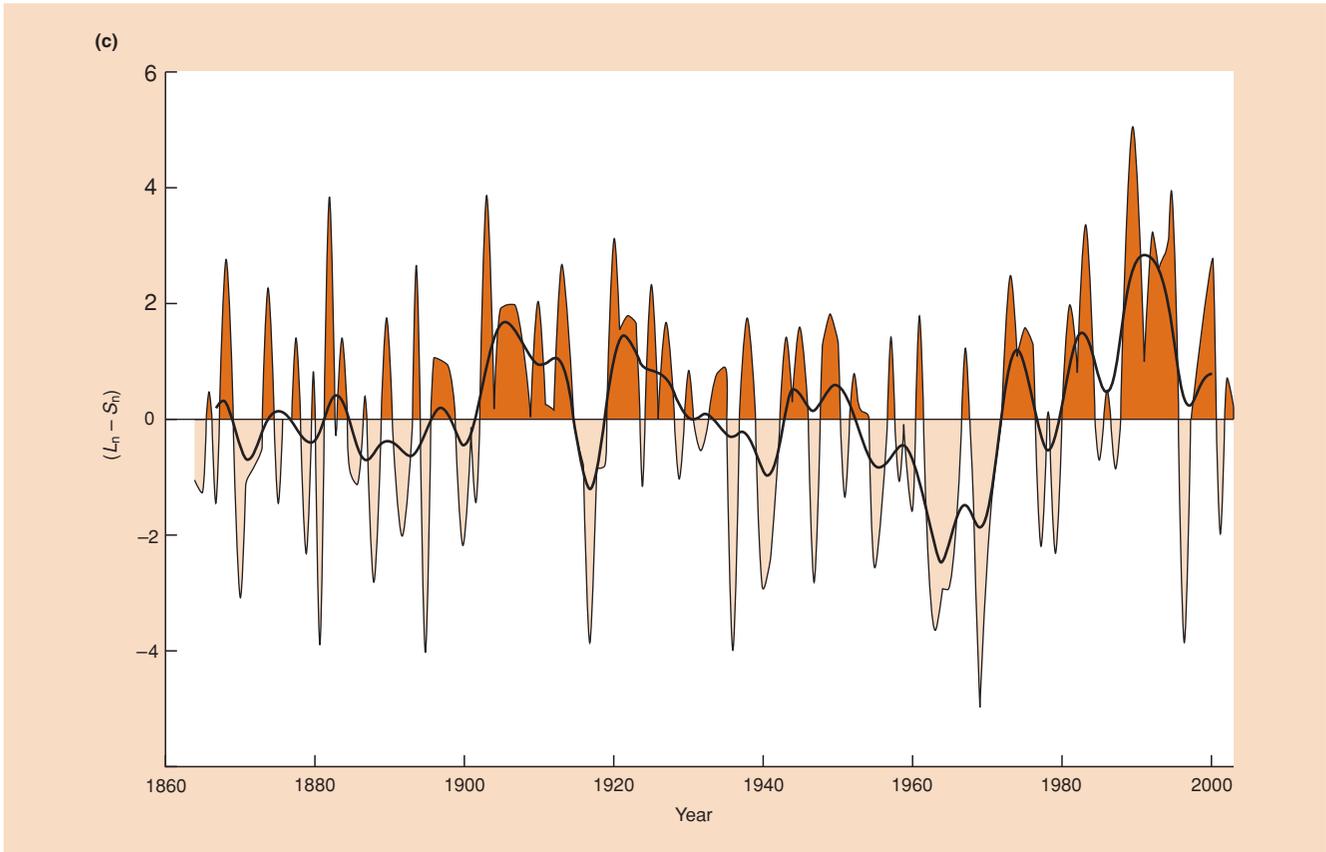
Figure 2.11 (*continued*) (c) The North Atlantic Oscillation (NAO) from 1864 to 2003 as measured by the normalized sea-level pressure difference ($L_n - S_n$) between Lisbon, Portugal and Reykjavik, Iceland. (Image from http://www.cgd.ucar.edu/~jhurrell/ nao.stat.winter.html#winter.) (d) Typical winter conditions when the NAO index is positive or negative. Conditions that are more than usually warm, cold, dry or wet are indicated. (Image from http://www.ldeo.columbia.edu/NAO/.) (For color, see Plate 2.2, between pp. 000 and 000.)
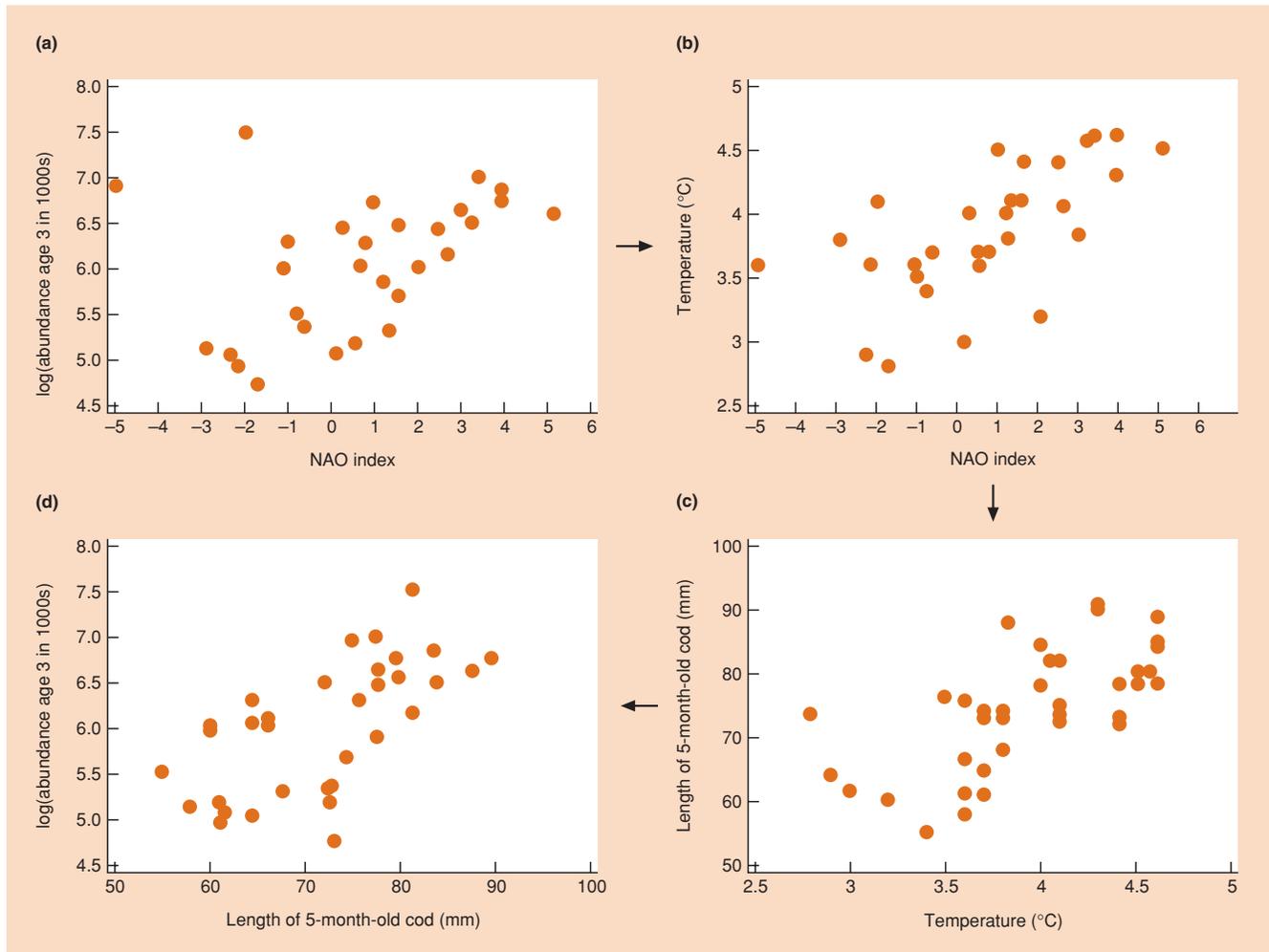
Figure 2.12   (a) The abundance of 3-year-old cod, *Gadus morhua*, in the Barents Sea is positively correlated with the value of the North Atlantic Oscillation (NAO) index for that year. The mechanism underlying this correlation is suggested in (b–d). (b) Annual mean temperature increases with the NAO index. (c) The length of 5-month-old cod increases with annual mean temperature. (d) The abundance of cod at age 3 increases with their length at 5 months. (After Ottersen *et al.*, 2001.)

isotherm (Figure 2.14a; an isotherm is a line on a map joining places that experience the same temperature – in this case a January mean of 4.5°C). However, we need to be very careful how we interpret such relationships: they can be extremely valuable in predicting where we might and might not find a particular species; they may suggest that some feature related to temperature is important in the life of the organisms; but they do not prove that temperature *causes* the limits to a species' distribution. The literature relevant to this and many other correlations between temperature and distribution patterns is reviewed by Hengeveld (1990), who also describes a more subtle graphical procedure. The minimum temperature of the coldest month and the maximum temperature of the hottest month are estimated for many places within and outside the range of a species. Each location is then plotted on a graph of maximum against minimum temperature, and a line is drawn that optimally discriminates between the presence and absence records (Figure 2.14b). This line is then used to define the geographic margin of the species distributions (Figure 2.14c). This may have powerful predictive value, but it still tells us nothing about the underlying forces that cause the distribution patterns.

One reason why we need to be cautious about reading too much into correlations of species distributions with maps of temperature is that the temperatures measured for constructing isotherms for a map are only rarely those that the organisms experience. In nature an organism may choose to lie in the sun or hide
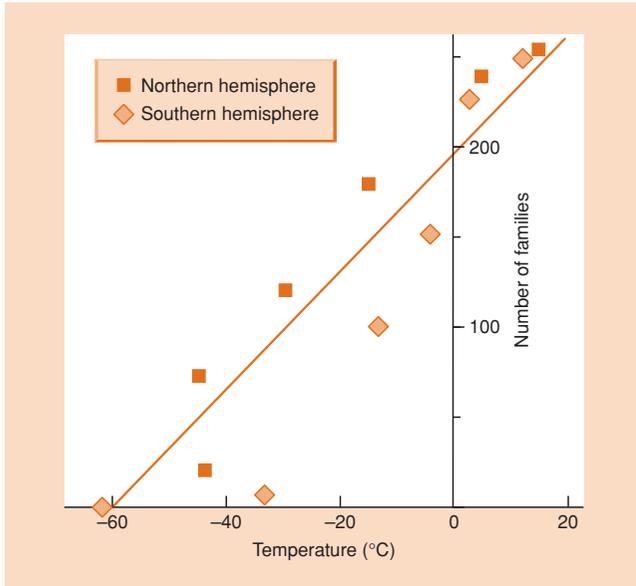
Figure 2.13 The relationship between absolute minimum temperature and the number of families of flowering plants in the northern and southern hemispheres. (After Woodward, 1987, who also discusses the limitations to this sort of analysis and how the history of continental isolation may account for the odd difference between northern and southern hemispheres.)

in the shade and, even in a single day, may experience a baking midday sun and a freezing night. Moreover, temperature varies from place to place on a far finer scale than will usually concern a geographer, but it is the conditions in these 'microclimates' that will be crucial in determining what is habitable for a particular species. For example, the prostrate shrub *Dryas octopetala* is restricted to altitudes exceeding 650 m in North Wales, UK, where it is close to its southern limit. But to the north, in Sutherland in Scotland, where it is generally colder, it is found right down to sea level.

2.4.3   Distributions and extreme conditions

For many species, distributions are accounted for not so much by average temperatures as by occasional extremes, especially occasional lethal temperatures that preclude its existence. For instance, injury by frost is probably the single most important factor limiting plant distribution. To take one example: the saguaro cactus (*Carnegiea gigantea*) is liable to be killed when temperatures remain below freezing for 36 h, but if there is a daily thaw it is under no threat. In Arizona, the northern and eastern edges of the cactus' distribution correspond to a line joining places where on occasional days it fails to thaw. Thus, the saguaro is absent where there are occasionally lethal conditions – an individual need only be killed once.

Similarly, there is scarcely any crop that is grown on a large commercial scale in the climatic conditions of its wild ancestors, and it is well known that crop failures are often caused by extreme events, especially frosts and drought. For instance, the climatic limit to the geographic range for the production of coffee (*Coffea arabica* and *C. robusta*) is defined by the 13°C isotherm for the coldest month of the year. Much of the world's crop is produced in the highland microclimates of the São Paulo and Paraná districts of

you only die once
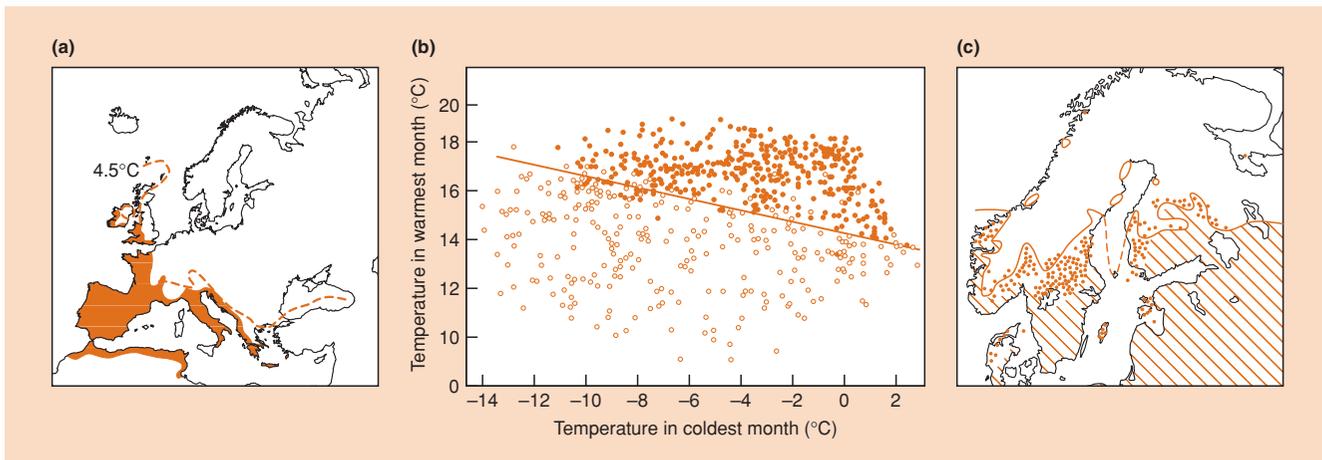
Figure 2.14   (a) The northern limit of the distribution of the wild madder (*Rubia peregrina*) is closely correlated with the position of the January 4.5°C isotherm. (After Cox *et al.*, 1976.) (b) A plot of places within the range of *Tilia cordat* (●), and outside its range (○) in the graphic space defined by the minimum temperature of the coldest month and the maximum temperature of the warmest month. (c) Margin of the geographic range of *T. cordata* in northern Europe defined by the straight line in (b). ((b, c) after Hintikka, 1963; from Hengeveld, 1990.)
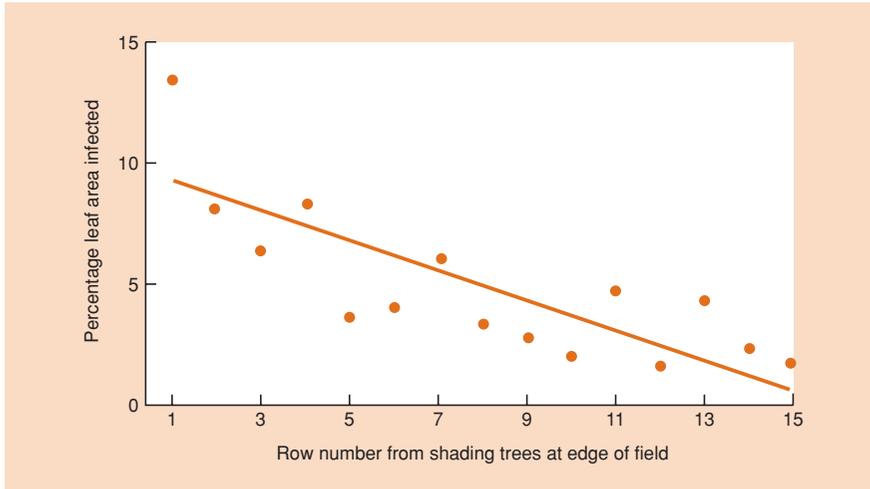
Figure 2.15   The incidence of southern corn leaf blight (*Helminthosporium maydis*) on corn growing in rows at various distances from trees that shaded them. Wind-borne fungal diseases were responsible for most of this mortality (Harper, 1955). (From Lukens & Mullany, 1972.)

Brazil. Here, the average minimum temperature is 20°C, but occasionally cold winds and just a few hours of temperature close to freezing are sufficient to kill or severely damage the trees (and play havoc with world coffee prices).

2.4.4   Distributions and the interaction of temperature with other factors

Although organisms respond to each condition in their environment, the effects of conditions may be determined largely by the responses of other community members. Temperature does not act on just one species: it also acts on its competitors, prey, parasites and so on. This, as we saw in Section 2.2, was the difference between a fundamental niche (where an organism *could* live) and a realized niche (where it *actually* lived). For example, an organism will suffer if its food is another species that cannot tolerate an environmental condition. This is illustrated by the distribution of the rush moth (*Coleophora alticolella*) in England. The moth lays its eggs on the flowers of the rush *Juncus squarrosus* and the caterpillars feed on the developing seeds. Above 600 m, the moths and caterpillars are little affected by the low temperatures, but the rush, although it grows, fails to ripen its seeds. This, in turn, limits the distribution of the moth, because caterpillars that hatch in the colder elevations will starve as a result of insufficient food (Randall, 1982).

disease

The effects of conditions on disease may also be important. Conditions may favor the spread of infection (winds carrying fungal spores), or favor the growth of the parasite, or weaken the defenses of the host. For example, during an epidemic of southern corn leaf blight (*Helminthosporium maydis*) in a corn field in Connecticut, the plants closest to the trees that were shaded for the longest periods were the most heavily diseased (Figure 2.15).

competition

Competition between species can also be profoundly influenced by environmental conditions, especially temperature. Two stream salmonid fishes, *Salvelinus malma* and *S. leucomaenis*, coexist at intermediate altitudes (and therefore intermediate temperatures) on Hokkaido Island, Japan, whereas only the former lives at higher altitudes (lower temperatures) and only the latter at lower altitudes (see also Section 8.2.1). A reversal, by a change in temperature, of the outcome of competition between the species appears to play a key role in this. For example, in experimental streams supporting the two species maintained at 6°C over a 191-day period (a typical high altitude temperature), the survival of *S. malma* was far superior to that of *S. leucomaenis*; whereas at 12°C (typical low altitude), both species survived less well, but the outcome was so far reversed that by around 90 days all of the *S. malma* had died (Figure 2.16). Both species are quite capable, alone, of living at either temperature.

temperature and humidity

Many of the interactions between temperature and other physical conditions are so strong that it is not sensible to consider them separately. The relative humidity of the atmosphere, for example, is an important condition in the life of terrestrial organisms because it plays a major part in determining the rate at which they lose water. In practice, it is rarely possible to make a clean distinction between the effects of relative humidity and of temperature. This is simply because a rise in temperature leads to an increased rate of evaporation. A relative humidity that is acceptable to an organism at a low temperature may therefore be unacceptable at a higher temperature. Microclimatic variations in relative humidity can be even more marked than those involving temperature. For instance, it is not unusual for the relative humidity to be almost 100% at ground level amongst dense vegetation and within the soil, whilst the air immediately above, perhaps 40 cm away, has a relative humidity

Figure 2.16 Changing temperature reverses the outcome of competition. At low temperature (6°C) on the left, the salmonid fish *Salvelinus malma* outsurvives cohabiting *S. leucomaenis*, whereas at 12°C, on the right, *S. leucomaenis* drives *S. malma* to extinction. Both species are quite capable, alone, of living at either temperature. (After Taniguchi & Nakano, 2000.)
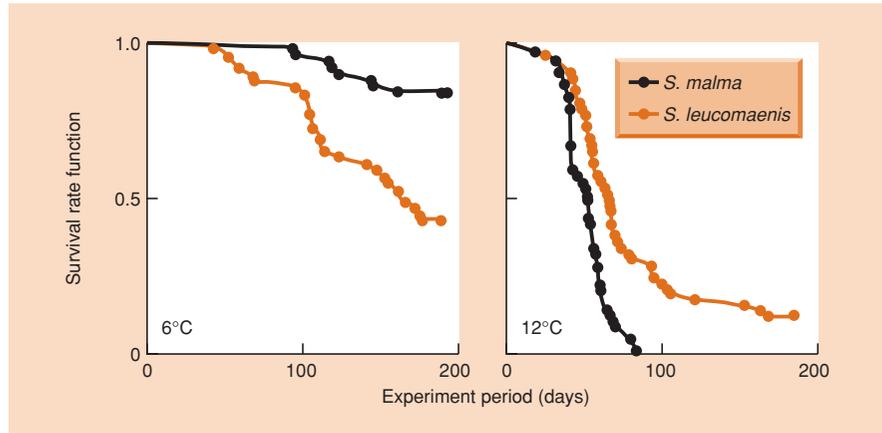
of only 50%. The organisms most obviously affected by humidity in their distribution are those 'terrestrial' animals that are actually, in terms of the way they control their water balance, 'aquatic'. Amphibians, terrestrial isopods, nematodes, earthworms and molluscs are all, at least in their active stages, confined to microenvironments where the relative humidity is at or very close to 100%. The major group of animals to escape such confinement are the terrestrial arthropods, especially insects. Even here though, the evaporative loss of water often confines their activities to habitats (e.g. woodlands) or times of day (e.g. dusk) when relative humidity is relatively high.

2.5   pH of soil and water

The pH of soil in terrestrial environments or of water in aquatic ones is a condition that can exert a powerful influence on the distribution and abundance of organisms. The protoplasm of the root cells of most vascular plants is damaged as a direct result of toxic concentrations of H^+ or OH^- ions in soils below pH 3 or above pH 9, respectively. Further, indirect effects occur because soil pH influences the availability of nutrients and/or the concentration of toxins (Figure 2.17).

Increased acidity (low pH) may act in three ways: (i) directly, by upsetting osmoregulation, enzyme activity or gaseous exchange across respiratory surfaces; (ii) indirectly, by increasing the concentration of toxic heavy metals, particularly aluminum (Al^{3+}) but also manganese (Mn^{2+}) and iron (Fe^{3+}), which are essential plant nutrients at higher pHs; and (iii) indirectly, by reducing the quality and range of food sources available to animals (e.g. fungal growth is reduced at low pH in streams (Hildrew *et al.*, 1984) and the aquatic flora is often absent or less diverse). Tolerance limits for pH vary amongst plant species, but only a minority are able to grow and reproduce at a pH below about 4.5.

In alkaline soils, iron (Fe^{3+}) and phosphate (PO_4^{3+}), and certain trace elements such as manganese (Mn^{2+}), are fixed in relatively
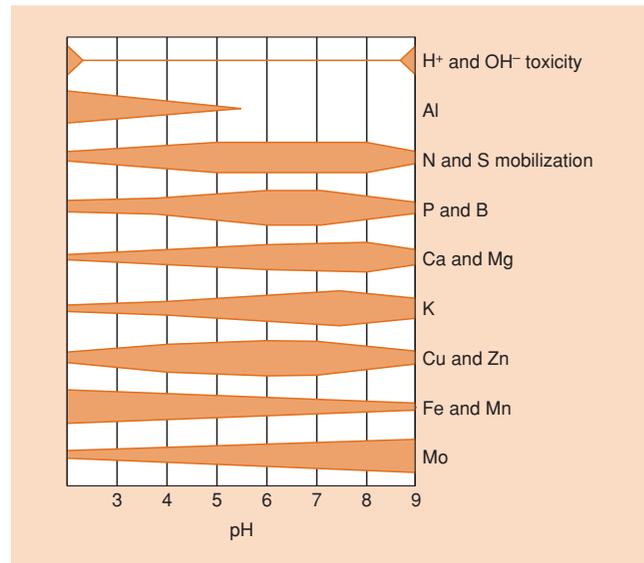
Fgiure 2.17 The toxicity of H^+ and OH^- to plants, and the availability to them of minerals (indicated by the widths of the bands) is influenced by soil pH. (After Larcher, 1980.)

insoluble compounds, and plants may then suffer because there is too little rather than too much of them. For example, calcifuge plants (those characteristic of acid soils) commonly show symptoms of iron deficiency when they are transplanted to more alkaline soils. In general, however, soils and waters with a pH above 7 tend to be hospitable to many more species than those that are more acid. Chalk and limestone grasslands carry a much richer flora (and associated fauna) than acid grasslands and the situation is similar for animals inhabiting streams, ponds and lakes.

Some prokaryotes, especially the Archaebacteria, can tolerate and even grow best in environments with a pH far outside the range tolerated by eukaryotes. Such environments are rare, but occur in volcanic lakes and geothermal springs where they are

dominated by sulfur-oxidizing bacteria whose pH optima lie between 2 and 4 and which cannot grow at neutrality (Stolp, 1988). *Thiobacillus ferroxidans* occurs in the waste from industrial metal-leaching processes and tolerates pH 1; *T. thiooxidans* cannot only tolerate but can grow at pH 0. Towards the other end of the pH range are the alkaline environments of soda lakes with pH values of 9–11, which are inhabited by cyanobacteria such as *Anabaenopsis arnoldii* and *Spirulina platensis*; *Plectonema nostocorum* can grow at pH 13.

2.6 Salinity

For terrestrial plants, the concentration of salts in the soil water offers osmotic resistance to water uptake. The most extreme saline conditions occur in arid zones where the predominant movement of soil water is towards the surface and cystalline salt accumulates. This occurs especially when crops have been grown in arid regions under irrigation; salt pans then develop and the land is lost to agriculture. The main effect of salinity is to create the same kind of osmoregulatory problems as drought and freezing and the problems are countered in much the same ways. For example, many of the higher plants that live in saline environments (halophytes) accumulate electrolytes in their vacuoles, but maintain a low concentration in the cytoplasm and organelles (Robinson *et al.*, 1983). Such plants maintain high osmotic pressures and so remain turgid, and are protected from the damaging action of the accumulated electrolytes by polyols and membrane protectants.

Freshwater environments present a set of specialized environmental conditions because water tends to move into organisms from the environment and this needs to be resisted. In marine habitats, the majority of organisms are isotonic to their environment so that there is no net flow of water, but there are many that are hypotonic so that water flows out from the organism to the environment, putting them in a similar position to terrestrial organisms. Thus, for many aquatic organisms the regulation of body fluid concentration is a vital and sometimes an energetically expensive process. The salinity of an aquatic environment can have an important influence on distribution and abundance, especially in places like estuaries where there is a particularly sharp gradient between truly marine and freshwater habitats.

The freshwater shrimps *Palaemonetes pugio* and *P. vulgaris*, for example, co-occur in estuaries on the eastern coat of the USA at a wide range of salinities, but the former seems to be more tolerant of lower salinities than the latter, occupying some habitats from which the latter is absent. Figure 2.18 shows the mechanism likely to be underlying this (Rowe, 2002). Over the low salinity range (though not at the effectively lethal lowest salinity) metabolic expenditure was significantly lower in *P. pugio*. *P. vulgaris* requires far more energy simply to maintain itself, putting it at a severe disadvantage in competition with *P. pugio* even when it is able to sustain such expenditure.

2.6.1 Conditions at the boundary between the sea and land

Salinity has important effects on the distribution of organisms in intertidal areas but it does so through interactions with other conditions – notably exposure to the air and the nature of the substrate.
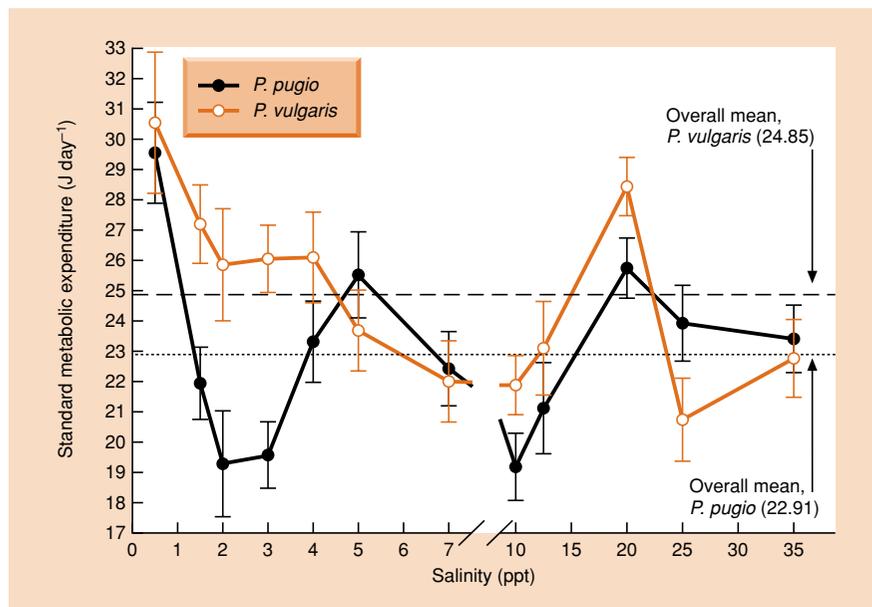

Figure 2.18 Standard metabolic expenditure (estimated through minimum oxygen consumption) in two species of shrimp, *Palaemonetes pugio* and *P. vulgaris*, at a range of salinities. There was significant mortality of both species over the experimental period at 0.5 ppt (parts per thousand), especially in *P. vulgaris* (75% compared with 25%). (After Rowe, 2002.)

Algae of all types have found suitable habitats permanently immersed in the sea, but permanently submerged higher plants are almost completely absent. This is a striking contrast with submerged freshwater habitats where a variety of flowering plants have a conspicuous role. The main reason seems to be that higher plants require a substrate in which their roots can find anchorage. Large marine algae, which are continuously submerged except at extremely low tides, largely take their place in marine communities. These do not have roots but attach themselves to rocks by specialized 'holdfasts'. They are excluded from regions where the substrates are soft and holdfasts cannot 'hold fast'. It is in such regions that the few truly marine flowering plants, for example sea grasses such as *Zostera* and *Posidonia*, form submerged communities that support complex animal communities.

algae and higher plants Most species of higher plants that root in seawater have leaves and shoots that are exposed to the atmosphere for a large part of the tidal cycle, such as mangroves, species of the grass genus *Spartina* and extreme halophytes such as species of *Salicornia* that have aerial shoots but whose roots are exposed to the full salinity of seawater. Where there is a stable substrate in which plants can root, communities of flowering plants may extend right through the intertidal zone in a continuum extending from those continuously immersed in full-strength seawater (like the sea grasses) through to totally nonsaline conditions. Salt marshes, in particular, encompass a range of salt concentrations running from full-strength seawater down to totally nonsaline conditions.

Higher plants are absent from intertidal rocky sea shores except where pockets of soft substrate may have formed in crevices. Instead, such habitats are dominated by the algae, which give way to lichens at and above the high tide level where the exposure to desiccation is highest. The plants and animals that live on rocky sea shores are influenced by environmental conditions in a very profound and often particularly obvious way by the extent to which they tolerate exposure to the aerial environment and the forces of waves and storms. This expresses itself in the *zonation* of the organisms, with different species at different heights up the shore (Figure 2.19).

zonation The extent of the intertidal zone depends on the height of tides and the slope of the shore. Away from the shore, the tidal rise and fall are rarely greater than 1 m, but closer to shore, the shape of the land mass can funnel the ebb and flow of the water to produce extraordinary spring tidal ranges of, for example, nearly 20 m in the Bay of Fundy (between Nova Scotia and New Brunswick, Canada). In contrast, the shores of the Mediterranean Sea
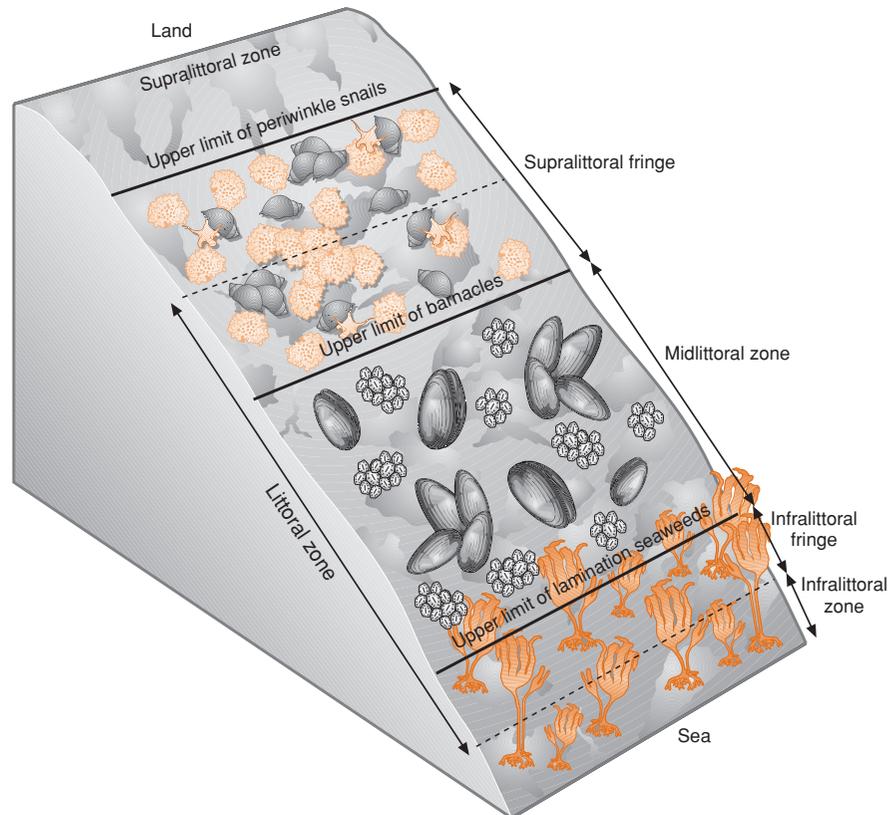

Figure 2.19   A general zonation scheme for the seashore determined by relative lengths of exposure to the air and to the action of waves. (After Raffaelli & Hawkins, 1996.)

experience scarcely any tidal range. On steep shores and rocky cliffs the intertidal zone is very short and zonation is compressed.

To talk of 'zonation as a result of exposure', however, is to oversimplify the matter greatly (Raffaelli & Hawkins, 1996). In the first place, 'exposure' can mean a variety, or a combination of, many different things: desiccation, extremes of temperature, changes in salinity, excessive illumination and the sheer physical forces of pounding waves and storms (to which we turn in Section 2.7). Furthermore, 'exposure' only really explains the *upper* limits of these essentially marine species, and yet zonation depends on them having lower limits too. For some species there can be *too little exposure* in the lower zones. For instance, green algae would be starved of blue and especially red light if they were submerged for long periods too low down the shore. For many other species though, a lower limit to distribution is set by competition and predation (see, for example, the discussion in Paine, 1994). The seaweed *Fucus spiralis* will readily extend lower down the shore than usual in Great Britain whenever other competing midshore fucoid seaweeds are scarce.

2.7   Physical forces of winds, waves and currents

In nature there are many forces of the environment that have their effect by virtue of the force of physical movement – wind and water are prime examples.

In streams and rivers, both plants and animals face the continual hazard of being washed away. The average velocity of flow generally increases in a downstream direction, but the greatest danger of members of the benthic (bottom-dwelling) community being washed away is in upstream regions, because the water here is turbulent and shallow. The only plants to be found in the most extreme flows are literally 'low profile' species like encrusting and filamentous algae, mosses and liverworts. Where the flow is slightly less extreme there are plants like the water crowfoot (*Ranunculus fluitans*), which is streamlined, offering little resistance to flow and which anchors itself around an immovable object by means of a dense development of adventitious roots. Plants such as the free-floating duckweed (*Lemna* spp.) are usually only found where there is negligible flow.

The conditions of exposure on sea shores place severe limits on the life forms and habits of species that can tolerate repeated pounding and the suction of wave action. Seaweeds anchored on rocks survive the repeated pull and push of wave action by a combination of powerful attachment by holdfasts and extreme flexibility of their thallus structure. Animals in the same environment either move with the mass of water or, like the algae, rely on subtle mechanisms of firm adhesion such as the powerful organic glues of barnacles and the muscular feet of limpets. A comparable diversity of morphological specializations is to be found amongst the invertebrates that tolerate the hazards of turbulent, freshwater streams.

2.7.1   Hazards, disasters and catastrophes: the ecology of extreme events

The wind and the tides are normal daily 'hazards' in the life of many organisms. The structure and behavior of these organisms bear some witness to the frequency and intensity of such hazards in the evolutionary history of their species. Thus, most trees withstand the force of most storms without falling over or losing their living branches. Most limpets, barnacles and kelps hold fast to the rocks through the normal day to day forces of the waves and tides. We can also recognize a scale of more severely damaging forces (we might call them 'disasters') that occur occasionally, but with sufficient frequency to have contributed repeatedly to the forces of natural selection. When such a force recurs it will meet a population that still has a genetic memory of the selection that acted on its ancestors – and may therefore suffer less than they did. In the woodlands and shrub communities of arid zones, fire has this quality, and tolerance of fire damage is a clearly evolved response (see Section 2.3.6).

When disasters strike natural communities it is only rarely that they have been carefully studied before the event. One exception is cyclone 'Hugo' which struck the Caribbean island of Guadeloupe in 1994. Detailed accounts of the dense humid forests of the island had been published only recently before (Ducrey & Labbé, 1985, 1986). The cyclone devastated the forests with mean maximum wind velocities of 270 km h^{-1} and gusts of 320 km h^{-1}. Up to 300 mm of rain fell in 40 h. The early stages of regeneration after the cyclone (Labbé, 1994) typify the responses of long-established communities on both land or sea to massive forces of destruction. Even in 'undisturbed' communities there is a continual creation of gaps as individuals (e.g. trees in a forest, kelps on a sea shore) die and the space they occupied is recolonized (see Section 16.7). After massive devastation by cyclones or other widespread disasters, recolonization follows much the same course. Species that normally colonize only natural gaps in the vegetation come to dominate a continuous community.

In contrast to conditions that we have called 'hazards' and 'disasters' there are natural occurrences that are enormously damaging, yet occur so rarely that they may have no lasting selective effect on the evolution of the species. We might call such events 'catastrophes', for example the volcanic eruption of Mt St Helens or of the island of Krakatau. The next time that Krakatau erupts there are unlikely to be any genes persisting that were selected for volcano tolerance!

2.8   Environmental pollution

A number of environmental conditions that are, regrettably, becoming increasingly important are due to the accumulation of toxic by-products of human activities. Sulfur dioxide emitted from power stations, and metals like copper, zinc and lead, dumped
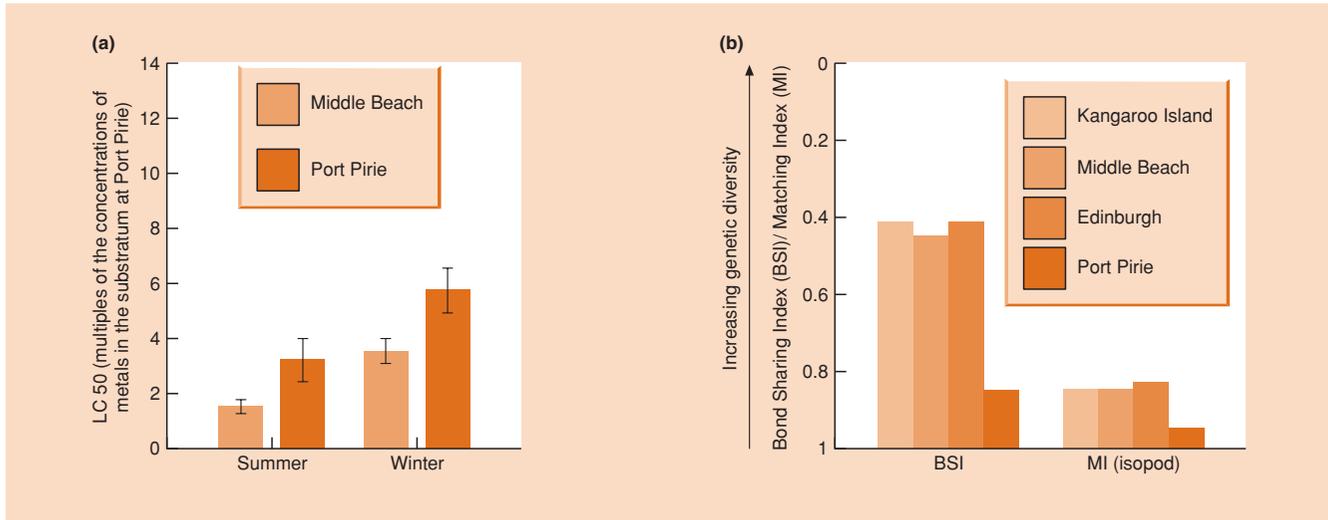
Figure 2.20 The response of the marine isopod, *Platynympha longicaudata*, to pollution around the largest lead smelting operation in the world, Port Pirie, South Australia. (a) Tolerance, both summer and winter, was significantly higher ($P < 0.05$) than for animals from a control (unpolluted) site, as measured by the concentration in food of a combination of metals (lead, copper, cadmium, zinc and manganese) required to kill 50% of the population (LC50). (b) Genetic diversity at Port Pirie was significantly lower than at three unpolluted sites, as measured by two indices of diversity based on RAPDs (random amplified polymorphic DNA). (After Ross *et al.*, 2002.)

around mines or deposited around refineries, are just some of the pollutants that limit distributions, especially of plants. Many such pollutants are present naturally but at low concentrations, and some are indeed essential nutrients for plants. But in polluted areas their concentrations can rise to lethal levels. The loss of species is often the first indication that pollution has occurred, and changes in the species richness of a river, lake or area of land provide bioassays of the extent of their pollution (see, for example, Lovett Doust *et al.*, 1994).

rare tolerators

Yet it is rare to find even the most inhospitable polluted areas entirely devoid of species; there are usually at least a few individuals of a few species that can tolerate the conditions. Even natural populations from unpolluted areas often contain a low frequency of individuals that tolerate the pollutant; this is part of the genetic variability present in natural populations. Such individuals may be the only ones to survive or colonize as pollutant levels rise. They may then become the founders of a tolerant population to which they have passed on their 'tolerance' genes, and, because they are the descendants of just a few founders, such populations may exhibit notably low genetic diversity overall (Figure 2.20). Moreover, species themselves may differ greatly in their ability to tolerate pollutants. Some plants, for example, are 'hyperaccumulators' of heavy metals – lead, cadmium and so on – with an ability not only to tolerate but also to accumulate much higher concentrations than the norm (Brooks, 1998). As a result, such plants may have an important role to play in 'bioremediation' (Salt *et al.*, 1998), removing pollutants from the soil so that

eventually other, less tolerant plants can grow there too (discussed further in Section 7.2.1).

Thus, in very simple terms, a pollutant has a twofold effect. When it is newly arisen or is at extremely high concentrations, there will be few individuals of any species present (the exceptions being naturally tolerant variants or their immediate descendants). Subsequently, however, the polluted area is likely to support a much higher density of individuals, but these will be representatives of a much smaller range of species than would be present in the absence of the pollutant. Such newly evolved, species-poor communities are now an established part of human environments (Bradshaw, 1987).

Pollution can of course have its effects far from the original source (Figure 2.21). Toxic effluents from a mine or a factory may enter a watercourse and affect its flora and fauna for its whole length downstream. Effluents from large industrial complexes can pollute and change the flora and fauna of many rivers and lakes in a region and cause international disputes.

A striking example is the creation of 'acid rain' – for example that falling in Ireland and Scandinavia from indus-

acid rain

trial activities in other countries. Since the Industrial Revolution, the burning of fossil fuels and the consequent emission to the atmosphere of various pollutants, notably sulfur dioxide, has produced a deposition of dry acidic particles and rain that is essentially dilute sulfuric acid. Our knowledge of the pH tolerances of diatom species enables an approximate pH history of a lake to be constructed. The history of the acidification of lakes is often
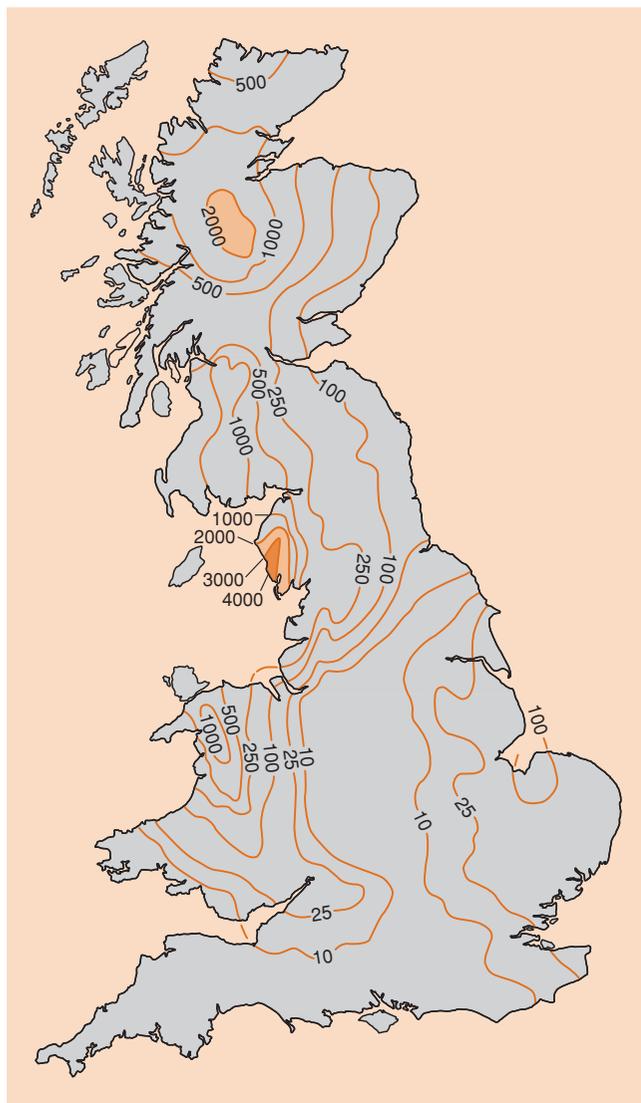
Figure 2.21 An example of long-distance environmental pollution. The distribution in Great Britain of fallout of radioactive caesium (Bq m^{-2}) from the Chernobyl nuclear accident in the Soviet Union in 1986. The map shows the persistence of the pollutant on acid upland soils where it is recycled through soils, plants and animals. Sheep in the upland areas contained more caesium-137 (^{137}Cs) in 1987 and 1988 (after recycling) than in 1986. ^{137}Cs has a half-life of 30 years! On typical lowland soils it is more quickly immobilized and does not persist in the food chains. (After NERC, 1990.)

recorded in the succession of diatom species accumulated in lake sediments (Flower *et al.*, 1994). Figure 2.22, for example, shows how diatom species composition has changed in Lough Maam, Ireland – far from major industrial sites. The percentage of various diatom species at different depths reflects the flora present

at various times in the past (four species are illustrated). The age of layers of sediment can be determined by the radioactive decay of lead-210 (and other elements). We know the pH tolerance of the diatom species from their present distribution and this can be used to reconstruct what the pH of the lake has been in the past. Note how the waters acidified since about 1900. The diatoms *Fragilaria virescens* and *Brachysira vitrea* have declined markedly during this period while the acid-tolerant *Cymbella perpusilla* and *Frustulia rhomboides* increased after 1900.

2.9   Global change

In Chapter 1 we discussed some of the ways in which global environments have changed over the long timescales involved in continental drift and the shorter timescales of the repeated ice ages. Over these timescales some organisms have failed to accommodate to the changes and have become extinct, others have migrated so that they continue to experience the same conditions but in a different place, and it is probable that others have changed their nature (evolved) and tolerated some of the changes. We now turn to consider global changes that are occurring in our own lifetimes – consequences of our own activities – and that are predicted, in most scenarios, to bring about profound changes in the ecology of the planet.

2.9.1   Industrial gases and the greenhouse effect

A major element of the Industrial Revolution was the switch from the use of sustainable fuels to the use of coal (and later, oil) as a source of power. Between the middle of the 19th and the middle of the 20th century the burning of fossil fuels, together with extensive deforestation, added about 9×10^{10} tonnes of carbon dioxide (CO_2) to the atmosphere and even more has been added since. The concentration of CO_2 in the atmosphere before the Industrial Revolution (measured in gas trapped in ice cores) was about 280 ppm, a fairly typical interglacial 'peak' (Figure 2.23), but this had risen to around 370 ppm by around the turn of the millennium and is still rising (see Figure 18.22).

Solar radiation incident on the earth's atmosphere is in part reflected, in part absorbed, and part is transmitted through to the earth's surface, which absorbs and is warmed by it. Some of this absorbed energy is radiated back to the atmosphere where atmospheric gases, mainly water vapor and CO_2 absorb about 70% of it. It is this trapped reradiated energy that heats the atmosphere in what is called the 'greenhouse effect'. The greenhouse effect was of course part of the normal environment before the Industrial Revolution and carried responsibility for some of the environmental warmth before industrial activity started to enhance it. At that time, atmospheric water vapor was responsible for the greater portion of the greenhouse effect.
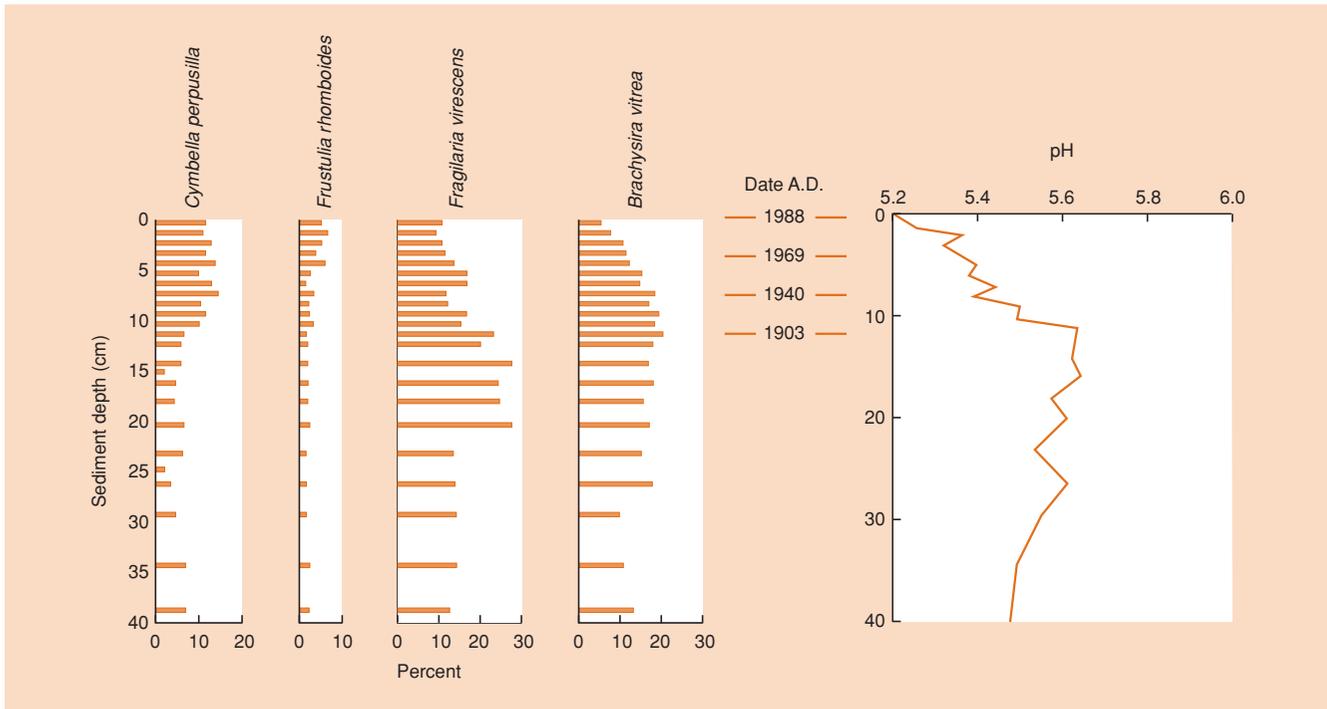
Figure 2.22   The history of the diatom flora of an Irish lake (Lough Maam, County Donegal) can be traced by taking cores from the sediment at the bottom of the lake. The percentage of various diatom species at different depths reflects the flora present at various times in the past (four species are illustrated). The age of the layers of sediment can be determined by the radioactive decay of lead-210 (and other elements). We know the pH tolerance of the diatom species from their present distribution and this can be used to reconstruct what the pH of the lake has been in the past. Note how the waters have been acidified since about 1900. The diatoms *Fragilaria virescens* and *Brachysira vitrea* have declined markedly during this period, while the acid-tolerant *Cymbella perpusilla* and *Frustulia rhomboides* have increased. (After Flower *et al.*, 1994.)
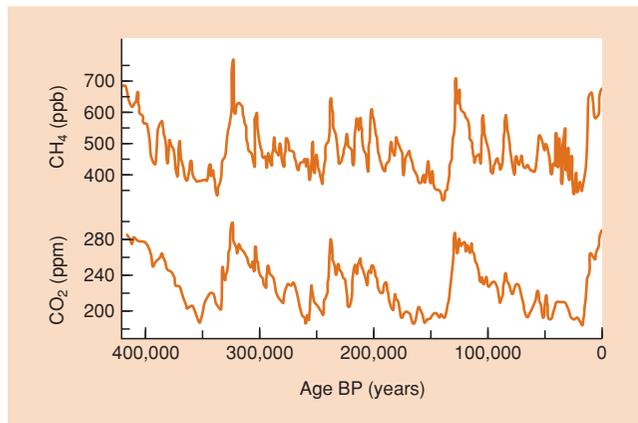

Figure 2.23   Concentrations of CO_2 and methane (CH_4) in gas trapped in ice cores from Vostok, Antarctica deposited over the past 420,000 years. Estimated temperatures are very strongly correlated with these. Thus, transitions between glacial and warm epochs occurred around 335,000, 245,000, 135,000 and 18,000 years ago. BP, before present; ppb, parts per billion; ppm, parts per million. (After Petit *et al.*, 1999; Stauffer, 2000.)

In addition to the enhancement of greenhouse effects by increased CO_2, other trace gases have increased markedly in the atmosphere, particularly methane (CH_4) (Figure 2.24a; and compare this with the historical record in Figure 2.23), nitrous oxide (N_2O) and the chlorofluorocarbons (CFCs, e.g. trichlorofluoromethane (CCl_3F) and dichlorodifluoromethane (CCl_2F_2)). Together, these and other gases contribute almost as much to enhancing the greenhouse effect as does the rise in CO_2 (Figure 2.24b). The increase in CH_4 is not all explained but probably has a microbial origin in intensive agriculture on anaerobic soils (especially increased rice production) and in the digestive process of ruminants (a cow produces approximately 40 litres of CH_4 each day); around 70% of its production is anthropogenic (Khalil, 1999). The effect of the CFCs from refrigerants, aerosol propellants and so on is potentially great, but international agreements at least appear to have halted further rises in their concentrations (Khalil, 1999).

It should be possible to draw up a balance sheet that shows how the CO_2 produced by human activities translates into the changes in concentration in the atmosphere. Human activities

> CO_2 – but not only CO_2
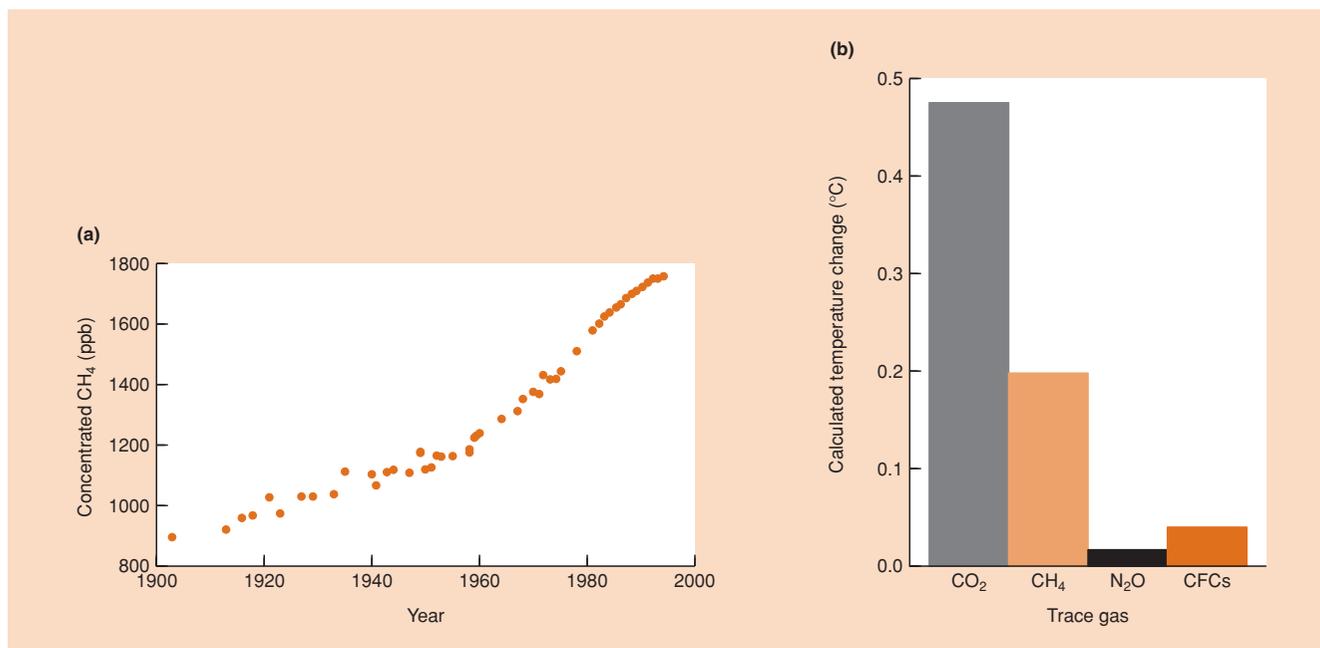
Figure 2.24   (a) Concentration of methane (CH_4) in the atmosphere through the 20th century. (b) Estimates of global warming over the period 1850–1990 caused by CO_2 and other major greenhouse gases. (After Khalil, 1999.)

release $5.1–7.5 \times 10^9$ metric tons of carbon to the atmosphere each year. But the increase in atmospheric CO_2 (2.9×10^9 metric tons) accounts for only 60% of this, a percentage that has remained remarkably constant for 40 years (Hansen *et al.*, 1999). The oceans absorb CO_2 from the atmosphere, and it is estimated that they may absorb $1.8–2.5 \times 10^9$ metric tons of the carbon released by human activities. Recent analyses also indicate that terrestrial vegetation has been 'fertilized' by the increased atmospheric CO_2, so that a considerable amount of extra carbon has been locked up in vegetation biomass (Kicklighter *et al.*, 1999). This softening of the blow by the oceans and terrestrial vegetation notwithstanding, however, atmospheric CO_2 and the greenhouse effect are increasing. We return to the question of global carbon budgets in Section 18.4.6.

2.9.2   Global warming

We started this chapter discussing temperature, moved through a number of other environmental conditions to pollutants, and now return to temperature because of the effects of those pollutants on global temperatures. It appears that the present air temperature at the land surface is $0.6 \pm 0.2°C$ warmer than in preindustrial times (Figure 2.25), and temperatures are predicted to continue to rise by a further $1.4–5.8°C$ by 2100 (IPCC, 2001). Such changes will probably lead to a melting of the ice caps, a consequent rising of sea level and large changes in the pattern of

global climates and the distribution of species. Predictions of the extent of global warming resulting from the enhanced greenhouse effect come from two sources: (i) predictions based on sophisticated computer models ('general circulation models') that simulate the world's climate; and (ii) trends detected in measured data sets, including the width of tree rings, sea-level records and measures of the rate of retreat of glaciers.

Not surprisingly, different global circulation models differ in their predictions of the rise in global temperature that will result from predicted increases in CO_2. However, most model predictions vary only from 2.3 to $5.2°C$ (most of the variation is accounted for by the way in which the effects of cloud cover are modeled), and a projected rise of $3–4°C$ in the next 100 years seems a reasonable value from which to make projections of ecological effects (Figure 2.26).

a 3–4°C rise in the next 100 years

But temperature regimes are, of course, only part of the set of conditions that determine which organisms live where. Unfortunately, we can place much less faith in computer projections of rainfall and evaporation because it is very hard to build good models of cloud behavior into a general model of climate. If we consider only temperature as a relevant variable, we would project a $3°C$ rise in temperature giving London (UK) the climate of Lisbon (Portugal) (with an appropriate vegetation of olives, vines, *Bougainvillea* and semiarid scrub). But with more reliable rain it would be nearly subtropical, and with a little less it might qualify for the status of an arid zone!

Figure 2.25   Global annual surface temperature variations from 1860 to 1998. The bars show departures from the mean at the end of the 19th century. The curve is a moving average obtained using a 21-year filter. Mean global temperatures are now higher than at any time since 1400. (After Saunders, 1999.)
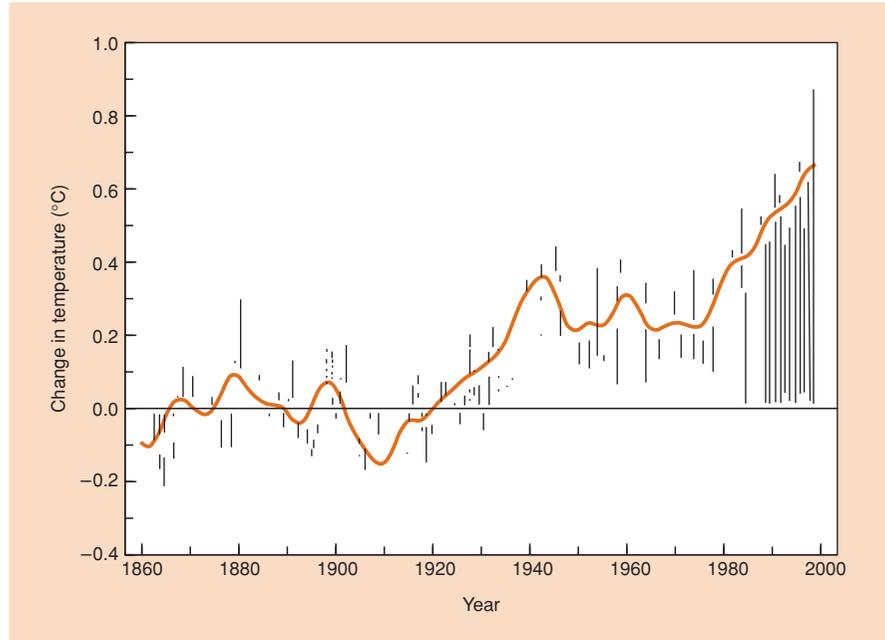

Figure 2.26   The rise in global mean surface temperature projected by the global coupled model (i.e. both the oceans and the atmosphere are modeled) for climate variability and change in use at the Geophysical Fluid Dynamics Laboratory, Princeton, USA. Observed increases in greenhouse gases are used for the period 1865–1990 (and clearly the projections match closely the observed trend in temperature); thereafter, greenhouse gases are assumed to increase at 1% per year. Since the model simulates the global behavior of the oceans and atmosphere, the precise behavior depends on the initial state of the system. The three 'experiments' were started from different states. (After Delworth *et al.*, 2002.)
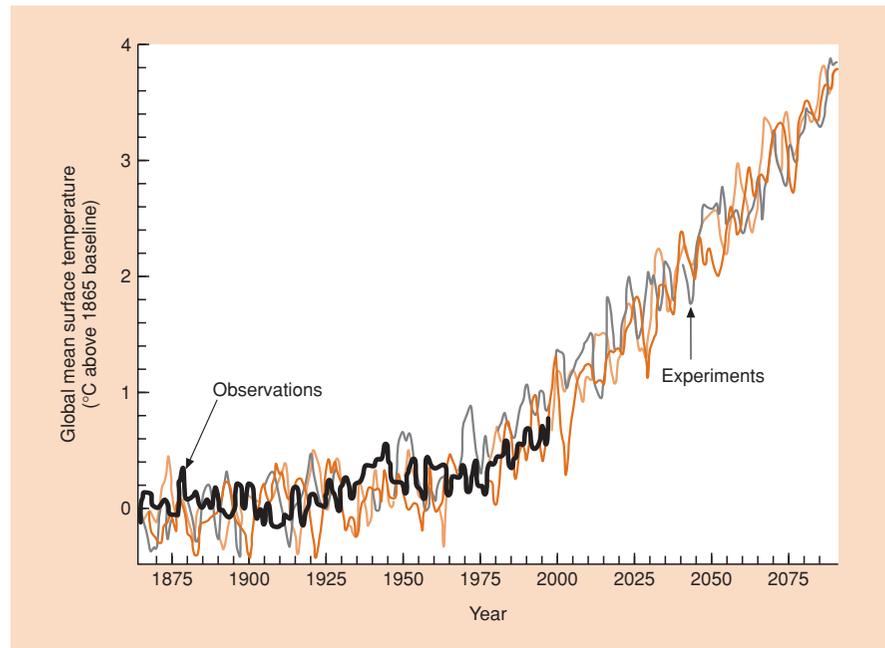
the global distribution of climate change

Also, global warming is not evenly distributed over the surface of the earth. Figure 2.27 shows the measured global change in the trends of surface temperature over the 46 years from 1951 to 1997. Areas of North America (Alaska) and Asia experienced rises of 1.5–2°C in that period, and these places are predicted to continue experiencing the fastest warming in the first half of the present century. In some regions the temperature has apparently not changed (New York, for example) and should not change greatly in the next 50 years. There are also some areas, notably Greenland and the northern Pacific Ocean, where surface temperatures have fallen.

We have emphasized, too, that the distribution of many organisms is determined by occasional extremes rather than by average conditions. Computer modeled projections imply that
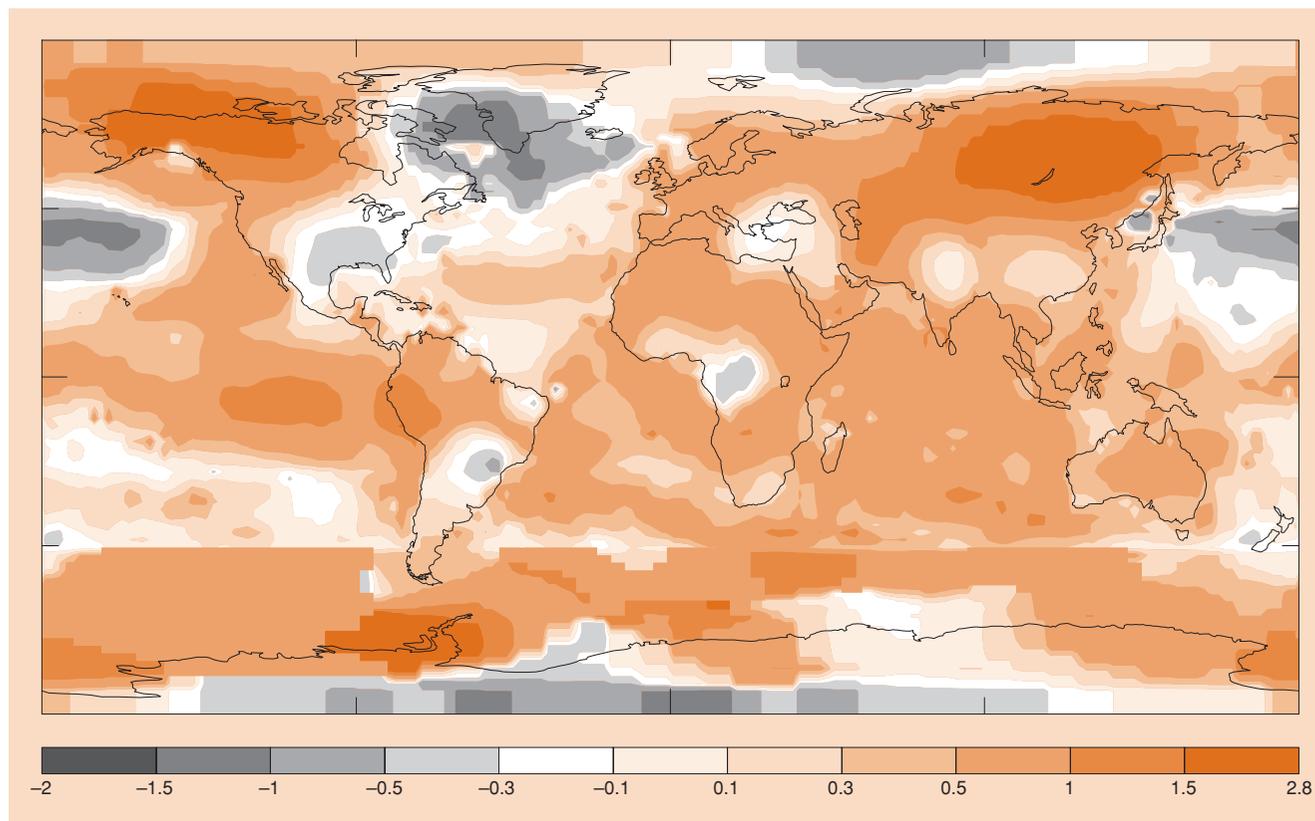
Figure 2.27   Change in the surface temperature of the globe expressed as the linear trend over 46 years from 1951 to 1997. The bar below gives the temperatures in °C. (From Hansen *et al.*, 1999.)

global climatic change will also bring greater variance in temperature. Timmerman *et al.* (1999), for example, modeled the effect of greenhouse warming on the ENSO (see Section 2.4.1). They found that not only was the mean climate in the tropical Pacific region predicted to move towards that presently represented by the (warmer) El Niño state, but that interannual variability was also predicted to increase and that variability was predicted to be more skewed towards unusually cold events.

can the biota keep up with the pace?

Global temperatures have changed naturally in the past, as we have seen. We are currently approaching the end of one of the warming periods that started around 20,000 years ago, during which global temperatures have risen by about 8°C. The greenhouse effect adds to global warming at a time when temperatures are already higher than they have been for 400,000 years. Buried pollen gives us evidence that North American forest boundaries have migrated north at rates of 100–500 m year^{-1} since the last ice age. However, this rate of advance has not been fast enough to keep pace with postglacial warming. The rate of warming forecast to result from the greenhouse effect is 50–100 times faster than postglacial warming. Thus, of all the types of environmental pollution caused by human activities, none may have such profound effects as global warming. We must expect latitudinal and altitudinal changes to species' distributions and widespread extinctions as floras and faunas fail to track and keep up with the rate of change in global temperatures (Hughes, 2000). What is more, large tracts of land over which vegetation might advance and retreat have been fragmented in the process of civilization, putting major barriers in the way of vegetational advance. It will be very surprising if many species do not get lost on the journey.

Summary

A condition is an abiotic environmental factor that influences the functioning of living organisms. For most, we can recognize an optimum level at which an organism performs best. Ultimately, we should define 'performs best' from an evolutionary point of view, but in practice we mostly measure the effect of conditions on some key property like the activity of an enzyme or the rate of reproduction.

The ecological niche is not a place but a summary of an organism's tolerances of conditions and requirements for resources. The

modern concept – Hutchinson's *n*-dimensional hypervolume – also distinguishes fundamental and realized niches.

Temperature is discussed in detail as a typical, and perhaps the most important, condition. Individuals respond to temperature with impaired function and ultimately death at upper and lower extremes, with a functional range between the extremes, within which there is an optimum, although these responses may be subject to evolutionary adaptation and to more immediate acclimatization.

The rates of biological enzymatic processes often increase exponentially with temperature (often $Q_{10} \approx 2$), but for rates of growth and development there are often only slight deviations from linearity: the basis for the day-degree concept. Because development usually increases more rapidly with temperature than does growth, final size tends to decrease with rearing temperature. Attempts to uncover universal rules of temperature dependence remain a matter of controversy.

We explain the differences between endotherms and ectotherms but also the similarities between them, ultimately, in their responses to a range of temperatures.

We examine variations in temperature on and within the surface of the earth with a variety of causes: latitudinal, altitudinal, continental, seasonal, diurnal and microclimatic effects, and, in soil and water, the effects of depth. Increasingly, the importance of medium-term temporal patterns have become apparent. Notable amongst these are the El Niño–Southern Oscillation (ENSO) and the North Atlantic Oscillation (NAO).

There are very many examples of plant and animal distributions that are strikingly correlated with some aspect of environmental temperature but these do not prove that temperature directly causes the limits to a species' distribution. The temperatures measured are only rarely those that the organisms experience. For many species, distributions are accounted for not so much by average temperatures as by occasional extremes; and the effects of temperature may be determined largely by the responses of other community members or by interactions with other conditions.

A range of other environmental conditions are also discussed: the pH of soil and water, salinity, conditions at the boundary between sea and land, and the physical forces of winds, waves and currents. Hazards, disasters and catastrophes are distinguished.

A number of environmental conditions are becoming increasingly important due to the accumulation of toxic by-products of human activities. A striking example is the creation of 'acid rain'. Another is the effect of industrial gases on the greenhouse effect and consequent effects on global warming. A projected rise of 3–4°C in the next 100 years seems a reasonable value from which to make projections of ecological effects, though global warming is not evenly distributed over the surface of the earth. This rate is 50–100 times faster than postglacial warming. We must expect latitudinal and altitudinal changes to species' distributions and widespread extinctions of floras and faunas.
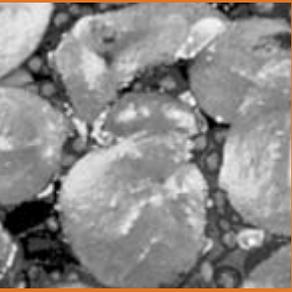
Chapter 3
Resources

3.1 Introduction

what are resources?

According to Tilman (1982), all things consumed by an organism are resources for it. But consumed does not simply mean 'eaten'. Bees and squirrels do not eat holes, but a hole that is occupied is no longer available to another bee or squirrel, just as an atom of nitrogen, a sip of nectar or a mouthful of acorn are no longer available to other consumers. Similarly, females that have already mated may be unavailable to other mates. All these things have been consumed in the sense that the stock or supply has been reduced. Thus, resources are entities required by an organism, the quantities of which can be reduced by the activity of the organism.

organisms may compete for resources

Green plants photosynthesize and obtain both energy and matter for growth and reproduction from inorganic materials. Their resources are solar radiation, carbon dioxide (CO_2), water and mineral nutrients. 'Chemosynthetic' organisms, such as many of the Archaebacteria, obtain energy by oxidizing methane, ammonium ions, hydrogen sulfide or ferrous iron; they live in environments such as hot springs and deep sea vents and use resources that were much more abundant during early phases of life on earth. All other organisms use as their food resource the bodies of other organisms. In each case, what has been consumed is no longer available to another consumer. The rabbit eaten by an eagle is no longer available to another eagle. The quantum of solar radiation absorbed and photosynthesized by a leaf is no longer available to another leaf. This has an important consequence: organisms may *compete* with each other to capture a share of a limited resource – a topic that will occupy us in Chapter 5.

A large part of ecology is about the assembly of inorganic resources by green plants and the reassembly of these packages at each successive stage in a web of consumer–resource inter-

actions. In this chapter we start with the resources of plants and focus especially on those most important in photosynthesis: radiation and CO_2. Together, plant resources fuel the growth of individual plants, which, collectively, determine the *primary productivity* of whole areas of land (or volumes of water): the rate, per unit area, at which plants produce biomass. Patterns of primary productivity are examined in Chapter 17. Relatively little space in this chapter is given to food as a resource for animals, simply because a series of later chapters (9–12) is devoted to the ecology of predators, grazers, parasites and saprotrophs (the consumers and decomposers of dead organisms). This chapter then closes where the previous chapter began: with the ecological niche, adding resource dimensions to the condition dimensions we have met already.

3.2 Radiation

Solar radiation is the only source of energy that can be used in metabolic activities by green plants. It comes to the plant as a flux of radiation from the sun, either directly having been diffused to a greater or lesser extent by the atmosphere, or after being reflected or transmitted by other objects. The direct fraction is highest at low latitudes (Figure 3.1). Moreover, for much of the year in temperate climates, and for the whole of the year in arid climates, the leaf canopy in terrestrial communities does not cover the land surface, so that most of the incident radiation falls on bare branches or on bare ground.

the fate of radiation

When a plant intercepts radiant energy it may be reflected (with its wavelength unchanged), transmitted (after some wavebands have been filtered out) or absorbed. Part of the fraction that is absorbed may raise the plant's temperature and be reradiated at much longer wavelengths; in terrestrial plants, part may contribute latent heat of evaporation of water and so power the transpiration

Figure 3.1 Global map of the solar radiation absorbed annually in the earth–atmosphere system: from data obtained with a radiometer on the Nimbus 3 meteorological satellite. The units are J cm^{-2} min^{-1}. (After Raushke *et al.*, 1973.)
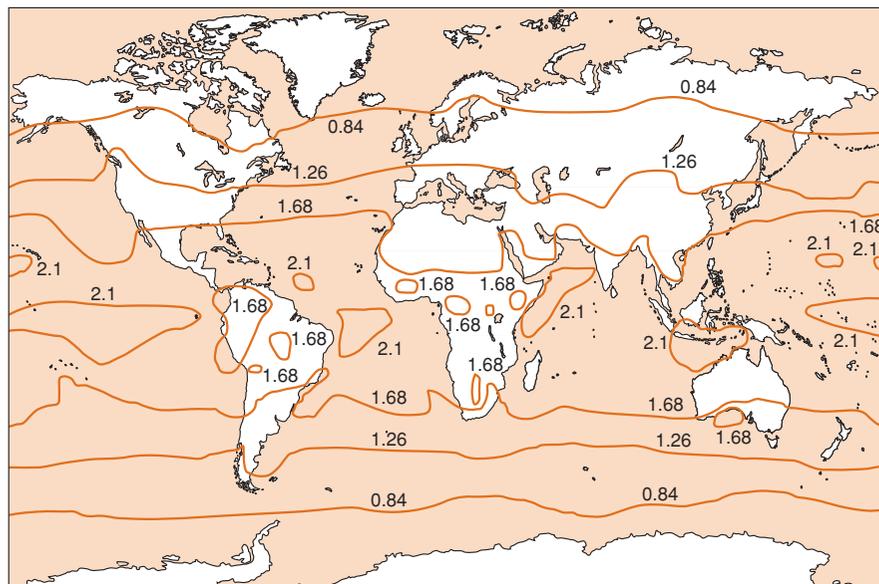
stream. A small part may reach the chloroplasts and drive the process of photosynthesis (Figure 3.2).

radiant energy must be captured or is lost forever

Radiant energy is converted during photosynthesis into energy-rich chemical compounds of carbon, which will subsequently be broken down in respiration (either by the plant itself or by organisms that consume it). But unless the radiation is captured and chemically fixed at the instant it falls on the leaf, it is irretrievably lost for photosynthesis. Radiant energy that has been fixed in photosynthesis passes just once through the world. This is in complete contrast to an atom of nitrogen or carbon or a molecule of water that may cycle repeatedly through endless generations of organisms.

photosynthetically active radiation

Solar radiation is a resource continuum: a spectrum of different wavelengths. But the photosynthetic apparatus is able to gain access to energy in only a restricted band of this spectrum. All green plants depend on chlorophyll and other pigments for the photosynthetic fixation of carbon, and these pigments fix radiation in a waveband between roughly 400 and 700 nm. This is the band of 'photosynthetically active radiation' (PAR). It corresponds broadly with the range of the spectrum visible to the human eye that we call 'light'. About 56% of the radiation incident on the earth's surface lies outside the PAR range and is thus unavailable as a resource for green plants. In other organisms there are pigments, for example bacteriochlorophyll in bacteria, that operate in photosynthesis outside the PAR range of green plants.

3.2.1 Variations in the intensity and quality of radiation

A major reason why plants seldom achieve their intrinsic photosynthetic capacity is that the intensity of radiation varies continually (Figure 3.3). Plant

photoinhibition at high intensities

morphology and physiology that are optimal for photosynthesis at one intensity of radiation will usually be inappropriate at another. In terrestrial habitats, leaves live in a radiation regime that varies throughout the day and the year, and they live in an environment of other leaves that modifies the quantity and quality of radiation received. As with all resources, the supply of radiation can vary both systematically (diurnal, annual) and unsystematically. Moreover, it is not the case simply that the intensity of radiation is a greater or lesser proportion of a maximum value at which photosynthesis would be most productive. At high intensities, *photoinhibition* of photosynthesis may occur (Long *et al.*, 1994), such that the rate of fixation of carbon decreases with increasing radiation intensity. High intensities of radiation may also lead to dangerous overheating of plants. Radiation is an essential resource for plants, but they can have too much as well as too little.

Annual and diurnal rhythms are systematic variations in solar radiation (Figure 3.3a, b). The green plant expe-

systematic variations in supply

riences periods of famine and glut in its radiation resource every 24 h (except near the poles) and seasons of famine and glut every year (except in the tropics). In aquatic habitats, an additional
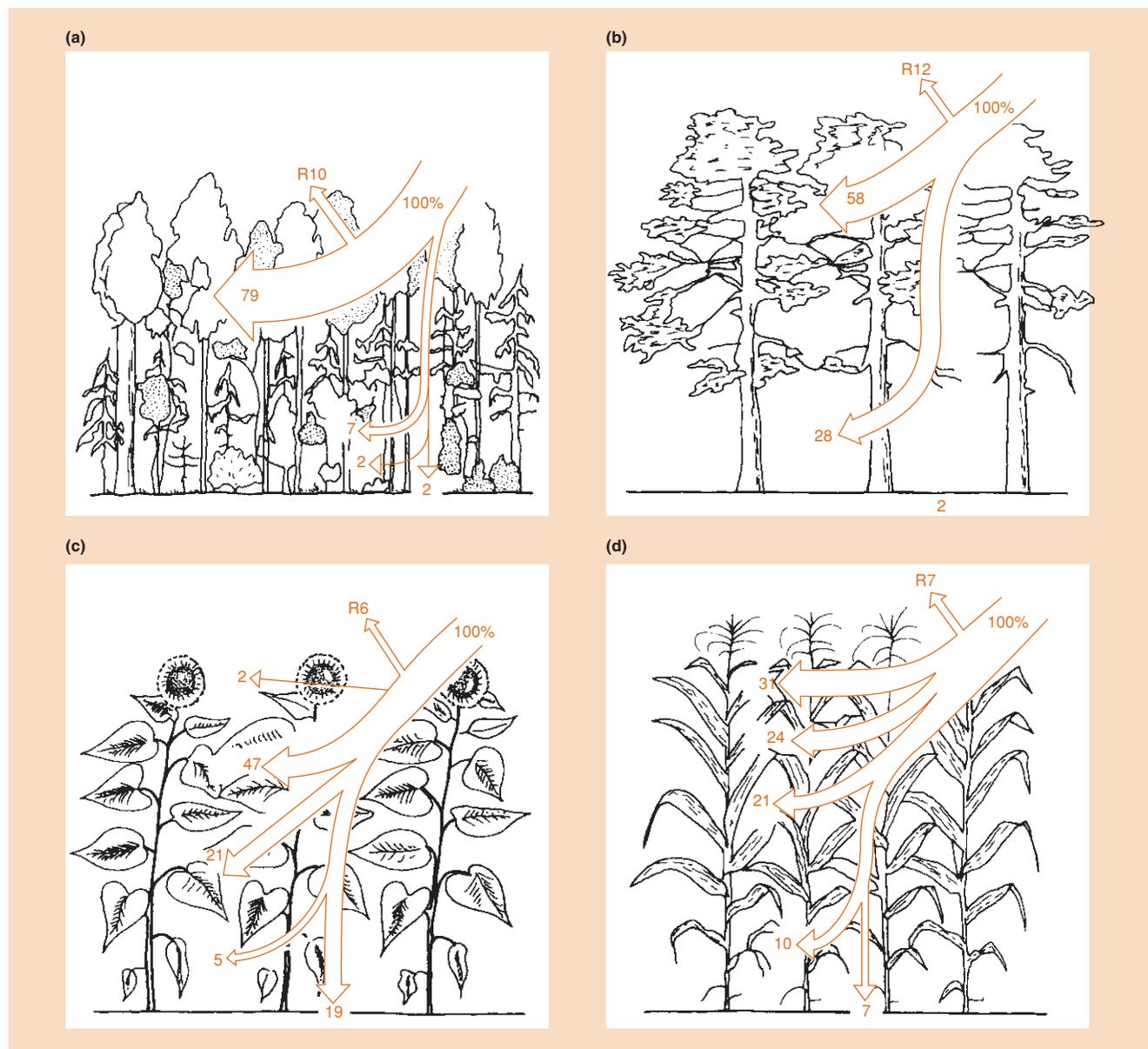
Figure 3.2 The reflection (R) and attenuation of solar radiation falling on various plant communities. The arrows show the percentage of incident radiation reaching various levels in the vegetation. (a) A boreal forest of mixed birch and spruce; (b) a pine forest; (c) a field of sunflowers; and (d) a field of corn (maize). These figures represent data obtained in particular communities and great variation will occur depending on the stage of growth of the forest or crop canopy, and on the time of day and season at which the measurements are taken. (After Larcher, 1980, and other sources.)

systematic and predictable source of variation in radiation intensity is the reduction in intensity with depth in the water column (Figure 3.3c), though the extent of this may vary greatly. For example, differences in water clarity mean that seagrasses may grow on solid substrates as much as 90 m below the surface in the relatively unproductive open ocean, whereas macrophytes in fresh waters rarely grow at depths below 10 m (Sorrell *et al.*, 2001), and

often only at considerably shallower locations, in large part because of differences in concentrations of suspended particles and also phytoplankton (see below).

The way in which an organism reacts to systematic, predictable variation in the supply of a resource reflects both its present physiology and its past evolution. The seasonal shedding of leaves by deciduous trees in temperate regions in part reflects the annual
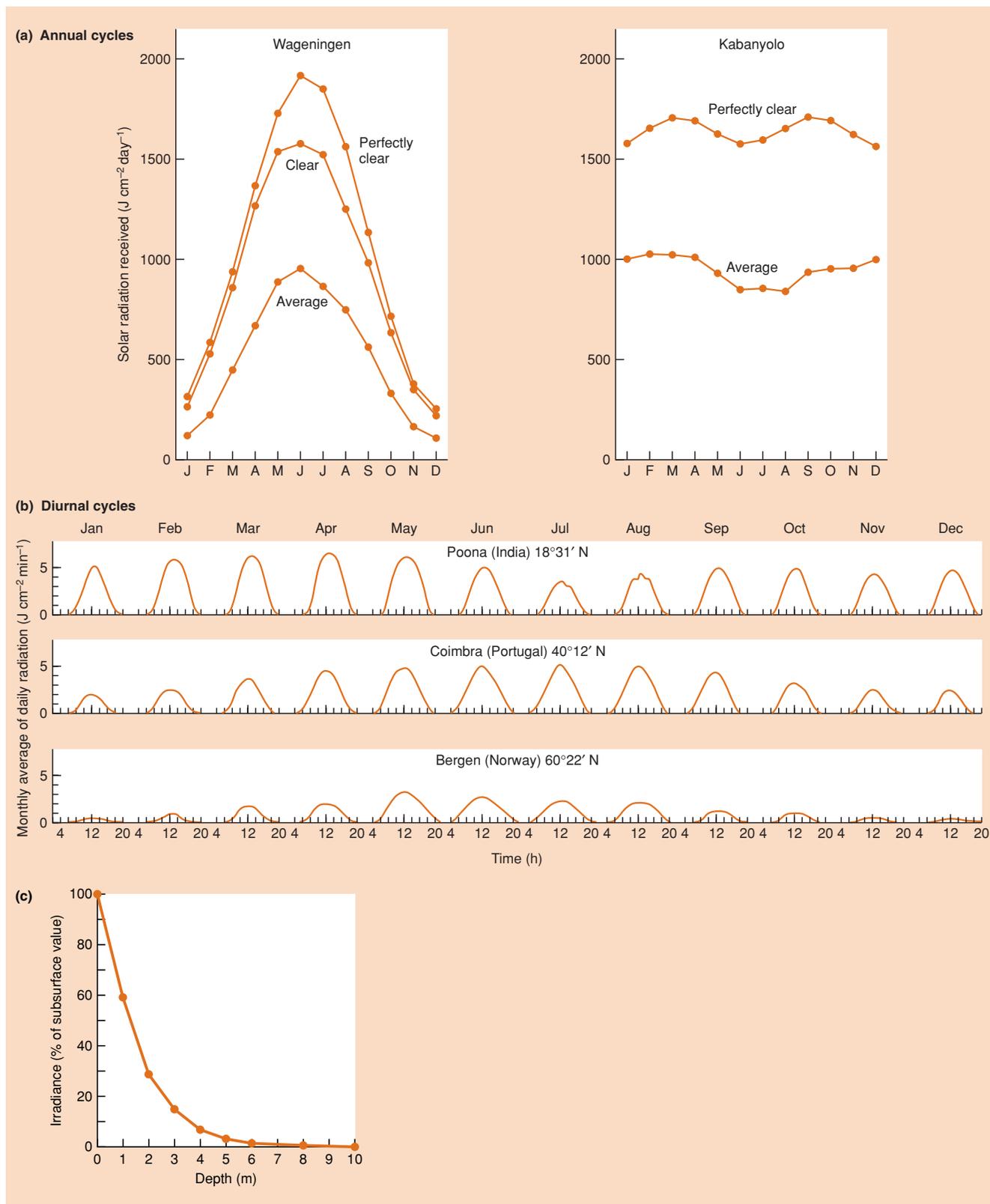
**Figure 3.3** (a) The daily totals of solar radiation received throughout the year at Wageningen (the Netherlands) and Kabanyolo (Uganda). (b) The monthly average of daily radiation recorded at Poona (India), Coimbra (Portugal) and Bergen (Norway). ((a, b) after de Wit, 1965, and other sources.) (c) Exponential diminution of radiation intensity in a freshwater habitat (Burrinjuck Dam, Australia). (After Kirk, 1994.)
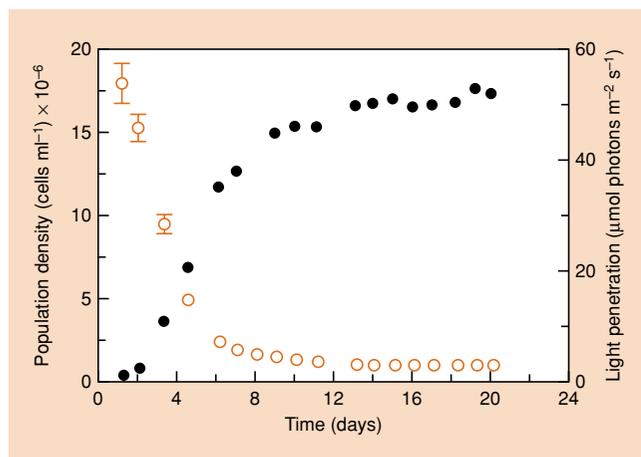
Figure 3.4   As population density (●) of the unicellular green alga, *Chlorella vulgaris*, increased in laboratory culture, this increased density reduced the penetration of light (○; its intensity at a set depth). Bars are standard deviations; they are omitted when they are smaller than the symbols. (After Huisman, 1999.)

Figure 3.5   Changing spectral distribution of radiation with depth in Lake Burley Griffin, Australia. Note that photosynthetically active radiation lies broadly within the range 400–700 nm. (After Kirk, 1994.)

rhythm in the intensity of radiation – they are shed when they are least useful. In consequence, an evergreen leaf of an understory species may experience a further systematic change, because the seasonal cycle of leaf production of overstory species determines what radiation remains to penetrate to the understory. The daily movement of leaves in many species also reflects the changing intensity and direction of incident radiation.

shade: a resource-depletion zone

Less systematic variations in the radiation environment of a leaf are caused by the nature and position of neighboring leaves. Each canopy, each plant and each leaf, by intercepting radiation, creates a resource-depletion zone (RDZ) – a moving band of shadow over other leaves of the same plant, or of others. Deep in a canopy, shadows become less well defined because much of the radiation loses its original direction by diffusion and reflection.

attenuation with depth, and plankton density, in aquatic habitats

Submerged vegetation in aquatic habitats is likely to have a much less systematic shading effect, simply because it is moved around by the flow of the water in which it lives, though vegetation floating on the surface, especially of ponds or lake, inevitably has a profound and largely unvarying effect on the radiation regime beneath it. Phytoplankton cells nearer the surface, too, shade the cells beneath them, such that the reduction of intensity with depth is greater, the greater the phytoplankton density. Figure 3.4, for example, shows the decline in light penetration, measured at a set depth in a laboratory system, as a population of the unicellular green alga, *Chlorella vulgaris*, built up over a 12-day period (Huisman, 1999).
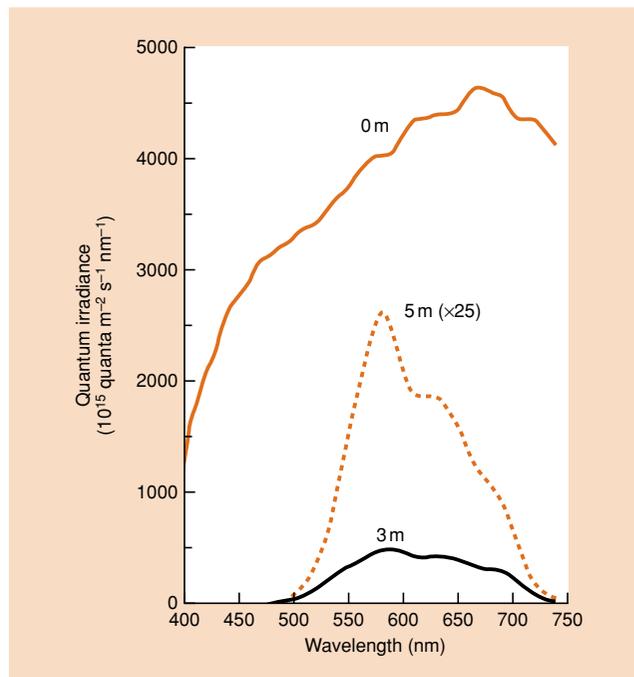
The composition of radiation that has passed through leaves in a canopy, or through a body of water, is also altered. It may be less useful photosynthetically because the PAR component has been reduced – though such reductions may also, of course, prevent photoinhibition and overheating. Figure 3.5 shows an example for the variation with depth in a freshwater habitat.

variations in quality as well as quantity

The major differences amongst terrestrial species in their reaction to systematic variations in the intensity of radiation are those that have evolved between 'sun species' and 'shade species'. In general, plant species that are characteristic of shaded habitats use radiation at low intensities more efficiently than sun species, but the reverse is true at high intensities (Figure 3.6). Part of the difference between them lies in the physiology of the leaves, but the morphology of the plants also influences the efficiency with which radiation is captured. The leaves of sun plants are commonly exposed at acute angles to the midday sun (Poulson & DeLucia, 1993). This spreads an incident beam of radiation over a larger leaf area, and effectively reduces its intensity. An intensity of radiation that is superoptimal for photosynthesis when it strikes a leaf at 90° may therefore be optimal for a leaf inclined at an acute angle. The leaves of sun plants are often superimposed into

sun and shade species
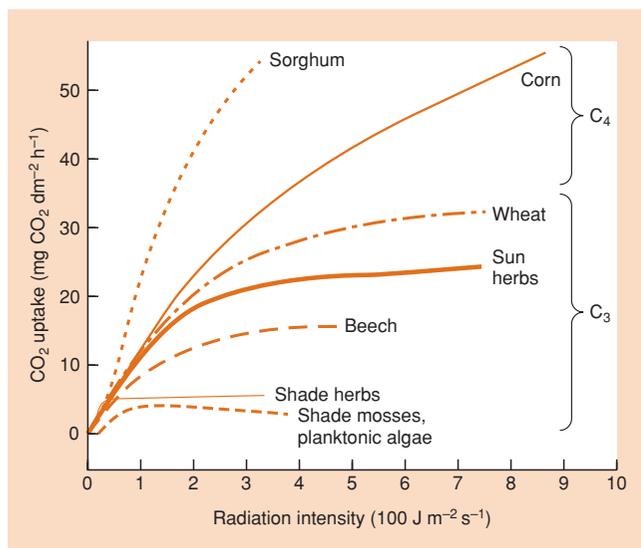
Figure 3.6 The response of photosynthesis to light intensity in various plants at optimal temperatures and with a natural supply of CO_2. Note that corn and sorghum are C_4 plants and the remainder are C_3 (the terms are explained in Sections 3.3.1 and 3.3.2). (After Larcher, 1980, and other sources.)

a multilayered canopy. In bright sunshine even the shaded leaves in lower layers may have positive rates of net photosynthesis. Shade plants commonly have leaves held near to the horizontal and in a single-layered canopy.

sun and shade leaves

In contrast to these 'strategic' differences, it may also happen that as a plant grows, its leaves develop differently as a 'tactical' response to the radiation environment in which it developed. This often leads to the formation of 'sun leaves' and 'shade leaves' within the canopy of a single plant. Sun leaves are typically smaller, thicker, have more cells per unit area, denser veins, more densely packed chloroplasts and a greater dry weight per unit area of leaf. These tactical maneuvers, then, tend to occur not at the level of the whole plant, but at the level of the individual leaf or even its parts. Nevertheless, they take time. To form sun or shade leaves as a tactical response, the plant, its bud or the developing leaf must sense the leaf's environment and respond by growing a leaf with an appropriate structure. For example, it is impossible for the plant to change its form fast enough to track the changes in intensity of radiation between a cloudy and a clear day. It can, however, change its rate of photosynthesis extremely rapidly, reacting even to the passing of a fleck of sunlight. The rate at which a leaf photosynthesizes also depends on the demands that are made on it by other vigorously growing parts. Photosynthesis may be reduced, even though conditions are otherwise ideal, if there is no demanding call on its products.

In aquatic habitats, much of the variation between species is accounted for by differences in photosynthetic pigments, which contribute significantly to the precise wavelengths of radiation that can be utilized (Kirk, 1994). Of the three types of pigment – chlorophylls, carotenoids and biliproteins – all photosynthetic plants contain the first two, but many algae also contain biliproteins; and within the chlorophylls, all higher plants have chlorophyll *a* and *b*, but many algae have only chlorophyll *a* and some have chlorophyll *a* and *c*. Examples of the absorption spectra of a number of pigments, the related contrasting absorption spectra of a number of groups of aquatic plants, and the related distributional differences (with depth) between a number of groups of aquatic plants are illustrated in Figure 3.7. A detailed assessment of the evidence for direct links between pigments, performance and distribution is given by Kirk (1994).

pigment variation in aquatic species

3.2.2 Net photosynthesis

The rate of photosynthesis is a gross measure of the rate at which a plant captures radiant energy and fixes it in organic carbon compounds. However, it is often more important to consider, and very much easier to measure, the net gain. Net photosynthesis is the increase (or decrease) in dry matter that results from the difference between gross photosynthesis and the losses due to respiration and the death of plant parts (Figure 3.8).

Net photosynthesis is negative in darkness, when respiration exceeds photosynthesis, and increases with the intensity of PAR. The *compensation point* is the intensity of PAR at which the gain from gross photosynthesis exactly balances the respiratory and other losses. The leaves of shade species tend to respire at lower rates than those of sun species. Thus, when both are growing in the shade the net photosynthesis of shade species is greater than that of sun species.

the compensation point

There is nearly a 100-fold variation in the *photosynthetic capacity* of leaves (Mooney & Gulmon, 1979). This is the rate of photosynthesis when incident radiation is saturating, temperature is optimal, relative humidity is high, and CO_2 and oxygen concentrations are normal. When the leaves of different species are compared under these ideal conditions, the ones with the highest photosynthetic capacity are generally those from environments where nutrients, water and radiation are seldom limiting (at least during the growing season). These include many agricultural crops and their weeds. Species from resource-poor environments (e.g. shade plants, desert perennials, heathland species) usually have low photosynthetic capacity – even when abundant resources are provided. Such patterns can be understood by noting that photosynthetic capacity, like all capacity, must be 'built'; and the investment in building

photosynthetic capacity
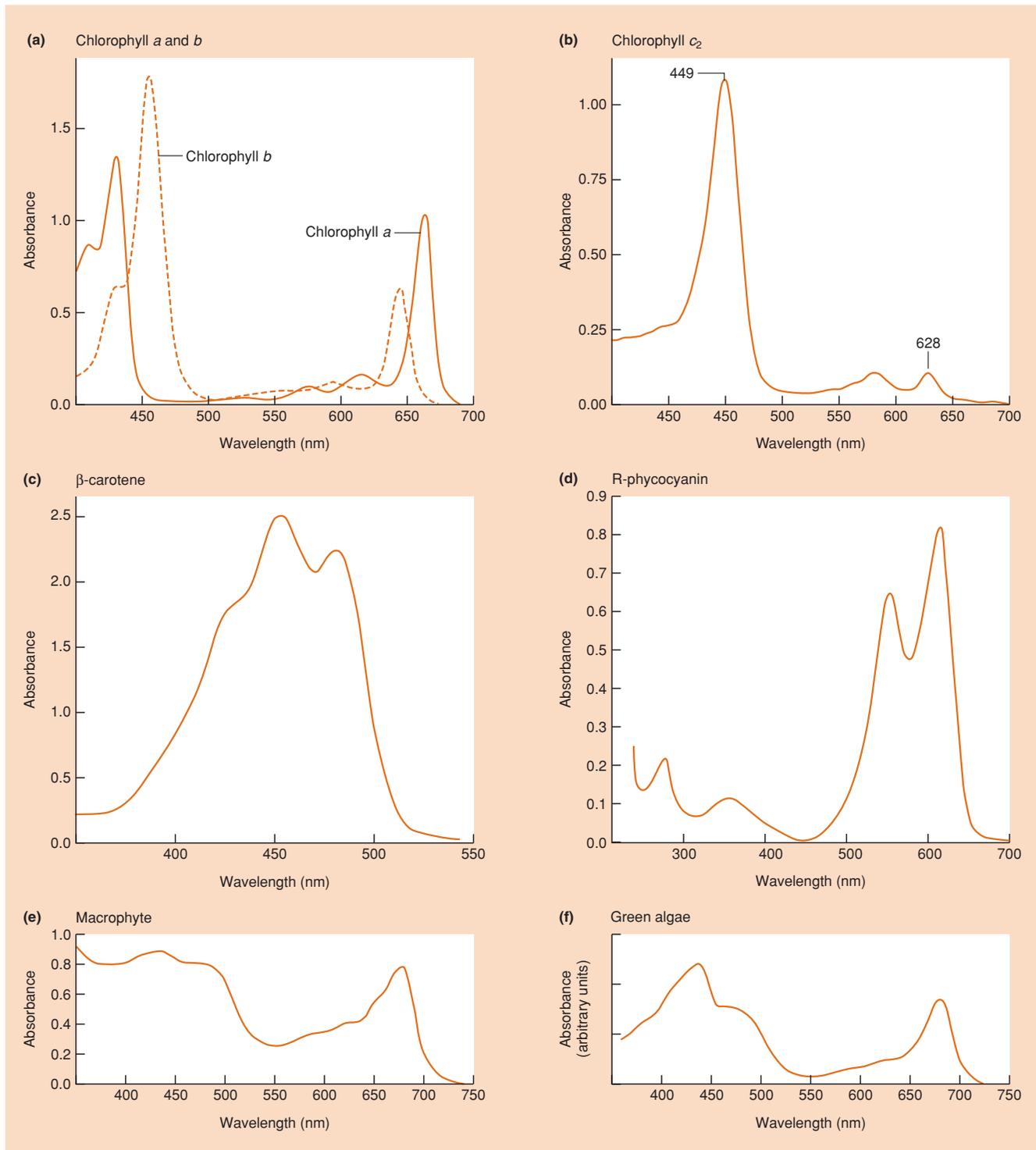
**Figure 3.7**   (a) Absorption spectra of chlorophylls *a* and *b*. (b) Absorption spectrum of chlorophyll *c₂*. (c) Absorption spectrum of β-carotene. (d) Absorption spectrum of the biliprotein, R-phycocyanin. (e) Absorption spectrum of a piece of leaf of the freshwater macrophyte, *Vallisneria spiralis*, from Lake Ginnindera, Australia. (f) Absorption spectrum of the planktonic alga *Chlorella pyrenoidos* (green).
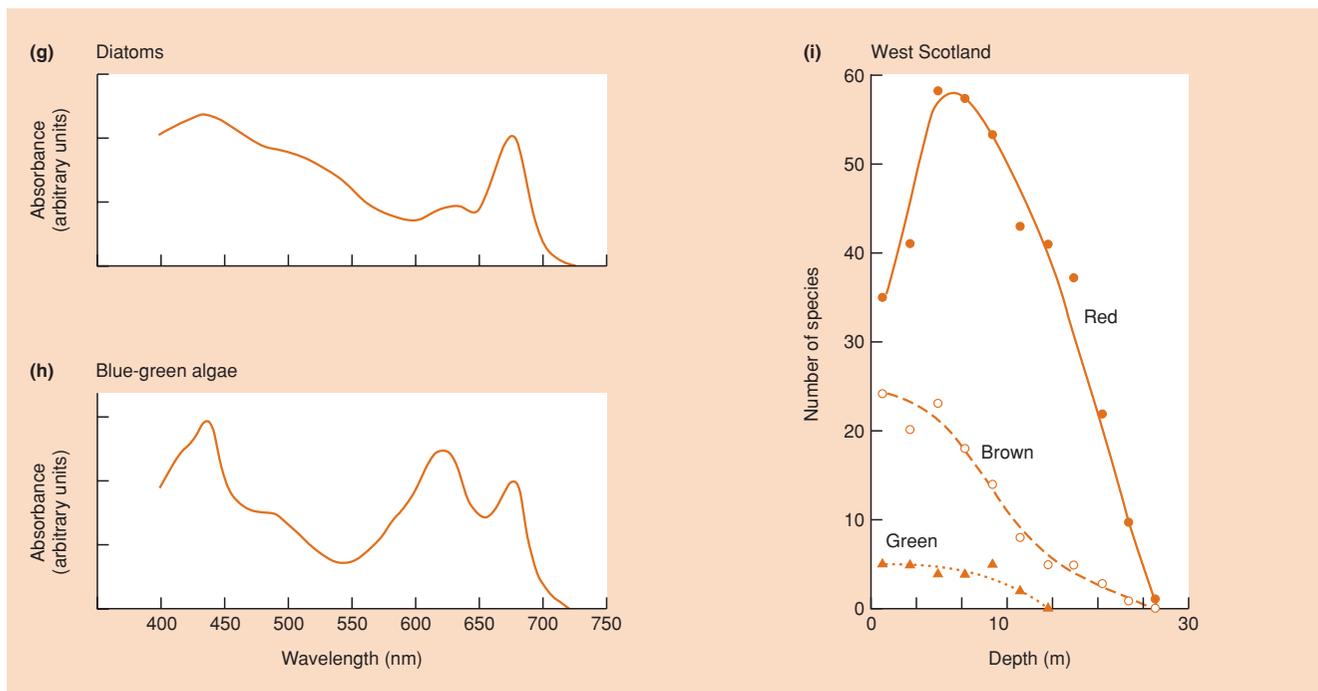
Figure 3.7 *(continued)* (g–h) Absorption spectra of the planktonic algae *Navicula minima* (diatom) and *Synechocystis* sp. (blue-green). (i) The numbers of species of benthic red, green and brown algae at various depths (and in various light regimes) off the west coast of Scotland (56–57°N). (After Kirk, 1994; data from various sources.)
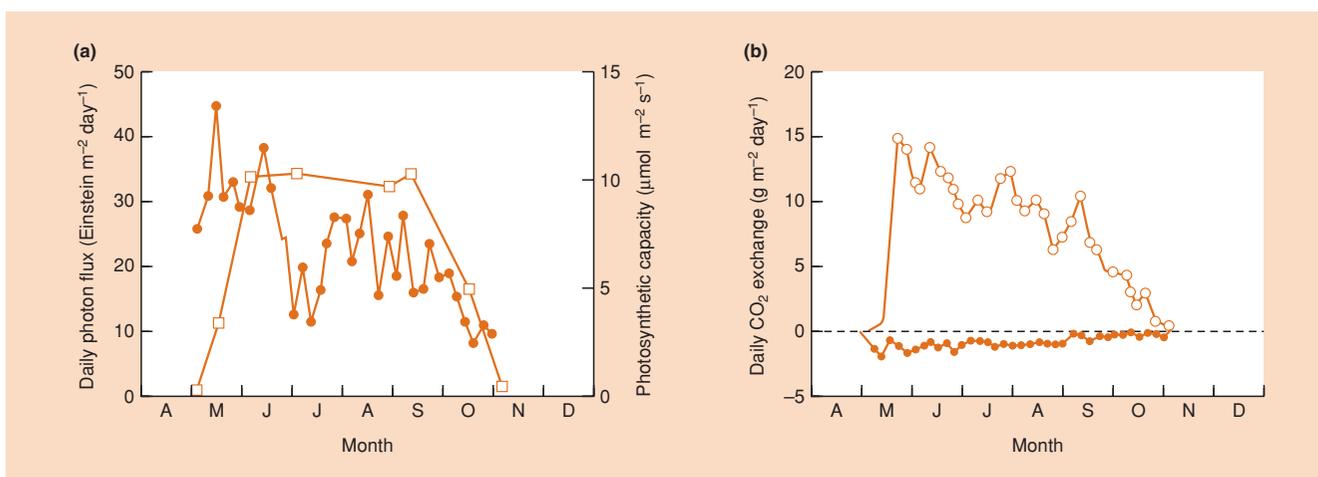
Figure 3.8 The annual course of events that determined the net photosynthetic rate of the foliage of maple (*Acer campestre*) in 1980. (a) Variations in the intensity of PAR (●), and changes in the photosynthetic capacity of the foliage (□) appearing in spring, rising to a plateau and then declining through late September and October. (b) The daily fixation of carbon dioxide (CO_2) (○) and its loss through respiration during the night (●). The annual total gross photosynthesis was 1342 g CO_2 m^{-2} and night respiration was 150 g CO_2 m^{-2}, giving a balance of 1192 g CO_2 m^{-2} net photosynthesis. (After Pearcy *et al.*, 1987.)

capacity is only likely to be repaid if ample opportunity exists for that capacity to be utilized.

Needless to say, ideal conditions in which plants may achieve their photosynthetic capacity are rarely present outside a physiologist's controlled environment chamber. In practice, the rate at which photosynthesis actually proceeds is limited by conditions (e.g. temperature) and by the availability of resources other than radiant energy. Leaves seem also to achieve their maximal photosynthetic rate only when the products are being actively withdrawn (to developing buds, tubers, etc.). In addition, the photosynthetic capacity of leaves is highly correlated with leaf nitrogen content, both between leaves on a single plant and between the leaves of different species (Woodward, 1994). Around 75% of leaf nitrogen is invested in chloroplasts. This suggests that the availability of nitrogen as a resource may place strict limits on the ability of plants to garner CO_2 and energy in photosynthesis. The rate of photosynthesis also increases with the intensity of PAR, but in most species ('C$_3$ plants' – see below) reaches a plateau at intensities of radiation well below that of full solar radiation.

The highest efficiency of utilization of radiation by green plants is 3–4.5%, obtained from cultured microalgae at low intensities of PAR. In tropical forests values fall within the range 1–3%, and in temperate forests 0.6–1.2%. The approximate efficiency of temperate crops is only about 0.6%. It is on such levels of efficiency that the energetics of all communities depend.

3.2.3 Sun and shade plants of an evergreen shrub

A number of the general points above are illustrated by a study of the evergreen shrub, *Heteromeles arbutifolia*. This plant grows both in chaparral habitats in California, where shoots in the upper crown are consistently exposed to full sunlight and high temperatures, especially during the dry season, and also in woodland habitats, where the plant grows both in open sites and in the shaded understory (Valladares & Pearcy, 1998). Shade plants from the understory were compared with sun plants from the chaparral, where they received around seven times as much radiation ('photon flux density', PFD). Compared to those from the shade (Figure 3.9 and Table 3.1a), sun plants had leaves that were inclined at a much steeper angle to the horizontal, were smaller but thicker, and were borne on shoots that were themselves shorter (smaller internode distances). The sun leaves also had a greater photosynthetic capacity (more chlorophyll and nitrogen) per unit leaf area but not per unit biomass.

The 'architectural' consequences of these differences (Table 3.1b) were first that shade plants had a much greater 'projection efficiency' in the summer, but a much lower efficiency in the winter. Projection efficiency expresses the degree to which the effective leaf area is reduced by being borne at an angle other than right angles to the incident radiation. Thus, the more angled leaves of sun plants absorbed the direct rays of the overhead summer
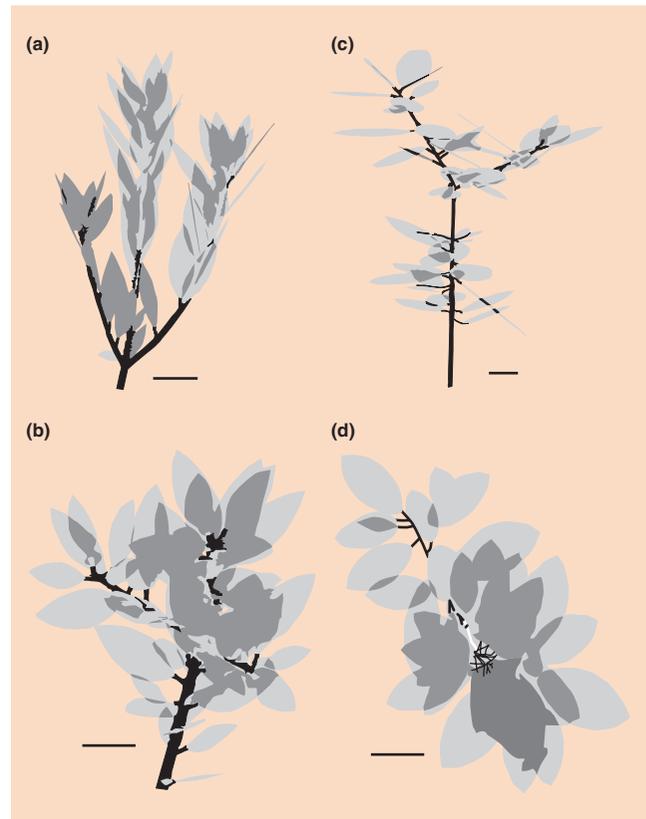
Figure 3.9 Computer reconstructions of stems of typical sun (a, c) and shade (b, d) plants of the evergreen shrub *Heteromeles arbutifolia*, viewed along the path of the sun's rays in the early morning (a, b) and at midday (c, d). Darker tones represent parts of leaves shaded by other leaves of the same plant. Bars = 4 cm. (After Valladares & Pearcy, 1998.)

sun over a wider leaf area than the more horizontal shade plant leaves, but the more sideways rays of the winter sun struck the sun plant leaves at closer to a right angle. Furthermore, these projection efficiencies can themselves be modified by the fraction of leaf area subject to self-shading, giving rise to 'display efficiencies'. These were higher in shade than in sun plants, in the summer because of the higher projection efficiency, but in the winter because of the relative absence of self-shading in shade plants.

Whole plant physiological properties (Table 3.1b), then, reflect both plant architecture and the morphologies and physiologies of individual leaves. The efficiency of light absorption, like display efficiency, reflects both leaf angles and self-shading. Hence, absorption efficiency was consistently higher for shade than for sun plants, though the efficiency for sun plants was significantly higher in winter compared to summer. The effective leaf ratio (the light absorption efficiency per unit of biomass) was then massively greater for shade than for sun plants (as a result of their thinner leaves), though again, somewhat higher for the latter in winter.

Table 3.1 (a) Observed differences in the shoots and leaves of sun and shade plants of the shrub *Heteromeles arbutifolia*. Standard deviations are given in parentheses; the significance of differences are given following analysis of variance. (b) Consequent whole plant properties of sun and shade plants. (After Valladares & Pearcy, 1998.)

(a)

	Sun		Shade		P
Internode distance (cm)	1.08	(0.06)	1.65	(0.02)	< 0.05
Leaf angle (degrees)	71.3	(16.3)	5.3	(4.3)	< 0.01
Leaf surface area (cm^2)	10.1	(0.3)	21.4	(0.8)	< 0.01
Leaf blade thickness (µm)	462.5	(10.9)	292.4	(9.5)	< 0.01
Photosynthetic capacity, area basis (µmol CO_2 m^{-2} s^{-1})	14.1	(2.0)	9.0	(1.7)	< 0.01
Photosynthetic capacity, mass basis (µmol CO_2 kg^{-1} s^{-1})	60.8	(10.1)	58.1	(11.2)	NS
Chlorophyll content, area basis (mg m^{-2})	280.5	(15.3)	226.7	(14.0)	< 0.01
Chlorophyll content, mass basis (mg g^{-1})	1.23	(0.04)	1.49	(0.03)	< 0.05
Leaf nitrogen content, area basis (g m^{-2})	1.97	(0.25)	1.71	(0.21)	< 0.05
Leaf nitrogen content, mass basis (% dry weight)	0.91	(0.31)	0.96	(0.30)	NS

(b)

	Sun plants		Shade plants	
	Summer	Winter	Summer	Winter
E_P	0.55[a]	0.80[b]	0.88[b]	0.54[a]
E_D	0.33[a]	0.38[a, b]	0.41[b]	0.43[b]
Fraction self-shaded	0.22[a]	0.42[b]	0.47[b]	0.11[a]
$E_{A, direct PFD}$	0.28[a]	0.44[b]	0.55[c]	0.53[c]
LAR$_c$ (cm^2 g^{-1})	7.1[a]	11.7[b]	20.5[c]	19.7[c]

E_P, projection efficiency; E_D, display efficiency; E_A, absorption efficiency; LAR$_e$, effective leaf area ratio; NS, not significant. Letter codes indicate groups that differed significantly in analyses of variance ($P < 0.05$).

Overall, therefore, despite receiving only one-seventh of the PFD of sun plants, shade plants reduced the differential in the amount absorbed to one-quarter, and reduced the differential in their daily rate of carbon gain to only a half. Shade plants successfully counterbalanced their reduced photosynthetic capacity at the leaf level with enhanced light-harvesting ability at the whole plant level. The sun plants can be seen as striking a compromise between maximizing whole plant photosynthesis on the one hand while avoiding photoinhibition and overheating of individual leaves on the other.

3.2.4 Photosynthesis or water conservation? Strategic and tactical solutions

stomatal opening

In fact, in terrestrial habitats especially, it is not sensible to consider radiation as a resource independently of water. Intercepted radiation does not result in photosynthesis unless there is CO_2 available, and the prime route of entry of CO_2 is through open stomata. But if the stomata are open to the air, water will evaporate through them. If water is lost faster than it can be gained, the leaf (and the plant) will sooner or later wilt and eventually die. But in most terrestrial communities, water is, at least sometimes, in short supply. Should a plant conserve water at the expense of present photosynthesis, or maximize photosynthesis at the risk of running out of water? Once again, we meet the problem of whether the optimal solution involves a strict strategy or the ability to make tactical responses. There are good examples of both solutions and also compromises.

Perhaps the most obvious strategy that plants may adopt is to have a short life and high photosynthetic activity during periods when water is abundant, but remain dormant as seeds during the rest of the year, neither photosynthesizing nor transpiring (e.g. many desert annuals, annual weeds and most annual crop plants).

short active interludes in a dormant life

Second, plants with long lives may produce leaves during periods when water is abundant and shed them during droughts (e.g. many species of *Acacia*). Some shrubs of the Israeli desert (e.g. *Teucrium polium*) bear finely divided, thin-cuticled leaves during the season when soil water is freely available. These are then replaced by undivided, small, thick-cuticled leaves in more drought-prone seasons, which in turn fall and may leave only green spines or thorns (Orshan, 1963): a sequential polymorphism through the season, with each leaf morph being replaced in turn by a less photosynthetically active but more water-tight structure.

Next, leaves may be produced that are long lived, transpire only slowly and tolerate a water deficit, but which are unable to photosynthesize rapidly even when water is abundant (e.g. evergreen desert shrubs). Structural features such as hairs, sunken stomata and the restriction of stomata to specialized areas on the lower surface of a leaf slow down water loss. But these same morphological features reduce the rate of entry of CO_2. Waxy and hairy leaf surfaces may, however, reflect a greater proportion of radiation that is not in the PAR range and so keep the leaf temperature down and reduce water loss.

Finally, some groups of plants have evolved particular physiologies: C_4 and crassulacean acid metabolism (CAM). We consider these in more detail in Sections 3.3.1–3.3.3. Here, we simply note that plants with 'normal' (i.e. C_3) photosynthesis are wasteful of water compared with plants that possess the modified C_4 and CAM physiologies. The water-use efficiency of C_4 plants (the amount of carbon fixed per unit of water transpired) may be double that of C_3 plants.

The viability of alternative strategies to solve a common problem is nicely illustrated by the trees of seasonally dry tropical forests and woodlands (Eamus, 1999). These communities are found naturally in Africa, the Americas, Australia and India, and as a result of human interference elsewhere in Asia. But whereas, for example, the savannas of Africa and India are dominated by deciduous species, and the Llanos of South America are dominated by evergreens, the savannas of Australia are occupied by roughly equal numbers of species from four groups (Figure 3.10a): evergreens (a full canopy all year), deciduous species (losing all leaves for at least 1 and usually 2–4 months each year), semideciduous species (losing around 50% or more of their leaves each year) and brevideciduous species (losing only about 20% of their leaves). At the ends of this continuum, the deciduous species avoid drought in the dry season (April–November in Australia) as a result of their vastly reduced rates of transpiration (Figure 3.10b), but the evergreens maintain a positive carbon balance throughout the year (Figure 3.10c), whereas the deciduous species make no net photosynthate at all for around 3 months.

The major tactical control of the rates of both photosynthesis and water loss is through changes in stomatal 'conductance' that may occur rapidly during the course of a day and allow a very rapid response to immediate water shortages. Rhythms of stomatal opening and closure may ensure that the above-ground parts of the plant remain more or less watertight except during controlled periods of active photosynthesis. These rhythms may

Figure 3.10   (a) Percentage canopy fullness for deciduous (♦), semideciduous (■), brevideciduous (▼) and evergreen (●) trees Australian savannas throughout the year. (Note that the southern hemisphere dry season runs from around April to November.) (b) Susceptibility to drought as measured by increasingly negative values of 'predawn water potential' for deciduous (♦) and evergreen (●) trees. (c) Net photosynthesis as measured by the carbon assimilation rate for deciduous (♦) and evergreen (●) trees. (After Eamus, 1999.)

be diurnal or may be quickly responsive to the plant's internal water status. Stomatal movement may even be triggered directly by conditions at the leaf surface itself – the plant then responds to desiccating conditions at the very site, and at the same time, as the conditions are first sensed.

3.3   Carbon dioxide

the rise in global levels

The CO_2 used in photosynthesis is obtained almost entirely from the atmosphere, where its concentration has risen from approximately 280 $\mu l\, l^{-1}$ in 1750 to about 370 $\mu l\, l^{-1}$ today and is still increasing by 0.4–0.5% year^{-1} (see Figure 18.22). In a terrestrial community, the flux of CO_2 at night is upwards, from the soil and vegetation to the atmosphere; on sunny days above a photosynthesizing canopy, there is a downward flux.

variations beneath a canopy

Above a vegetation canopy, the air becomes rapidly mixed. However, the situation is quite different within and beneath canopies. Changes in CO_2 concentration in the air within a mixed deciduous forest in New England were measured at various heights above ground level during the year (Figure 3.11a) (Bazzaz & Williams, 1991). Highest concentrations, up to around 1800 $\mu l\, l^{-1}$, were measured near the

surface of the ground, tapering off to around 400 $\mu l\, l^{-1}$ at 1 m above the ground. These high values near ground level were achieved in the summer when high temperatures allowed the rapid decomposition of litter and soil organic matter. At greater heights within the forest, the CO_2 concentrations scarcely ever (even in winter) reached the value of 370 $\mu l\, l^{-1}$ which is the atmospheric concentration of bulk air measured at the Mauna Loa laboratory in Hawaii (see Figure 18.22). During the winter months, concentrations remained virtually constant through the day and night at all heights. But in the summer, major diurnal cycles of concentration developed that reflected the interaction between the production of CO_2 by decomposition and its consumption in photosynthesis (Figure 3.11b).

That CO_2 concentrations vary so widely within vegetation means that plants growing in different parts of a forest will experience quite different CO_2 environments. Indeed the lower leaves on a forest shrub will usually experience higher CO_2 concentrations than its upper leaves, and seedlings will live in environments richer in CO_2 than mature trees.

In aquatic environments, variations in CO_2 concentration can be just as striking, especially when water mixing is limited, for example during the sum-

variations in aquatic habitats . . .

mer 'stratification' of lakes, with layers of warm water towards the surface and colder layers beneath (Figure 3.12).
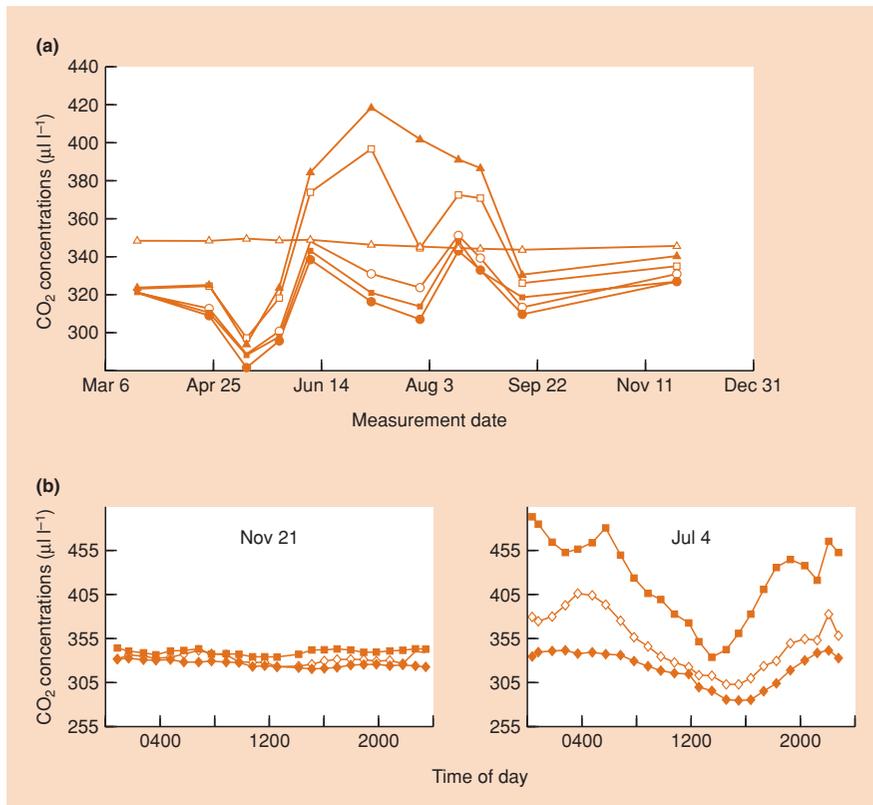
Figure 3.11   (a) CO_2 concentrations in a mixed deciduous forest (Harvard Forest, Massachusetts, USA) at various times of year at five heights above ground: ▲, 0.05 m; □, 0.20 m; ■, 3.00 m; ○, 6.00 m; ●, 12.00 m. Data from the Mauna Loa CO_2 observatory (△) are given on the same axis for comparison. (b) CO_2 concentrations for each hour of the day (averaged over 3–7-day periods) on November 21 and July 4. (After Bazzaz & Williams, 1991.)
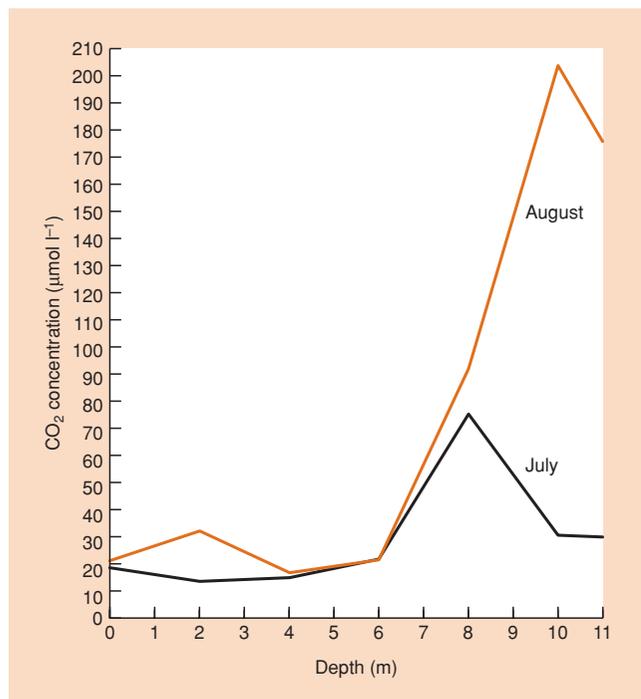
Figure 3.12 Variation in CO_2 concentration with depth in Lake Grane Langsø, Denmark in early July and again in late August after the lake becomes stratified with little mixing between the warm water at the surface and the colder water beneath. (After Riis & Sand-Jensen, 1997.)


Figure 3.13 The increase (to a plateau) in photosynthetic rate with artificially manipulated CO_2 concentrations in moss, *Sphagnum subsecundum*, taken from depths of 9.5 m (●) and 0.7 m (○) in Lake Grane Langsø, Denmark, in early July. These concentrations – and hence the rates of photosynthesis – are much higher than those occurring naturally (see Figure 3.12). (After Riis & Sand-Jensen, 1997.)

. . . setting a limit on photosynthetic rates

Also, in aquatic habitats, dissolved CO_2 tends to react with water to form carbonic acid, which in turn ionizes, and these tendencies increase with pH, such that 50% or more of inorganic carbon in water may be in the form of bicarbonate ions. Many aquatic plants can utilize carbon in this form, but since it must ultimately be reconverted to CO_2 for photosynthesis, this is likely to be less useful as a source of inorganic carbon, and in practice, many plants will be limited in their photosynthetic rate by the availability of CO_2. Figure 3.13, for example, shows the response of the moss, *Sphagnum subsecundum*, taken from two depths in a Danish lake, to increases in CO_2 concentration. At the time they were sampled (July 1995), the natural concentrations in the waters from which they were taken (Figure 3.12) were 5–10 times less than those eliciting maximum rates of photosynthesis. Even the much higher concentrations that occurred at the lower depths during summer stratification would not have maximized photosynthetic rate.

One might expect a process as fundamental to life on earth as carbon fixation in photosynthesis to be underpinned by a single unique biochemical pathway. In fact, there are three such pathways (and variants within them): the C_3 pathway (the most common), the C_4 pathway and CAM (crassulacean acid metabolism). The ecological consequences of the different pathways are profound, especially as they affect the reconciliation of photosynthetic activity and controlled water loss (see Section 3.2.4). Even in aquatic plants, where water conservation is not normally an issue, and most plants use the C_3 pathway, there are many CO_2-concentrating mechanisms that serve to enhance the effectiveness of CO_2 utilization (Badger *et al.*, 1997).

3.3.1 The C_3 pathway

In this, the Calvin–Benson cycle, CO_2 is fixed into a three-carbon acid (phosphoglyceric acid) by the enzyme Rubisco, which is present in massive amounts in the leaves (25–30% of the total leaf nitrogen). This same enzyme can also act as an oxygenase, and this activity (photorespiration) can result in a wasteful release of CO_2 – reducing by about one-third the net amounts of CO_2 that are fixed. Photorespiration increases with temperature with the consequence that the overall efficiency of carbon fixation declines with increasing temperature.

The rate of photosynthesis of C_3 plants increases with the intensity of radiation, but reaches a plateau. In many species, particularly shade species, this plateau occurs at radiation intensities far below that of full solar radiation (see Figure 3.6). Plants with C_3 metabolism have low water-use efficiency compared with C_4 and CAM plants (see below), mainly because in a C_3 plant, CO_2 diffuses rather slowly into the leaf and so allows time for a lot of water vapor to diffuse out of it.

3.3.2   The C_4 pathway

In this, the Hatch–Slack cycle, the C_3 pathway is present but it is confined to cells deep in the body of the leaf. CO_2 that diffuses into the leaves via the stomata meets mesophyll cells containing the enzyme phosphoenolpyruvate (PEP) carboxylase. This enzyme combines atmospheric CO_2 with PEP to produce a four-carbon acid. This diffuses, and releases CO_2 to the inner cells where it enters the traditional C_3 pathway. PEP carboxylase has a much greater affinity than Rubisco for CO_2. There are profound consequences.

First, C_4 plants can absorb atmospheric CO_2 much more effectively than C_3 plants. As a result, C_4 plants may lose much less water per unit of carbon fixed. Furthermore, the wasteful release of CO_2 by photorespiration is almost wholly prevented and, as a consequence, the efficiency of the overall process of carbon fixation does not change with temperature. Finally, the concentration of Rubisco in the leaves is a third to a sixth of that in C_3 plants, and the leaf nitrogen content is correspondingly lower. As a consequence of this, C_4 plants are much less attractive to many herbivores and also achieve more photosynthesis per unit of nitrogen absorbed.

One might wonder how C_4 plants, with such high water-use efficiency, have failed to dominate the vegetation of the world, but there are clear costs to set against the gains. The C_4 system has a high light compensation point and is inefficient at low light intensities; C_4 species are therefore ineffective as shade plants. Moreover, C_4 plants have higher temperature optima for growth than C_3 species: most C_4 plants are found in arid regions or the tropics. In North America, C_4 dicotyledonous species appear to be favored in sites of limited water supply (Figure 3.14) (Stowe & Teeri, 1978), whereas the abundance of C_4 monocotyledonous species is strongly correlated with maximum daily temperatures during the growing season (Teeri & Stowe, 1976). But these correlations are not universal. More generally, where there are mixed populations of C_3 and C_4 plants, the proportion of C_4 species tends to fall with elevation on mountain ranges, and in seasonal climates it is C_4 species that tend to dominate the vegetation in the hot dry seasons and C_3 species in the cooler wetter seasons. The few C_4 species that extend into temperate regions (e.g. *Spartina* spp.) are found in marine or other saline environments where osmotic conditions may especially favor species with efficient water use.

Perhaps the most remarkable feature of C_4 plants is that they do not seem to use their high water-use efficiency in faster shoot growth, but instead devote a greater fraction of the plant body to a well-developed root system. This is one of the hints that the rate of carbon assimilation is not the major limit to their growth, but that the shortage of water and/or nutrients matters more.

3.3.3   The CAM pathway

Plants with a crassulacean acid metabolism (CAM) pathway also use PEP carboxylase with its strong power of concentrating CO_2. In contrast to C_3 and C_4 plants, though, they open their stomata and fix CO_2 at night (as malic acid). During the daytime the stomata are closed and the CO_2 is released within the leaf and fixed by Rubisco. However, because the CO_2 is then at a high concentration within the leaf, photorespiration is prevented, just as it is in plants using the C_4 pathway. Plants using the CAM photosynthetic pathway have obvious advantages when water is in short supply, because their stomata are closed during the daytime when evaporative forces are strongest. The system is now known in a wide variety of families, not just the Crassulaceae. This appears to be a highly effective means of water conservation, but CAM species have not come to inherit the earth. One cost to CAM plants is the problem of storing the malic acid that is formed at night: most CAM plants are succulents with extensive water-storage tissues that cope with this problem.

In general, CAM plants are found in arid environments where strict stomatal control of daytime water is vital for survival (desert succulents) and where CO_2 is in short supply during the daytime, for example in submerged aquatic plants, and in photosynthetic organs that lack stomata (e.g. the aerial photosynthetic roots of orchids). In some CAM plants, such as *Opuntia basilaris*, the stomata remain closed both day and night during drought. The CAM process then simply allows the plant to 'idle' – photosynthesizing only the CO_2 produced internally by respiration (Szarek *et al.*, 1973).

A taxonomic and systematic survey of C_3, C_4 and CAM photosynthetic systems is given by Ehleringer and Monson (1993). They describe the very strong evidence that the C_3 pathway is evolutionarily primitive and, very surprisingly, that the C_4 and CAM systems must have arisen repeatedly and independently during the evolution of the plant kingdom.

3.3.4   The response of plants to changing atmospheric concentrations of CO_2

Of all the various resources required by plants, CO_2 is the only one that is increasing on a global scale. This rise is strongly correlated with the increased rate of consumption of fossil fuels
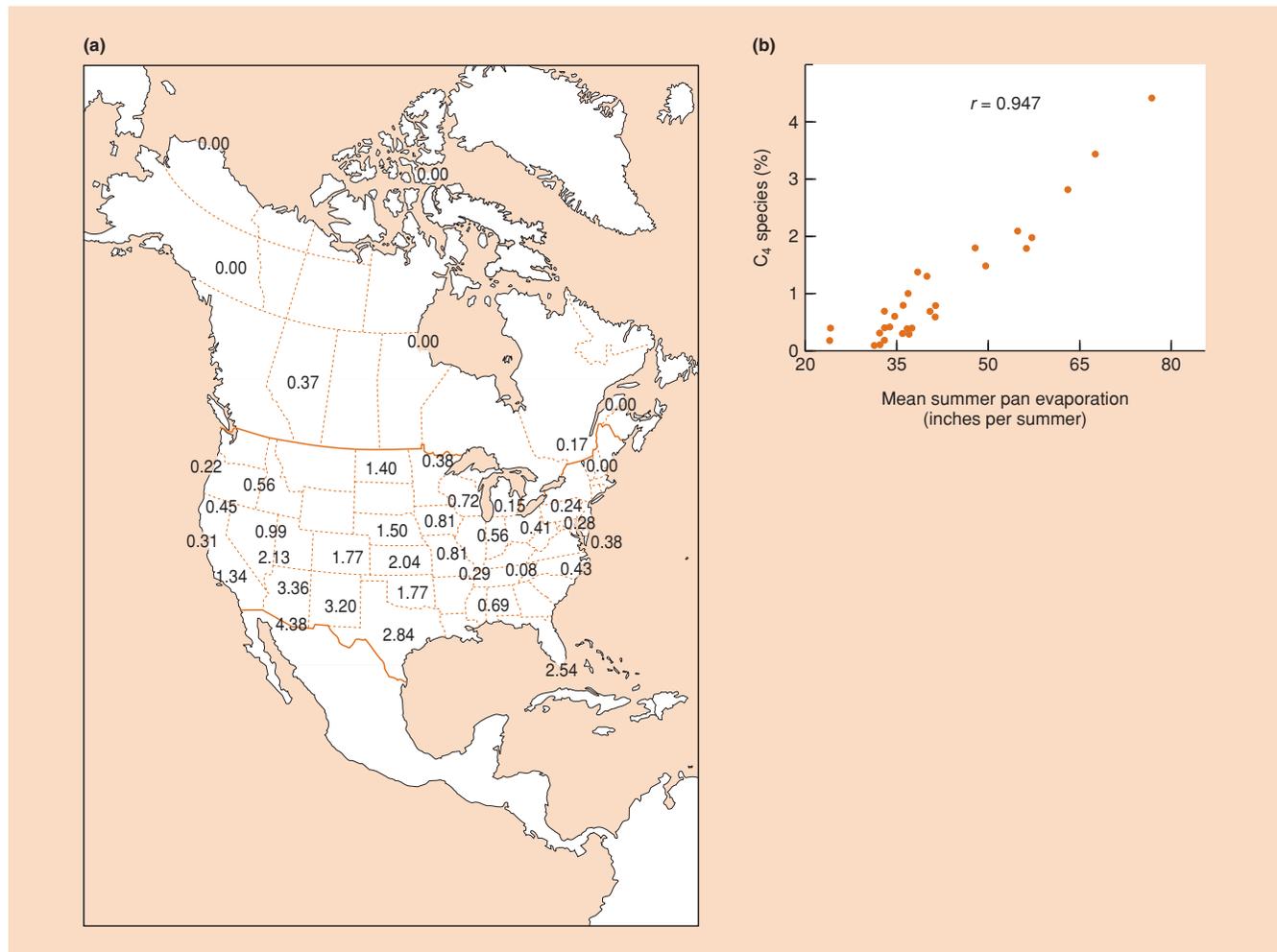
Figure 3.14 (a) The percentage of native C_4 dicot species in various regions of North America. (b) The relationship between the percentage of native C_4 species in 31 geographic regions of North America, and the mean summer (May–October) pan evaporation – a climatic indicator of plant/water balance. Regions for which appropriate climatic data were unavailable were excluded, together with south Florida, where the peculiar geography and climate may explain the aberrant composition of the flora. (After Stowe & Teeri, 1978.)

and the clearing of forests. As Loladze (2002) points out, while consequential changes to global climate may be controversial in some quarters, marked increases in CO_2 concentration itself are not. Plants now are experiencing around a 30% higher concentration compared to the pre-industrial period – effectively instantaneous on geological timescales; trees living now may experience a doubling in concentration over their lifetimes – effectively an instantaneous change on an *evolutionary* timescale; and high mixing rates in the atmosphere mean that these are changes that will affect *all* plants.

changes in geological time

There is also evidence of large-scale changes in atmospheric CO_2 over much longer timescales. Carbon balance models suggest that during the Triassic, Jurassic and Cretaceous periods, atmospheric concentrations of CO_2 were four to eight times greater than at present, falling after the Cretaceous from between 1400 and 2800 $\mu l\, l^{-1}$ to below 1000 $\mu l\, l^{-1}$ in the Eocene, Miocene and Pliocene, and fluctuating between 180 and 280 $\mu l\, l^{-1}$ during subsequent glacial and interglacial periods (Ehleringer & Monson, 1993).

The declines in CO_2 concentration in the atmosphere after the Cretaceous may have been the primary force that favored the evolution of plants with C_4 physiology (Ehleringer et al., 1991), because at low concentrations of CO_2, photorespiration places C_3 plants at a particular disadvantage. The steady rise in CO_2 since the Industrial Revolution is therefore a partial return to pre-Pleistocene conditions and C_4 plants may begin to lose some of their advantage.

When other resources are present at adequate levels, additional CO_2 scarcely influences the rate of photosynthesis of C_4 plants but increases the rate of C_3 plants. Indeed, artificially increasing the CO_2 concentration in greenhouses is a commercial technique to increase crop (C_3) yields. We might reasonably predict dramatic increases in the productivity of individual plants and of whole crops, forests and natural communities as atmospheric concentrations of CO_2 continue to increase. In the 1990s alone, results from more than 2700 studies on free-air CO_2 enrichment (FACE) experiments were published, and it is clear that, for example, doubling CO_2 concentration generally stimulates photosynthesis and increases agricultural yield by an average of 41% (Loladze, 2002). However, there is also much evidence that the responses may be complicated (Bazzaz, 1990). For example, when six species of temperate forest tree were grown for 3 years in a CO_2-enriched atmosphere in a glasshouse, they were generally larger than controls, but the CO_2 enhancement of growth declined even within the relatively short timescale of the experiment (Bazzaz *et al.*, 1993).

Moreover, there is a general tendency for CO_2 enrichment to change the composition of plants, and in particular to reduce nitrogen concentration in above-ground plant tissues – around 14% on average under CO_2 enhancement (Cotrufo *et al.*, 1998). This in turn may have indirect effects on plant–animal interactions, because insect herbivores may then eat 20–80% more foliage to maintain their nitrogen intake and fail to gain weight as fast (Figure 3.15).

CO_2 enhancement may also reduce concentrations in plants of other essential nutrients and micronutrients (Figure 3.16) (see Section 3.5), contributing in turn to 'micronutrient malnutrition', which diminishes the health and economy of more than one-half of the world's human population (Loladze, 2002).
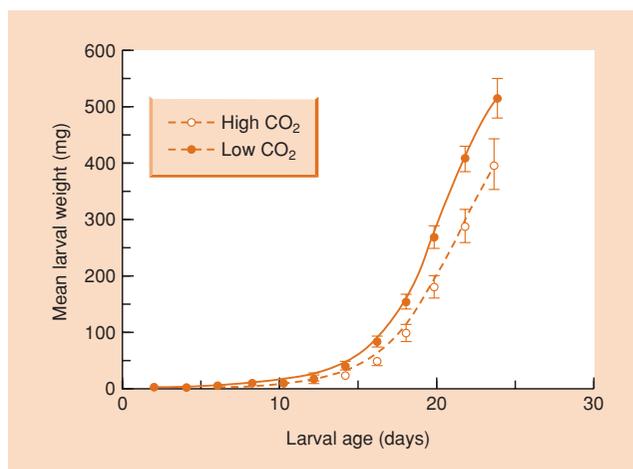

Figure 3.15   Growth of larvae of the buckeye butterfly (*Junonia coenia*) feeding on *Plantago lanceolata* that had been grown at ambient and elevated CO_2 concentrations. (After Fajer, 1989.)

3.4   Water

The volume of water that becomes incorporated in higher plants during growth is infinitesimal in comparison to the volume that flows through the plant in the transpiration stream. Nevertheless, water is a critical resource. Hydration is a necessary condition for metabolic reactions to proceed, and because no organism is completely watertight its water content needs continual replenishment. Most terrestrial animals drink free water and also generate some from the metabolism of food and body materials; there are extreme cases in which animals of arid zones may obtain all their water from their food.

3.4.1   Roots as water foragers

For most terrestrial plants, the main source of water is the soil and they gain access to it through a root system. We proceed here

Figure 3.16   Changes in the concentrations of nutrients in plant material grown at twice-ambient atmospheric CO_2 concentrations, based on 25 studies on leaves of a variety of plants (colored bars) and five studies of wheat grains (gray bars). Black lines indicate the standard errors. (After Loladze, 2002.)
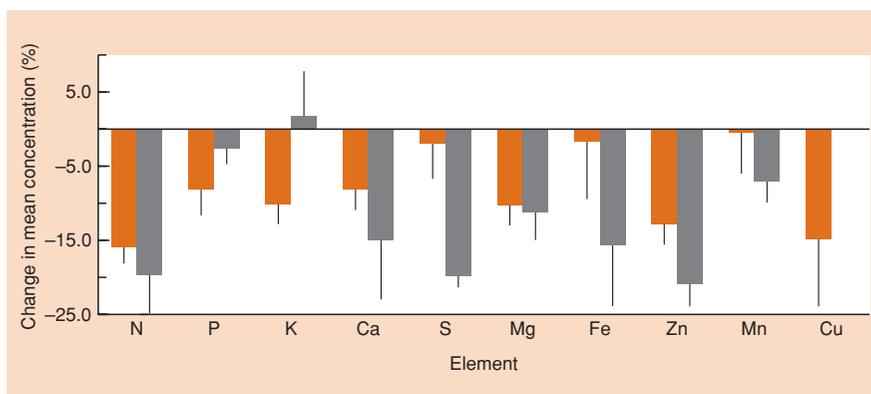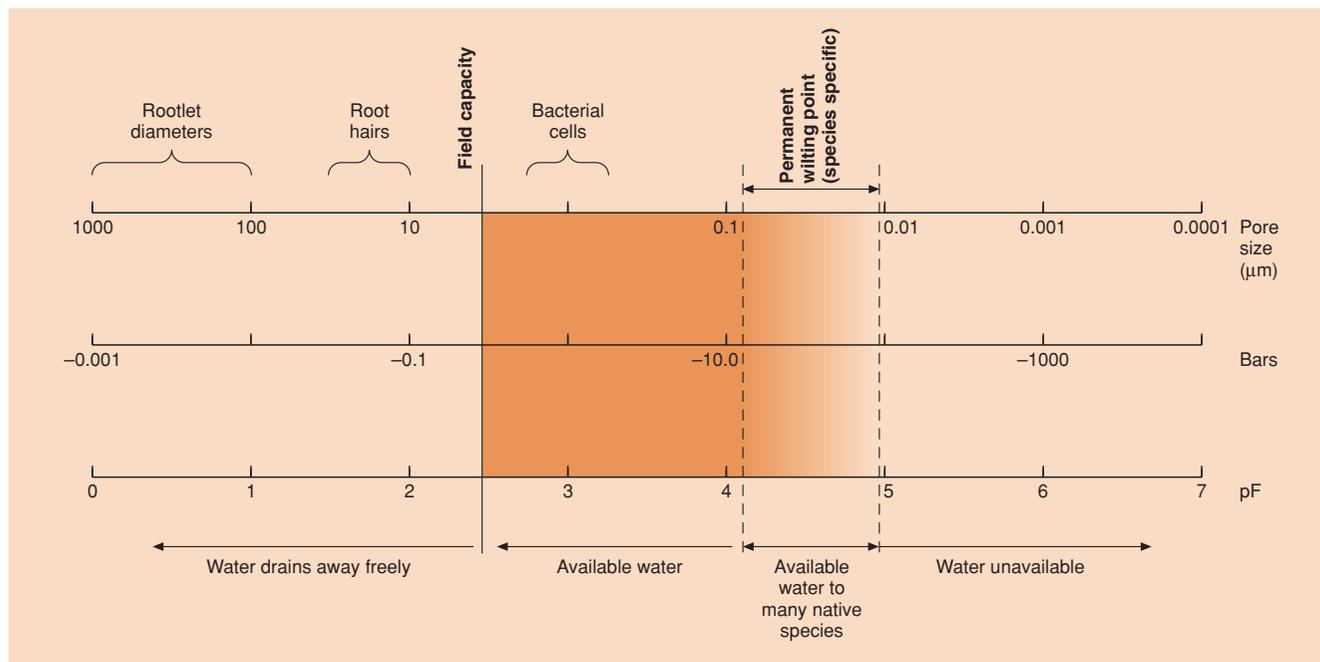
Figure 3.17   The status of water in the soil, showing the relationship between three measures of water status: (i) pF, the logarithm of the height (cm) of the column of water that the soil would support; (ii) water status expressed as atmospheres or bars; (iii) the diameter of soil pores that remain water-filled. The size of water-filled pores may be compared in the figure with the sizes of rootlets, root hairs and bacterial cells. Note that for most species of crop plant the permanent wilting point is at approximately −15 bars (−1.5 × 10⁶ Pa), but in many other species it reaches −80 bars (−8 ×10⁶ Pa), depending on the osmotic potentials that the species can develop.

(and in the next section on plant nutrient resources) on the basis of plants simply having 'roots'. In fact, most plants do not have roots – they have mycorrhizae: associations of fungal and root tissue in which both partners are crucial to the resource-gathering properties of the whole. Mycorrhizae, and the respective roles of the plants and the fungi, are discussed in Chapter 13.

It is not easy to see how roots evolved by the modification of any more primitive organ (Harper *et al.*, 1991), yet the evolution of the root was almost certainly the most influential event that made an extensive land flora and fauna possible. Once roots had evolved they provided secure anchorage for structures the size of trees and a means for making intimate contact with mineral nutrients and water within the soil.

field capacity and the permanent wilting point   Water enters the soil as rain or melting snow and forms a reservoir in the pores between soil particles. What happens to it then depends on the size of the pores, which may hold it by capillary forces against gravity. If the pores are wide, as in a sandy soil, much of the water will drain away until it reaches some impediment and accumulates as a rising watertable or finds its way into streams or rivers. The water held by soil pores against the force of gravity is called the 'field capacity' of the soil. This is the upper limit of the water that a freely drained soil will retain.

There is a less clearly defined lower limit to the water that can be used in plant growth (Figure 3.17). This is determined by the ability of plants to extract water from the narrower soil pores, and is known as the 'permanent wilting point' – the soil water content at which plants wilt and are unable to recover. The permanent wilting point does not differ much between the plant species of mesic environments (i.e. with a moderate amount of water) or between species of crop plants, but many species native to arid regions can extract significantly more water from the soil.

As a root withdraws water from the soil pores at its surface, it creates water-depletion zones around it. These determine gradients of water potential between the interconnected soil pores. Water flows along the gradient into the depleted zones, supplying further water to the root. This simple process is made much more complex because the more the soil around the roots is depleted of water, the more resistance there is to water flow. As the root starts to withdraw water from the soil, the first water that it obtains is from the wider pores because they hold the water with weaker capillary forces. This leaves only the narrower, more tortuous water-filled paths through which flow can occur, and so the resistance to water flow increases. Thus, when the root draws water from the soil very rapidly, the resource depletion zone (RDZ; see Section 3.2.1) becomes sharply defined and water can move across it only slowly. For this reason, rapidly transpiring

plants may wilt in a soil that contains abundant water. The fineness and degree of ramification of the root system through the soil then become important in determining the access of the plant to the water in the soil reservoir.

roots and the dynamics of water depletion zones

Water that arrives on a soil surface as rain or as melting snow does not distribute itself evenly. Instead, it tends to bring the surface layer to field capacity, and further rain extends this layer further and further down into the soil profile. This means that different parts of the same plant root system may encounter water held with quite different forces, and indeed the roots can move water between soil layers (Caldwell & Richards, 1986). In arid areas, where rainfall is in rare, short showers, the surface layers may be brought to field capacity whilst the rest of the soil stays at or below wilting point. This is a potential hazard in the life of a seedling that may, after rain, germinate in the wet surface layers lying above a soil mass that cannot provide the water resource to support its further growth. A variety of specialized dormancy-breaking mechanisms are found in species living in such habitats, protecting them against too quick a response to insufficient rain.

The root system that a plant establishes early in its life can determine its responsiveness to future events. Where most water is received as occasional showers on a dry substrate, a seedling with a developmental program that puts its early energy into a deep taproot will gain little from subsequent showers. By contrast, a program that determines that the taproot is formed early in life may guarantee continual access to water in an environment in which heavy rains fill a soil reservoir to depth in the spring, but there is then a long period of drought.

3.4.2 Scale, and two views of the loss of plant water to the atmosphere

There are two very different ways in which we can analyze and explain the loss of water from plants to the atmosphere. Plant physiologists going back at least to Brown and Escombe in 1900 have emphasized the way in which the behavior of the stomata determines the rate at which a leaf loses water. It now seems obvious that it is the frequency and aperture of pores in an otherwise mainly waterproof surface that will control the rate at which water diffuses from a leaf to the outside atmosphere. But micrometerologists take a quite different viewpoint, focusing on vegetation as a whole rather than on the single stoma, leaf or plant. Their approach emphasizes that water will be lost by evaporation only if there is latent heat available for this evaporation. This may be from solar radiation received directly by the transpiring leaves or as 'advective' energy, i.e. heat received as solar radiation elsewhere but transported in moving air. The micrometeorologists have developed formulae for the rate of water loss that are based entirely on the weather: wind speed, solar radiation, temperature and so

on. They wholly ignore both the species of plants and their physiology, but their models nevertheless prove to be powerful predictors of the evaporation of water from vegetation that is not suffering from drought. Neither approach is right or wrong: which to use depends on the question being asked. Large-scale, climatically based models, for example, are likely to be the most relevant in predicting the evapotranspiration and photosynthesis that might occur in areas of vegetation as a result of global warming and changes in precipitation (Aber & Federer, 1992).

3.5 Mineral nutrients

It takes more than light, CO_2 and water to make a plant. Mineral resources are also needed. The mineral resources that the plant must obtain from the soil (or,

macronutrients and trace elements

in the case of aquatic plants, from the surrounding water) include macronutrients (i.e. those needed in relatively large amounts) – nitrogen (N), phosphorus (P), sulfur (S), potassium (K), calcium (Ca), magnesium (Mg) and iron (Fe) – and a series of trace elements – for example, manganese (Mn), zinc (Zn), copper (Cu), boron (B) and molybdenum (Mo) (Figure 3.18). (Many of these elements are also essential to animals, although it is more common for animals to obtain them in organic form in their food than as inorganic chemicals.) Some plant groups have special requirements. For example, aluminum is a necessary nutrient for some ferns, silicon for diatoms and selenium for certain planktonic algae.

Green plants do not obtain their mineral resources as a single package. Each element enters the plant independently as an ion or a molecule, and each has its own characteristic properties of absorption in the soil and of diffusion, which affect its accessibility to the plant even before any selective processes of uptake occur at the root membranes. All green plants require all of the 'essential' elements listed in Figure 3.18, although not in the same proportion, and there are some quite striking differences between the mineral compositions of plant tissues of different species and between the different parts of a single plant (Figure 3.19).

Many of the points made about water as a resource, and about roots as extractors of this resource, apply

roots as foragers

equally to mineral nutrients. Strategic differences in developmental programs can be recognized between the roots of different species (Figure 3.20a), but it is the ability of root systems to override strict programs and be opportunistic that makes them effective exploiters of the soil. Most roots elongate before they produce laterals, and this ensures that exploration precedes exploitation. Branch roots usually emerge on radii of the parent root, secondary roots radiate from these primaries and tertiaries from the secondaries. These rules reduce the chance that two branches of the same root will forage in the same soil particle and enter each other's RDZs.

Essential for most organisms

☐ Essential to most living organisms

☐ Essential to animals

Essential to restricted groups of organisms

(a) Boron – Some vascular plants and algae
(b) Chromium – Probably essential in higher animals
(c) Cobalt – Essential in ruminants and N-fixing legumes
(d) Fluorine – Beneficial to bone and tooth formation
(e) Iodine – Higher animals
(f) Selenium – Some higher animals?
(g) Silicon – Diatoms
(h) Vanadium – Tunicates, echinoderms and some algae

Figure 3.18 Periodic table of the elements showing those that are essential resources in the life of various organisms.

Roots pass through a medium in which they meet obstacles and encounter heterogeneity – patches of nutrient that vary on the same scale as the diameter of a root itself. In 1 cm of growth, a root may encounter a boulder, pebbles and sand grains, a dead or living root, or the decomposing body of a worm. As a root passes through a heterogeneous soil (and all soils are heterogeneous seen from a 'root's-eye view'), it responds by branching freely in zones that supply resources, and scarcely branching in less rewarding patches (Figure 3.20b). That it can do so depends on the individual rootlet's ability to react on an extremely local scale to the conditions that it meets.

interactions between foraging for water and nutrients

There are strong interactions between water and nutrients as resources for plant growth. Roots will not grow freely into soil zones that lack available water, and so nutrients in these zones will not be exploited. Plants deprived of essential minerals make less growth and may then fail to reach volumes of soil that contain available water. There are similar interactions between mineral resources. A plant starved of nitrogen makes poor root growth and so may fail to 'forage' in areas that contain available phosphate or indeed contain more nitrogen.

Of all the major plant nutrients, nitrates move most freely in the soil solution and are carried from as far away from the root surface as water is carried. Hence nitrates will be most mobile in soils at or near field capacity, and in soils with wide pores. The RDZs for nitrates will then be wide, and those produced around neighboring roots will be more likely to overlap. Competition can then occur – even between the roots of a single plant.

The concept of RDZs is important not only in visualizing how one organism influences the resources available to another, but also in understanding how the architecture of the root system affects the capture of these resources. For a plant growing in an environment in which water moves freely to the root surface, those nutrients that are freely in solution will move with the water. They will then be most effectively captured by wide ranging, but not
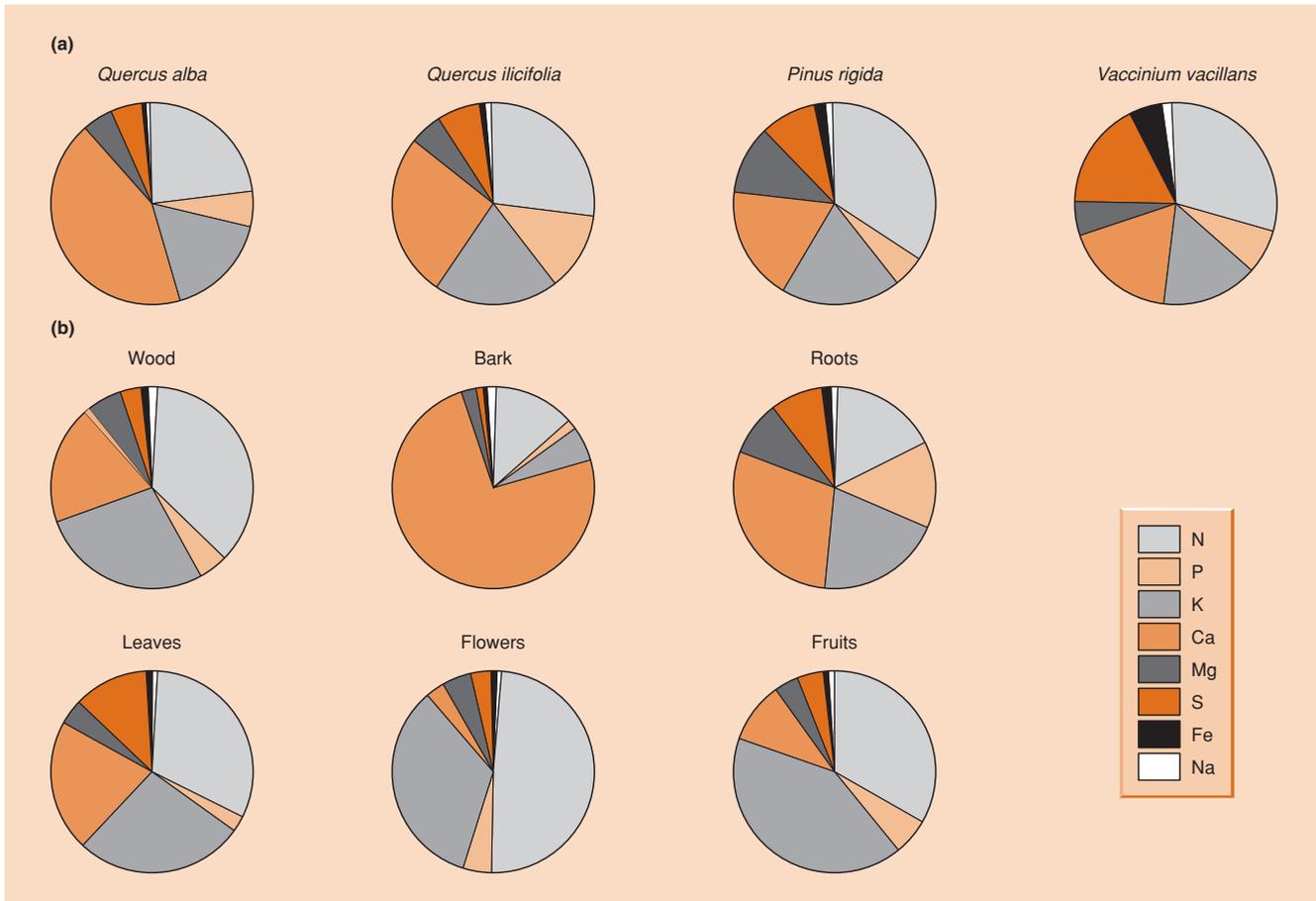
Figure 3.19 (a) The relative concentration of various minerals in whole plants of four species in the Brookhaven Forest, New York. (b) The relative concentration of various minerals in different tissues of the white oak (*Quercus alba*) in the Brookhaven Forest. Note that the differences between species are much less than between the parts of a single species. (After Woodwell *et al.*, 1975).

Figure 3.20 (a) The root systems of plants in a typical short-grass prairie after a run of years with average rainfall (Hays, Kansas). Ap, *Aristida purpurea*; Aps, *Ambrosia psilostachya*; Bd, *Buchloe dactyloides*; Bg, *Bouteloua gracilis*; Mc, *Malvastrum coccineum*; Pt, *Psoralia tenuiflora*; Sm, *Solidago mollis*. (After Albertson, 1937; Weaver & Albertson, 1943.) (b) The root system developed by a plant of wheat grown through a sandy soil containing a layer of clay. Note the responsiveness of root development to the localized environment that it encounters. (Courtesy of J.V. Lake.)
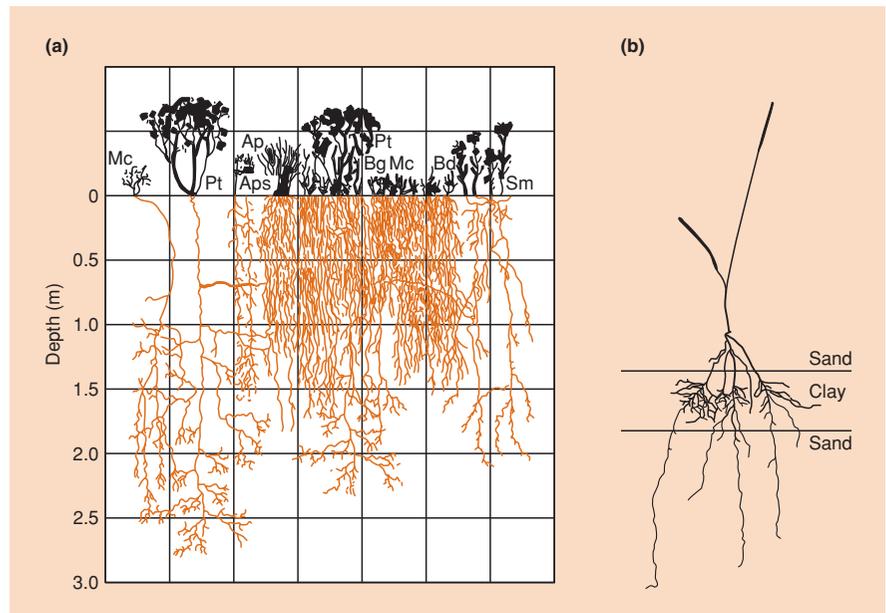
intimately branched, root systems. The less freely that water moves in the soil, the narrower will be the RDZs, and the more it will pay the plant to explore the soil intensively rather than extensively.

The soil solution that flows through soil pores to the root surface has a biased mineral composition compared with what is potentially available. This is because different mineral ions are held by different forces in the soil. Ions such as nitrate, calcium and sodium may, in a fertile agricultural soil, be carried to the root surface faster than they are accumulated in the body of the plant. By contrast, the phosphate and potassium content of the soil solution will often fall far short of the plant's requirements. Phosphate is bound on soil colloids by surfaces that bear calcium, aluminum and ferric ions, and the rate at which it can be extracted by plants then depends on the rate at which its concentration is replenished by release from the colloids. In dilute solutions, the diffusion coefficients of ions that are not absorbed, such as nitrate, are of the order of 10^{-5} cm^2 s^{-1}, and for cations such as calcium, magnesium, ammonium and potassium they are 10^{-7} cm^2 s^{-1}. For strongly absorbed anions such as phosphate, the coefficients are as low as 10^{-9} cm^2 s^{-1}. The diffusion rate is the main factor that determines the width of an RDZ.

For resources like phosphate that have low diffusion coefficients, the RDZs will be narrow (Figure 3.21); roots or root hairs will only tap common pools of resource (i.e. will compete) if they
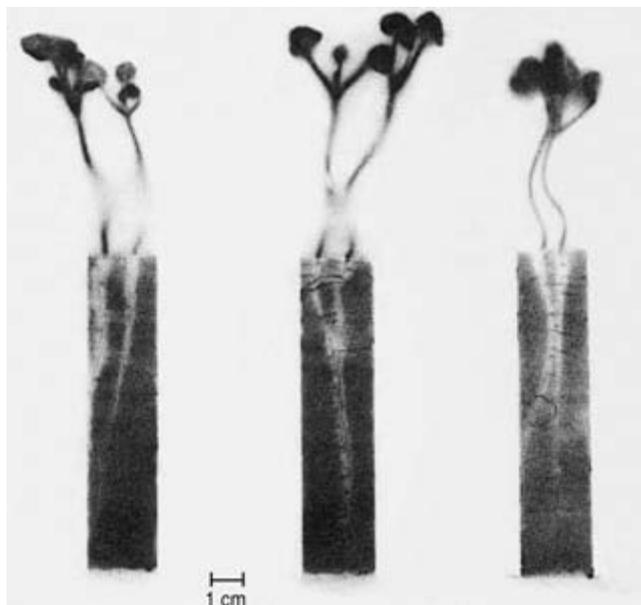

Figure 3.21 Radioautograph of soil in which seedlings of mustard have been grown. The soil was supplied with radioactively labeled phosphate (^{32}PO$_4^-$) and the zones that have been depleted by the activity of the roots show up clearly as white. (After Nye & Tinker, 1977.)

are very close together. It has been estimated that more than 90% of the phosphate absorbed by a root hair in a 4-day period will have come from the soil within 0.1 mm of its surface. Two roots will therefore only draw on the same phosphate resource in this period if they are less than 0.2 mm apart. A widely spaced, extensive root system tends to maximize access to nitrate, whilst a narrowly spaced, intensively branched root system tends to maximize access to phosphates (Nye & Tinker, 1977). Plants with different shapes of root system may therefore tolerate different levels of soil mineral resources, and different species may deplete different mineral resources to different extents. This may be of great importance in allowing a variety of plant species to cohabit in the same area (coexistence of competitors is discussed in Chapters 8 and 19).

3.6   Oxygen

Oxygen is a resource for both animals and plants. Only a few prokaryotes can do without it. Its diffusibility and solubility in water are very low and so it becomes limiting most quickly in aquatic and waterlogged environments. Its solubility in water also decreases rapidly with increasing temperature. When organic matter decomposes in an aquatic environment, microbial respiration makes a demand for oxygen and this 'biological oxygen demand' may constrain the types of higher animal that can persist. High biological oxygen demands are particularly characteristic of still waters into which leaf litter or organic pollutants are deposited and they become most acute during periods of high temperature.

Because oxygen diffuses so slowly in water, aquatic animals must either maintain a continual flow of water over their respiratory surfaces (e.g. the gills of fish), or have very large surface areas relative to body volume (e.g. many aquatic crustacea have large feathery appendages), or have specialized respiratory pigments or a slow respiration rate (e.g. the midge larvae that live in still and nutrient-rich waters), or continually return to the surface to breathe (e.g. whales, dolphins, turtles and newts).

The roots of many higher plants fail to grow into waterlogged soil, or die if the water table rises after they have penetrated deeply. These reactions may be direct responses to oxygen deficiency or responses to the accumulation of gases such as hydrogen sulfide, methane and ethylene, which are produced by microorganisms engaged in anaerobic decomposition. Even if roots do not die when starved of oxygen, they may cease to absorb mineral nutrients so that the plants suffer from mineral deficiencies.

3.7   Organisms as food resources

Autotrophic organisms (green plants and certain bacteria) assimilate inorganic resources into packages of organic

molecules (proteins, carbohydrates, etc.). These become the resources for *heterotrophic* organisms (decomposers, parasites, predators and grazers), which take part in a chain of events in which each consumer of a resource becomes, in turn, a resource for another consumer. At each link in this food chain the most obvious distinction is between saprotrophs and predators (defined broadly).

Saprotrophs – bacteria, fungi and detritivorous animals (see Chapter 11) – use other organisms, or parts of other organisms, as food but only after they have died, or they consume another organism's waste or secretory products.

|saprotrophs, predators, grazers and parasites|

Predators use other living organisms, or parts of other living organisms, as food. True predators predictably kill their prey. Examples include a mountain lion consuming a rabbit but also consumers that we may not refer to as predators in everyday speech: a water flea consuming phytoplankton cells, a squirrel eating an acorn, and even a pitcherplant drowning a mosquito. *Grazing* can also be regarded as a type of predation, but the food (prey) organism is not killed; only part of the prey is taken, leaving the remainder with the potential to regenerate. Grazers feed on (or from) many prey during their lifetime. True predation and grazing are discussed in detail in Chapter 9. *Parasitism*, too, is a form of predation in which the consumer usually does not kill its food organism; but unlike a grazer, a parasite feeds from only one or a very few host organisms in its lifetime (see Chapter 12).

|specialists and generalists|

An important distinction amongst animal consumers is whether they are specialized or generalized in their diet. Generalists (*polyphagous* species) take a wide variety of prey species, though they very often have clear preferences and a rank order of what they will choose when there are alternatives available. Specialists may consume only particular parts of their prey though they range over a number of species. This is most common among herbivores because, as we shall see, different parts of plants are quite different in their composition. Thus, many birds specialize on eating seeds though they are seldom restricted to a particular species. Other specialists, however, may feed on only a narrow range of closely related species or even just a single species (when they are said to be *monophagous*). Examples are caterpillars of the cinnabar moth (which eat the leaves, flower buds and very young stems of species of ragwort, *Senecio*) and many species of host-specific parasites.

Many of the resource-use patterns found among animals reflect the different lifespans of the consumer and what it consumes. Individuals of long-lived species are likely to be generalists: they cannot depend on one food resource being available throughout their life. Specialization is increasingly likely if a consumer has a short lifespan. Evolutionary forces can then shape the timing of the consumer's food demands to match the

timetable of its prey. Specialization also allows the evolution of structures that make it possible to deal very efficiently with particular resources – this is especially the case with mouthparts. A structure like the stylet of an aphid (Figure 3.22) can be interpreted as an exquisite product of the evolutionary process that has given the aphid access to a valuable food resource – or as an example of the ever-deepening rut of specialization that has constrained what aphids can feed on. The more specialized the food resource required by an organism, the more it is constrained to live in patches of that resource *or* to spend time and energy in searching for it among a mixture of resources. This is one of the costs of specialization.

3.7.1   The nutritional content of plants and animals as food

As a 'package' of resources, the body of a green plant is quite different from the body of an animal. This has a tremendous effect on the value of these resources as potential food (Figure 3.23). The most important contrast is that plant cells are bounded by walls of cellulose, lignin and/or other structural materials. It is these cell walls that give plant material its high fiber content. The presence of cell walls is also largely responsible for the high fixed carbon content of plant tissues and the high ratio of carbon to other important elements. For example, the carbon : nitrogen (C : N) ratio of plant tissues commonly exceeds 40 : 1, in contrast to the ratios of approximately 10 : 1 in bacteria, fungi and animals. Unlike plants, animal tissues contain no structural carbohydrate or fiber component but are rich in fat and, in particular, protein.

|C : N ratios in animals and plants|

The various parts of a plant have very different compositions (Figure 3.23) and so offer quite different resources. Bark, for example, is largely composed of dead cells with corky and lignified walls and is quite useless as a food for most herbivores (even species of 'bark beetle' specialize on the nutritious cambium layer just beneath the bark, rather than on the bark itself). The richest concentrations of plant proteins (and hence of nitrogen) are in the meristems in the buds at shoot apices and in leaf axils. Not surprisingly, these are usually heavily protected with bud scales and defended from herbivores by thorns and spines. Seeds are usually dried, packaged reserves rich in starch or oils as well as specialized storage proteins. And the very sugary and fleshy fruits are resources provided by the plant as 'payment' to the animals that disperse the seeds. Very little of the plants' nitrogen is 'spent' on these rewards.

|different plant parts represent very different resources . . .|

The dietary value of different tissues and organs is so different that it is no surprise to find that most small herbivores are specialists – not only on particular species or plant groups, but on particular plant parts: meristems, leaves, roots, stems, etc. The smaller the herbivore, the finer is the scale of heterogeneity of
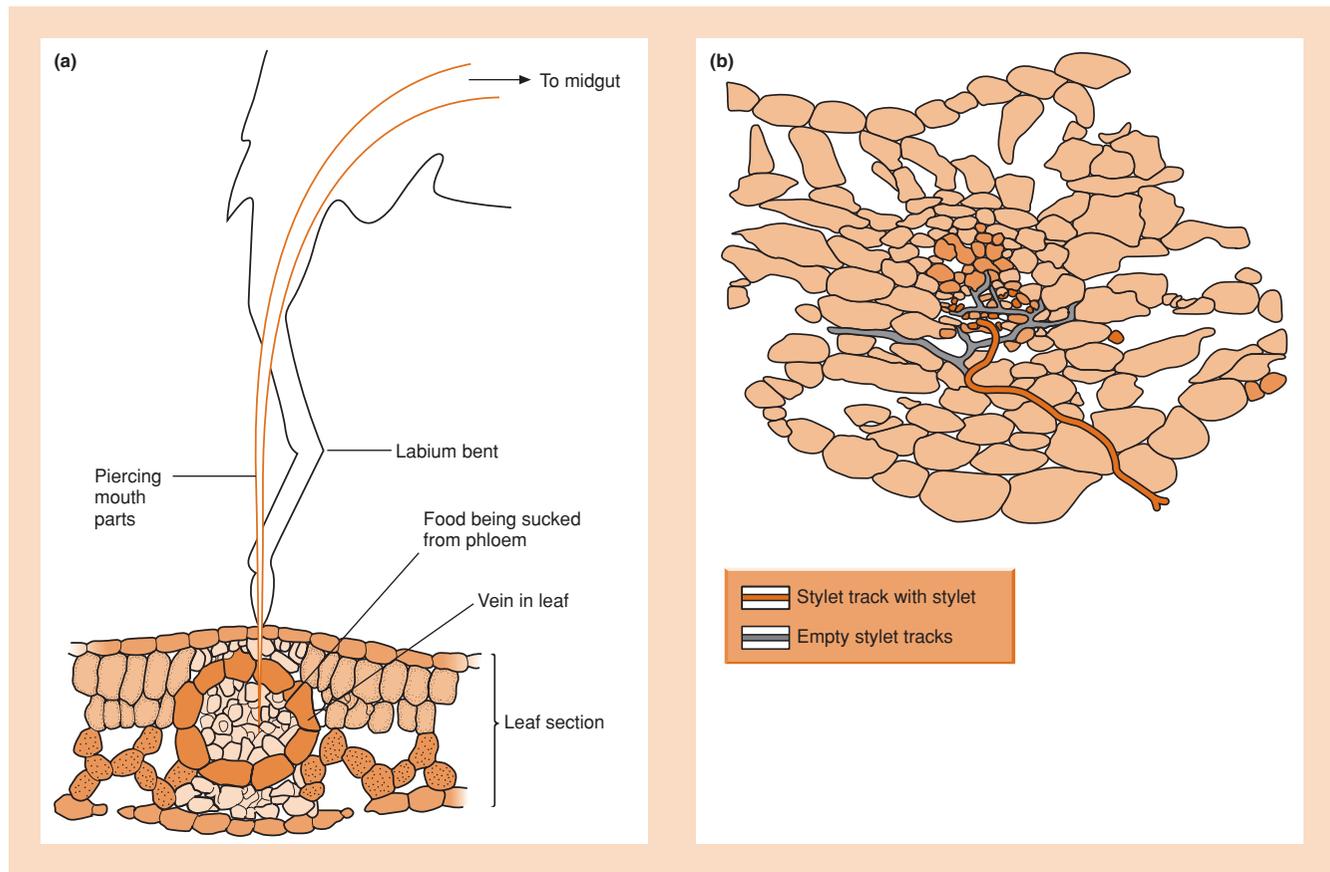
Figure 3.22 The stylet of an aphid penetrating the host tissues and reaching the sugar-rich phloem cells in the leaf veins. (a) Aphid mouthparts and cross-section of a leaf. (b) A stylet, showing its circuitous path through a leaf. (After Tjallingii & Hogen Esch, 1993.)

the plant on which it may specialize. Extreme examples can be found in the larvae of various species of oak gall wasps, some of which may specialize on young leaves, some on old leaves, some on vegetative buds, some on male flowers and others on root tissues.

> . . . but the composition of all herbivores is remarkably similar

Although plants and their parts may differ widely in the resources they offer to potential consumers, the composition of the bodies of different herbivores is remarkably similar. In terms of the content of protein, carbohydrate, fat, water and minerals per gram there is very little to choose between a diet of caterpillars, cod or venison. The packages may be differently parceled (and the taste may be different), but the contents are essentially the same. Carnivores, then, are not faced with problems of digestion (and they vary rather little in their digestive apparatus), but rather with difficulties in finding, catching and handling their prey (see Chapter 9).

Differences in detail aside, herbivores that consume living plant material – and saprotrophs that consume dead plant material – all utilize a food resource that is rich in carbon and poor in protein. Hence, the transition from plant to consumer involves a massive burning off of carbon as the C : N ratio is lowered. This is the realm of ecological stoichiometry (Elser & Urabe 1999): the analysis of constraints and consequences in ecological interactions of the mass balance of multiple chemical elements (particularly the ratios of carbon to nitrogen and of carbon to phosphorus – see Sections 11.2.4 and 18.2.5). The main waste products of organisms that consume plants are carbon-rich compounds: CO_2, fiber, and in the case of aphids, for example, carbon-rich honeydew dripping from infested trees. By contrast, the greater part of the energy requirements of carnivores is obtained from the protein and fats of their prey, and their main excretory products are in consequence nitrogenous.

The differential in C : N ratios between plants and microbial decomposers also means that the long-term effects of CO_2 enhancement (see Section 3.3.4) are not as straightforward as might be imagined (Figure 3.24): that is, it is not necessarily the case that plant

> C : N ratios and the effects of CO_2 enhancement
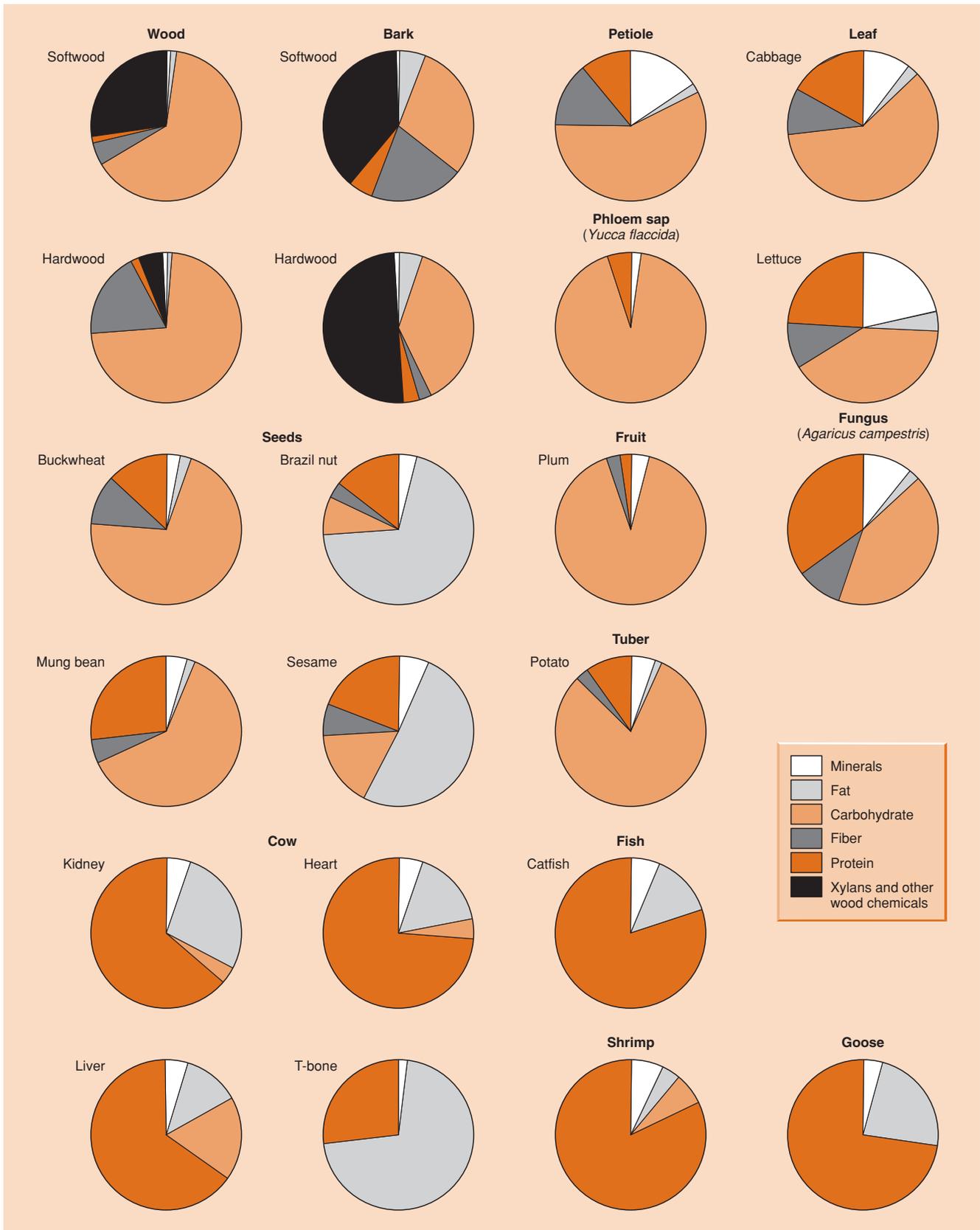
Figure 3.23 The composition of various plant parts and of the bodies of animals that serve as food resources for other organisms. (Data from various sources.)
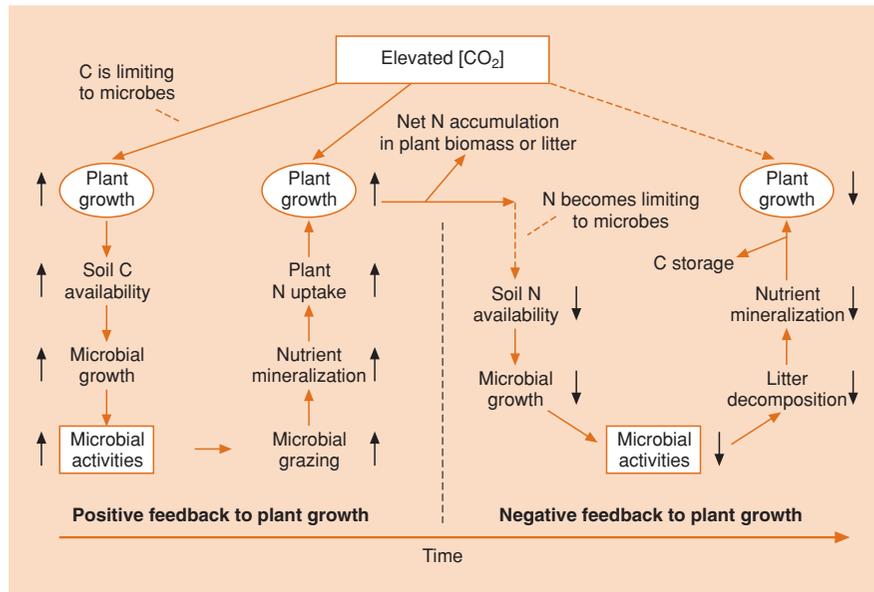
Figure 3.24 Potential positive and negative feedback from elevated CO_2 concentrations to plant growth, to microbial activity and back to plant growth. The arrows between descriptors indicate causation; the black arrows alongside descriptors indicate increases or decreases in activity. The dashed arrow from elevated $[CO_2]$ to plant growth indicates that any effect may be absent as a result of nutrient-limitation. (After Hu *et al.*, 1999.)

biomass is increased. If the microbes themselves are carbon-limited, then increased CO_2 concentrations, apart from their direct effects on plants, might stimulate microbial activity, making other nutrients, especially nitrogen, available to plants, further stimulating plant growth. Certainly, short-term experiments have demonstrated this kind of effect on decomposer communities. On the other hand, though, decomposers may be nitrogen-limited, either initially or following a period of enhanced plant growth during which nitrogen accumulates in plant biomass and litter. Then, microbial activity would be depressed, diminishing the release of nutrients to plants and potentially preventing their enhanced growth in spite of elevated CO_2 concentrations. These, though, are longer term effects and to date very few data have been collected to detect them. The more general issue of local and global 'carbon budgets' is taken up again in Section 18.4.6.

3.7.2 Digestion and assimilation of plant material

the large amounts of fixed carbon in plant materials mean that they are potentially rich sources of energy. It is other components of the diet (e.g. nitrogen) that are more likely to be limiting. Yet most of that energy is only directly available to consumers if they have enzymes capable of mobilizing cellulose and lignins, whereas the overwhelming majority of species in both the plant and animal kingdoms lack these enzymes. Of all the many constraints that put limits on what living organisms can do, the failure of so many to have evolved cellulolytic enzymes is a particular evolutionary puzzle. It may be that gut-inhabiting, cellulolytic prokaryotes have so readily formed

intimate, 'symbiotic' relationships with herbivores (see Chapter 13) that there has been little selection pressure to evolve cellulases of their own (Martin, 1991). It is now recognized that a number of insects do indeed produce their own cellulases but the vast majority nevertheless depend on symbionts.

Because most animals lack cellulases, the cell wall material of plants hinders the access of digestive enzymes to the contents of plant cells. The acts of chewing by the grazing mammal, cooking by humans and grinding in the gizzard of birds allow digestive enzymes to reach cell contents more easily. The carnivore, by contrast, can more safely gulp its food.

When plant parts are decomposed, material with a high carbon content is converted to microbial bodies with a relatively low carbon content – the limitations on microbial growth and multiplication are resources other than carbon. Thus, when microbes multiply on a decaying plant part, they withdraw nitrogen and other mineral resources from their surroundings and build them into their own microbial bodies. For this reason, and because microbial tissue is more readily digested and assimilated, plant detritus that has been richly colonized by microorganisms is generally preferred by detritivorous animals.

In herbivorous vertebrates the rate of energy gain from different dietary resources is determined by the structure of the gut – in particular, the balance between a well-stirred anterior chamber in which microbial fermentation occurs (AF), a connecting tube in which there is digestion but no fermentation (D), and a posterior fermentation chamber, the colon and cecum (PF). Models of such three-part digestive systems (Alexander, 1991) suggest that large AF, small D and small PF (e.g. the ruminant) would give near-optimal gains

from poor-quality food, and that large PF, as in horses, is more appropriate for food with less cell wall material and more cell contents. For very high-quality food (a very high proportion of cell contents and little cell wall material) the optimum gut has long D and no AF or PF.

Elephants, lagomorphs and some rodents eat their own feces and so double the distance traveled by the food resource through the digestive system. This allows further fermentation and digestion but may also allow time for dietary deficiencies (e.g. of vitamins) to be made good by microbial synthesis. These issues are picked up again in Section 13.5.

3.7.3   Physical defenses

coevolution

All organisms are potentially food resources for others and so it is not surprising that many organisms have evolved physical, chemical, morphological and/or behavioral defenses that reduce the chance of an encounter with a consumer and/or increase the chance of surviving such an encounter. But the interaction does not necessarily stop there. A better defended food resource itself exerts a selection pressure on consumers to overcome that defense; though in overcoming that defense, rather than the defenses of other species, the consumer is likely to become relatively specialized on that resource – which is then under particular pressure to defend itself against that particular consumer, and so on. A continuing interaction can therefore be envisaged in which the evolution of both the consumer and the organism consumed depend crucially on the evolution of the other: a coevolutionary 'arms race' (Ehrlich & Raven, 1964), which, in its most extreme form, has a coadapted pair of species locked together in perpetual struggle.

Of course, the resources of green plants (and of autotrophs in general) are not alive and cannot therefore evolve defenses. Coevolution is also not possible between decomposer organisms and their dead food resources, although bacteria, fungi and detritivorous animals will often have to contend with the residual effects of physical and, in particular, chemical defenses in their food.

spines

Simple spines can be an effective deterrent. The spiny leaves of holly are not eaten by oak eggar moth larvae (*Lasiocampa quercus*), but if the spines are removed the leaves are eaten readily. No doubt a similar result would be achieved with foxes as predators and de-spined hedgehogs as prey. In many small planktonic invertebrates that live in lakes, the development of spines, crests and other appendages that reduce their vulnerability to predation can be induced by a predator's presence. Thus, for example, spine development in the offspring of brachionid rotifers, including *Keratella cochlearis*, is promoted if their mother was cultured in a medium conditioned by the predatory rotifer, *Asplachna priodonta* (Stemberger & Gilbert, 1984; Snell, 1998). At a smaller scale still, many plant surfaces are clothed in epidermal hairs (trichomes) and in some species these develop thick secondary walls to form strong hooks or points that may trap or impale insects.

shells

Any feature that increases the energy a consumer spends in discovering or handling a food item – the thick shell of a nut or the fibrous cone on a pine – is a defense if, as a consequence, the consumer eats less of it. The green plant uses none of its energetic resources in running away and so may have relatively more available to invest in energy-rich defense structures. Moreover, most green plants are probably relatively overprovided with energy resources, making it cheap to build shells around seeds and woody spines on stems – mainly out of cellulose and lignin – and so protecting the real riches: the scarce resources of nitrogen, phosphorus, potassium, etc. in the embryos and meristems.

seeds: dissipation or protection

Seeds are most at risk to predators when they have just ripened and are still attached, in a cone or ovary, to the parent plant, but their value is literally dissipated as soon as the capsule opens and the seeds are shed. The poppies illustrate this point. The seeds of wild poppies are shed through a series of pores at the apex of the capsule as it waves in the wind. Two of the species, *Papaver rhoeas* and *P. dubium*, open these pores as soon as the seed is ripe and the capsules are often empty by the following day. Two other species, *P. argemone* and *P. hybridum*, have seeds that are large relative to the size of the capsule pores and dispersal is a slow process over the fall and winter months. The capsules of these species are defended by spines. The cultivated poppy (*P. somniferum*) by contrast, has been selected by humans not to disperse its seeds – the capsule pores do not open. Birds can therefore be a serious pest of the cultivated poppy; they tear open the capsules to reach an oil- and protein-rich reward. Humans, in fact, have selected most of their crops to retain rather than disperse their seeds and these represent sitting targets for seed-eating birds.

3.7.4   Chemical defenses

secondary chemicals: protectants?

The plant kingdom is very rich in chemicals that apparently play no role in the normal pathways of plant biochemistry. These 'secondary' chemicals range from simple molecules like oxalic acid and cyanide to the more complex glucosinolates, alkaloids, terpenoids, saponins, flavonoids and tannins (Futuyma, 1983). Many of these have been shown to be toxic to a wide range of potential consumers. For example, populations of white clover, *Trifolium repens*, are commonly polymorphic for the ability to release hydrogen cyanide when the tissues are attacked. Plants that lack the ability to generate hydrogen cyanide are eaten by slugs and snails: the cyanogenic forms are nibbled but then rejected. Many researchers have assumed that

protection against consumers has provided the selective pressure favoring the production of such chemicals. Many others, however, have questioned whether the selective force of herbivory is powerful enough for this (their production may be costly to the plants in terms of essential nutrients) and have pointed to other properties that they possess: for example as protectants against ultraviolet radiation (Shirley, 1996). None the less, in the few cases where selection experiments have been carried out, plants reared in the presence of consumers have evolved enhanced defenses against these enemies, relative to control plants reared in the absence of consumers (Rausher, 2001). Later, in Chapter 9 when we look in more detail at the *inter*action between predators and their prey, we will look at the costs and benefits of prey (especially plant) defense to both the prey itself and its consumers. Here, we focus more on the nature of those defenses.

| apparency theory | If the attentions of herbivores select for plant defensive chemicals, then equally, those chemicals will select for |

adaptations in herbivores that can overcome them: a classic coevolutionary 'arms race'. This, though, suggests that plants should become ever more noxious and herbivores ever more specialized, leaving unanswered the question of why there are so many generalist herbivores, capable of feeding from many plants (Cornell & Hawkins, 2003). An answer has been suggested by 'apparency theory' (Feeny, 1976; Rhoades & Cates, 1976). This is based on the observation that noxious plant chemicals can be classified broadly into two types: (i) toxic (or qualitative) chemicals, which are poisonous even in small quantities; and (ii) digestion-reducing (or quantitative) chemicals, which act in proportion to their concentration. Tannins are an example of the second type. They bind proteins, rendering tissues such as mature oak leaves relatively indigestible. The theory further supposes that toxic chemicals, by virtue of their specificity, are likely to be the foundation of an arms race, requiring an equally simple and specific response from a herbivore; whereas chemicals that make plants generally indigestible are much more difficult to overcome.

Apparency theory then proposes that relatively short-lived, ephemeral plants (said to be 'unapparent') gain a measure of protection from consumers because of the unpredictability of their appearance in space and time. They therefore need to invest less in defense than predictable, long-lived ('apparent') species like forest trees. Moreover, the apparent species, precisely because they are apparent for long, predictable periods to a large number of herbivores, should invest in digestion-reducing chemicals that, while costly, will afford them broad protection; whereas unapparent plants should produce toxins since it is only likely to pay a few specialist species to coevolve against them.

Apparency theory, incorporating ideas on coevolution, therefore makes a number of predictions (Cornell & Hawkins, 2003). The most obvious is that more unapparent plants are more likely to be protected by simple, toxic compounds than by more complex, digestion-inhibiting compounds. This can even be seen in the changing balance of chemical defense in some plants as the season progresses. For example, in the bracken fern (*Pteridium aquilinum*), the young leaves that push up through the soil in spring are less apparent to potential herbivores than the luxuriant foliage in late summer. The young leaves are rich in cyanogenic glucosinolates, whilst the tannin content steadily increases in concentration to its maximum in mature leaves (Rhoades & Cates, 1976).

A more subtle prediction of the theory is that specialist herbivores, having invested evolutionarily in overcoming particular chemicals, should perform best when faced with those chemicals (compared to chemicals they would not normally encounter); whereas generalists, having invested in performing well when faced with a wide range of chemicals, should perform least well when faced with chemicals that have provoked coevolutionary responses from specialists. This is supported by an analysis of a wide range of data sets for insect herbivores fed on artificial diets with added chemicals (892 insect/chemical combinations) shown in Figure 3.25.

Furthermore, plants are predicted to differ in their chemical defenses not only from species to species but also within an individual plant. 'Optimal defense theory' predicts that the more

| | optimal defense theory: constitutive and inducible defenses |

important an organ or tissue is for an organism's fitness, the better protected it will be; and in the present context, it predicts that more important plant parts should be protected by *constitutive chemicals* (produced all the time), whereas less important parts should rely on *inducible chemicals*, only produced in response to damage itself, and hence with far lower fixed costs to the plants (McKey, 1979; Strauss *et al.*, 2004). This is confirmed, for example, by a study of wild radish, *Raphanus sativus*, in which plants were either subjected to herbivory by caterpillars of the butterfly, *Pieris rapae*, or left as unmanipulated controls (Strauss *et al.*, 2004). Petals (and all parts of the flower) are known in this insect-pollinated plant to be highly important to fitness. Concentrations of protective glucosinolates were twice as high in petals as in undamaged leaves, and these levels were maintained constitutively, irrespective of whether the petals were damaged by the caterpillars (Figure 3.26). Leaves, on the other hand, have a much less direct influence on fitness: high levels of leaf damage can be sustained without any measurable effect on reproductive output. Constitutive levels of glucosinolates, as already noted, were low; but if the leaves were damaged the (induced) concentrations were even higher than in the petals.

Similar results were found for the brown seaweed, *Sargassum filipendula*, where the holdfast at its base was the most valuable tissue: without it the plant would be cast adrift in the water (Taylor *et al.*, 2002). This was protected by costly constitutive, quantitative chemicals, whereas the much less valuable youngest stipes (effectively stems) near the tip of the plant were protected only by toxic chemicals induced by grazing.
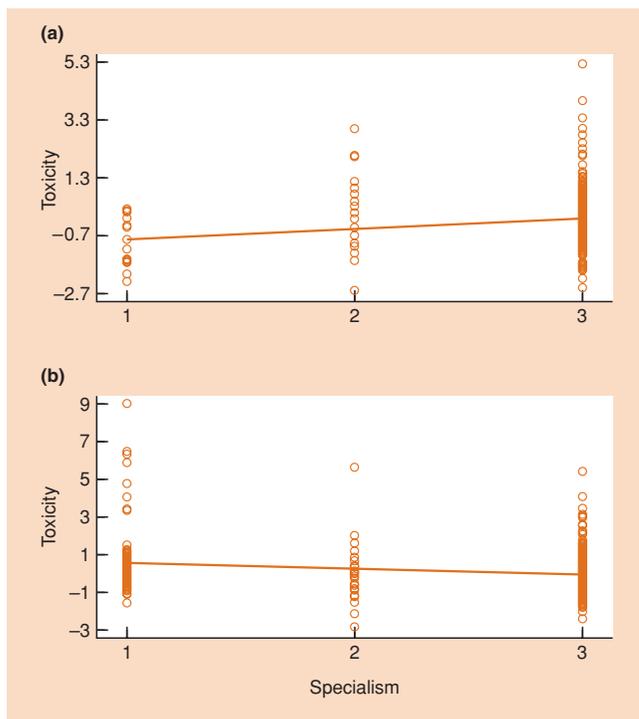
Figure 3.25 Combining data from a wide range of published studies, herbivores were split into three groups: 1, specialists (feeding from one or two plant families), 2, oligophages (3–9 families) and 3, generalists (more than nine families). Chemicals were split into two groups: (a) those that are, and (b) those that are not, found in the normal hosts of specialists and oligophages. With increasing specialization, (a) herbivores suffered decreased mortality on chemicals that have not provoked a coevolutionary response from specialist herbivores, but (b) suffered higher mortality on chemicals that have not provoked such a response. Regressions: (a) $y = 0.33x - 1.12$; $r^2 = 0.032$; $t = 3.25$; $P = 0.0013$; (b) $y = 0.93 - 0.36x$; $r^2 = 0.049$; $t = -4.35$; $P < 0.00001$. (After Cornell & Hawkins, 2003.)

animal defenses

Animals have more options than plants when it comes to defending themselves, but some still make use of chemicals. For example, defensive secretions of sulfuric acid of pH 1 or 2 occur in some marine gastropod groups, including the cowries. Other animals that can tolerate the chemical defenses of their plant food, store and use them in their own defense. A classic example is the monarch butterfly (*Danaus plexippus*), whose caterpillars feed on milkweeds (*Asclepias* spp.). Milkweeds contain secondary chemicals, cardiac glycosides, which affect the vertebrate heartbeat and are poisonous to mammals and birds. Monarch caterpillars can store the poison, and it is still present in the adults, which in consequence are completely unacceptable to bird predators. A naive blue jay (*Cyanocitta cristata*) (i.e. one that has not tried a monarch
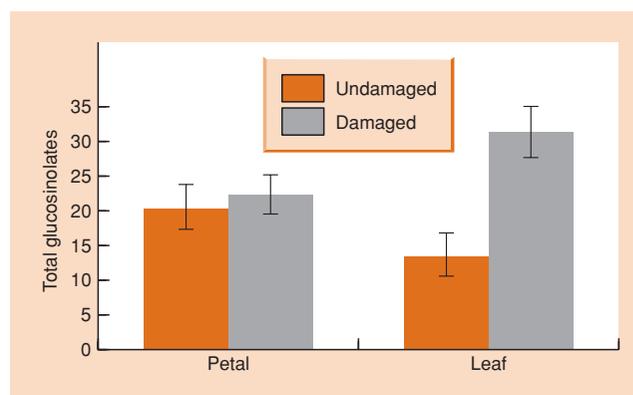

Figure 3.26 Concentrations of glucosinolates ($\mu g\ mg^{-1}$ dry mass) in the petals and leaves of wild radish, *Raphanus sativus*, either undamaged or damaged by caterpillars of *Pieris rapae*. Bars are standard errors. (After Strauss *et al.*, 2004.)

butterfly before) will vomit violently after eating one, and once it recovers it will reject all others on sight. In contrast, monarchs reared on cabbage are edible (Brower & Corvinó, 1967).

Chemical defenses are not equally effective against all consumers. Indeed, what is unacceptable to most animals may be the chosen, even unique, diet of others. It is, after all, an inevitable consequence of having evolved resistance to a plant's defenses that a consumer will have gained access to a resource unavailable to most (or all) other species. For example, the tropical legume *Dioclea metacarpa* is toxic to almost all insect species because it contains a nonprotein amino acid, L-canavanine, which insects incorporate into their proteins in place of arginine. But a species of bruchid beetle, *Caryedes brasiliensis*, has evolved a modified tRNA synthetase that distinguishes between L-canavanine and arginine, and the larvae of these beetles feed solely on *D. metacarpa* (Rosenthal *et al.*, 1976).

one man's poison is another man's meat

3.7.5 Crypsis, aposematism and mimicry

An animal may be less obvious to a predator if it matches its background, or possesses a pattern that disrupts its outline, or resembles an inedible feature of its environment. Straightforward examples of such *crypsis* are the green coloration of many grasshoppers and caterpillars, and the transparency of many planktonic animals that inhabit the surface layers of oceans and lakes. More dramatic cases are the sargassum fish (*Histrio pictus*), whose body outline mimics the sargassum weed in which it is found, or the caterpillar of the viceroy butterfly (*Limenitis archippus*) that resembles a bird dropping. Cryptic animals may be highly palatable, but their morphology and color (and their choice of the appropriate background) reduce the likelihood that they will be used as a resource.

crypsis

aposematism

Whilst crypsis may be a defense strategy for a palatable organism, noxious or dangerous animals often seem to advertize the fact by bright, conspicuous colors and patterns. This phenomenon is referred to as *aposematism*. The monarch butterfly, discussed above, is aposematically colored, as is its caterpillar, which actually sequesters the defensive cardiac glucosinolates from its food. The usual evolutionary argument for this runs as follows: conspicuous coloration will be favored because noxious prey will be recognized (memorized) as such by experienced predators, and thus will be protected, whereas the costs of 'educating' the predator will have been shared amongst the whole population of conspicuous prey. This argument, however, leaves unanswered the question of how conspicuous, noxious prey arose in the first place, since when initially rare, they seem likely to be repeatedly eliminated by naive (i.e. 'uneducated') predators (Speed & Ruxton, 2002). One possible answer is that predators and prey have coevolved: in each generation – from an original mixture of conspicuous and inconspicuous, noxious and edible prey – conspicuous edible prey are eliminated, and, with conspicuous prey therefore becoming disproportionately noxious, predators evolve an increased wariness for conspicuous prey (Sherratt, 2002).

Batesian and Müllerian mimicry

The adoption of memorable body patterns by distasteful prey also immediately opens the door for deceit by other species, because there will be a clear evolutionary advantage to a palatable prey, 'the mimic', if it looks like an unpalatable species, 'the model' (Batesian mimicry). Developing the story of the monarch butterfly a little further, the adult of the palatable viceroy butterfly mimics the distasteful monarch, and a blue jay that has learned to avoid monarchs will also avoid viceroys. There will also be an advantage to aposematically colored, distasteful prey in looking like one another (Müllerian mimicry), though many unanswered questions remain as to where exactly Batesian mimicry ends and Müllerian mimicry begins, in part because there are more theoretical viewpoints than impeccable data sets that might distinguish between them (Speed, 1999).

By living in holes (e.g. millipedes and moles) animals may avoid stimulating the sensory receptors of predators, and by 'playing dead' (e.g. the opossum *Didelphis virginiana* and African ground squirrels) animals may fail to stimulate a killing response. Animals that withdraw to a prepared retreat (e.g. rabbits and prairie dogs to their burrows, snails to their shells), or which roll up and protect their vulnerable parts by a tough exterior (e.g. armadillos and pill millipedes), reduce their chance of capture but stake their lives on the chance that the attacker will not be able to breach their defenses. Other animals seem to try to bluff themselves out of trouble by threat displays. The startle response of moths and butterflies that suddenly expose eye-spots on their wings is one example. No doubt the most common behavioral response of an animal in danger of being preyed upon is to flee.

## 3.8   A classification of resources, and the ecological niche

We have seen that every plant requires many distinct resources to complete its life cycle, and most plants require the same set of resources, although in subtly different proportions. Each of the resources has to be obtained independently of the others, and often by quite different uptake mechanisms – some as ions (potassium), some as molecules (CO_2), some in solution, some as gases. Carbon cannot be substituted by nitrogen, nor phosphorus by potassium. Nitrogen can be taken up by most plants as either nitrate or ammonium ions, but there is no substitute for nitrogen itself. In complete contrast, for many carnivores, most prey of about the same size are wholly interchangeable as articles of diet. This contrast between resources that are individually *essential* for an organism, and those that are *substitutable*, can be extended into a classification of resources taken in pairs (Figure 3.27).

In this classification, the concentration or quantity of one resource is plotted on the *x*-axis, and that of the other resource on the *y*-axis. We know

zero net growth isoclines

that different combinations of the two resources will support different growth rates for the organism in question (this can be individual growth or population growth). Thus, we can join together points (i.e. combinations of resources) with the same growth rates, and these are therefore contours or 'isoclines' of equal growth. In Figure 3.27, line B in each case is an isocline of *zero* net growth: each of the resource combinations on these lines allows the organism just to maintain itself, neither increasing nor decreasing. The A isoclines, then, with less resources than B, join combinations giving the same *negative* growth rate; whilst the C isoclines, with more resources than B, join combinations giving the same *positive* growth rate. As we shall see, the shapes of the isoclines vary with the nature of the resources.

3.8.1   Essential resources

Two resources are said to be *essential* when neither can substitute for the other. Thus, the growth that can be supported on resource 1 is absolutely dependent on the amount available of resource 2 and vice versa. This is denoted in Figure 3.27a by the isoclines running parallel to both axes. They do so because the amount available of one resource defines a maximum possible growth rate, irrespective of the amount of the other resource. This growth rate is achieved unless the amount available of the other resource defines an even lower growth rate. It will be true for nitrogen and potassium as resources in the growth of green plants, and for two obligate hosts in the life of a parasite or pathogen that are required to alternate in its life cycle (see Chapter 12).

Figure 3.27   Resource-dependent growth isoclines. Each of the growth isoclines represents the amounts of two resources (R_1 and R_2) that would have to exist in a habitat for a population to have a given growth rate. Because this rate increases with resource availability, isoclines further from the origin represent higher population growth rates – isocline A has a negative growth rate, isocline B a zero growth rate and isocline C a positive growth rate. (a) Essential resources; (b) perfectly substitutable; (c) complementary; (d) antagonistic; and (e) inhibition. (After Tilman, 1982.)
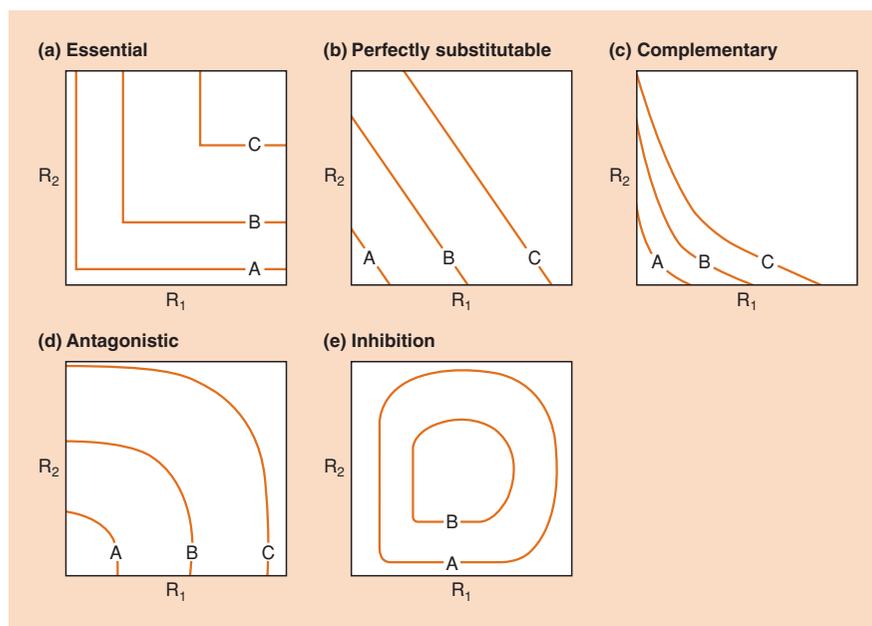
3.8.2   Other categories of resource

Two resources are said to be *perfectly substitutable* when either can wholly replace the other. This will be true for seeds of wheat or barley in the diet of a farmyard chicken, or for zebra and gazelle in the diet of a lion. Note that we do not imply that the two resources are as good as each other. This feature (perfectly substitutable but not necessarily as good as each other) is included in Figure 3.27b by the isoclines having slopes that do not cut both axes at the same distance from the origin. Thus, in Figure 3.27b, in the absence of resource 2, the organism needs relatively little of resource 1, but in the absence of resource 1 it needs a relatively large amount of resource 2.

complementary resources

Substitutable resources are defined as *complementary* if the isoclines bow inwards towards the origin (Figure 3.27c). This shape means that a species requires less of two resources when taken together than when consumed separately. A good example is human vegetarians combining beans and rice in their diet. The beans are rich in lysine, an essential amino acid poorly represented in rice, whilst rice is rich in sulfur-containing amino acids that are present only in low abundance in beans.

antagonistic resources

A pair of substitutable resources with isoclines that bow away from the origin are defined as *antagonistic* (Figure 3.27d). The shape indicates that a species requires proportionately more resource to maintain a given rate of increase when two resources are consumed together than when consumed separately. This could arise, for example, if the resources contain different toxic compounds that act synergistically (more than just additively) on their consumer. For example, D, L-pipecolic acid and djenkolic acid (two defensive chemicals found in certain seeds) had no significant effect on the growth of the seed-eating larva of a bruchid beetle if consumed separately, but they had a pronounced effect if taken together (Janzen *et al.,* 1977).

Finally, Figure 3.27e illustrates the phenomenon of *inhibition* at high resource levels for a pair of essential resources: resources that are essential but become damaging when in excess. CO_2, water and mineral nutrients such as iron are all required for photosynthesis, but each is lethal in excess. Similarly, light leads to increased growth rates in plants through a broad range of intensities, but can inhibit growth at very high intensities. In such cases, the isoclines form closed curves because growth decreases with an increase in resources at very high levels.

inhibition

3.8.3   Resource dimensions of the ecological niche

In Chapter 2 we developed the concept of the ecological niche as an *n*-dimensional hypervolume. This defines the limits within which a given species can survive and reproduce, for a number (*n*) of environmental factors, including both conditions and resources. Note, therefore, that the zero growth isoclines in Figure 3.27 define niche boundaries in two dimensions. Resource combinations to one side of line B allow the organisms to thrive – but to the other side of the line the organisms decline.