Phylogenetics: Theory and Practice of Phylogenetic Systematics, Second Edition

E. O. Wiley
Bruce S. Lieberman

WILEY-BLACKWELL

PHYLOGENETICS

PHYLOGENETICS

Theory and Practice of Phylogenetic Systematics

Second Edition

E. O. WILEY
BRUCE S. LIEBERMAN

CONTENTS

PREFACE TO THE SECOND EDITION

It has been over 25 years since the first edition of *Phylogenetics*. During that time, phylogenetic systematics has taken its place as the dominant paradigm of systematic biology and fundamentally influenced how scientists study evolution. Moreover, during the intervening years since that first edition, there have been many theoretical and technical advances and the field of phylogenetics has continued to grow. The great philosopher Marcus Aurelius's recognition that "time is a sort of river of passing events, and strong is its current" is doubly true in this area of scientific research. For instance, there are now new approaches to reconstructing the pattern of evolution designed to take character conflict and the uncertainty of phylogenetic estimates into account. The fallout from the molecular systematics revolution is a prominent part of this. Phylogeneticists have also moved beyond solely employing Hennig's argumentation schemes and now use more formal parsimony analysis or parametric methods such as likelihood and Bayesian inference in an attempt to reconstruct evolutionary relationships among organisms and find a fit between Earth history and descent with modification. We have tried to capture the essence of the evolving discipline that is phylogenetics in this new edition. If current trajectories imply anything, they suggest that the next 25 years of phylogenetic research will continue to prove exciting, with many fascinating theoretical and technical developments yet to come.

We also recognize that this disciplinary growth has not been without acrimony, and there have at times been battles waged between those advocates of parsimony analysis and those who argue for more statistical approaches to estimating phylogenies. We present the view here, however, that there is room for all of these approaches within the phylogenetic community. The principles used in these different approaches are closely similar. Relationship still means genealogical relationship, synapomorphy is still the mark of common ancestry, and monophyletic groups are the only natural groups regardless of whether one uses a parsimony algorithm or a likelihood algorithm to analyze one's data. That make us all phylogeneticists, and if you wish to use a label, it make us all Hennigians.

We have written this book for the practicing systematist and phylogeneticist. Our focus is on both philosophical and technical issues, and the philosophical issues discussed are those that we believe all working systematists need to address; these involve issues of the nature of species, the nature of characters, the nature of names, and the nature of biogeographic areas. While we cover what we think are the basics of parsimony, likelihood, and Bayesian analyses, we do not pretend that our coverage is more than basic. There are other texts, some highly mathematical, others less so, that cover these topics in more depth. We have tried to broadly cite this literature, at least up through 2009, but the field of parametric phylogenetics continues to advance faster than any one book can hope to capture without being out of date before publication. However, we hope that working systematists will be able to understand the basics we present and use these as an entré to a rapidly evolving discipline.

Over the long course of producing this second edition of *Phylogenetics*, we have greatly benefited from the comments of many colleagues. First and foremost are Mark Holder (University of Kansas) and Peter Midford (now at NEScent) who reviewed, page-by-page, most of the chapters dealing with taxa, characters, and methods of analysis. Mark Holder paid special attention to our chapter on parametric phylogenetics, patiently guiding us through much of the technical literature and attempting to keep us out of trouble in an area where we have no special expertise. We also gratefully acknowledge Mark Holder for his contributions in the area of biogeography. In particular, he helped figure out exactly how Modified Brooks Parsimony could be placed in a formal, algorithmic framework. We are very grateful for all of his insights and help. We also thank Norman MacLeod (Natural History Museum, London) for his insightful comments and suggested revisions on the subject of morphometric analysis. In addition, Francine Abe and Matthew Davis (University of Kansas) helped us understand morphometrics well enough to get a draft of this section to Norman. We thank John Wiens for his insights on missing data. We thank Dr. Randy J. Read for permission to use and adapt his examples illustrating maximum likelihood and Bayesian inference from WWW material that formed part of a course he taught at the University of Cambridge 1999–2000. Special thanks go to two philosophers of science for taking the time to consider philosophical issues with one of us (EOW). David Hull (Northwestern University) has always been willing to discuss issues of individuality and species. Elliott Sober (University of Wisconsin) kindly reviewed an earlier draft of the section on philosophy. We are also grateful to the folks at Wiley-Blackwell, especially our editor Karen Chambers, for helping to bring this project to fruition.

E. O. Wiley

Lawrence, Kansas   *Bruce S. Lieberman*

PREFACE TO THE FIRST EDITION

This is a book about systematics and how the results of systematic research can be applied to studying the pattern and processes of evolution. The past twenty or so years have seen tremendous changes in biological systematics. Although some of these changes have occurred because of the discovery of previously unobservable characters, the most profound changes have taken place on the methodological and philosophical levels. Systematists have become more critical about the methods they employ and the biological and philosophical bases for these methods.

The first half of this century saw evolutionary theory march ahead of systematics, but in a rather curious manner. Evolutionary theorists became disinterested in the pattern of organic descent and concentrated on various processes purported to occur on the populational level of analysis. This resulted in the generally accepted theory known as the Synthetic Theory of Evolution, or neo-Darwinism. In itself the neo-Darwinian theory is an admirable accomplishment. However, it is not enough. What is needed now is a better understanding of the origin of species, and, as Waddington (1957) says, why there are tigers and elephants and other such things. To approach such an understanding we must first have something to understand. This something is a phylogenetic tree, a pattern of organismic descent.

Phylogenetic systematics, or simply "phylogenetics," is not just another approach to systematics. It is an approach to systematics designed to estimate the pattern of phylogenetic descent that is needed to deduce the processes of evolution concerned with the origin of species. The classifications that result from phylogenetic analysis are critical tools for evolutionary studies. Phylogenetics is also more than the hand-maiden of evolution, however, for its underlying philosophy provides a way of viewing nature, asking questions and solving problems associated with the evolution of organisms.

I wrote this book to outline what I perceive as the philosophy and methodology of phylogenetics as a systematic discipline. As such, it is not restricted simply to the methods for reconstructing phylogenetic relationships and presenting these relationships in the form of a classification. Rather, it is also directed toward an

understanding of the evolution of species and the biological entities that comprise the history of descent with modification. Further, the phylogeneticist must also be a taxonomic scholar familiar with methods for dealing with specimens and characters, ways of assessing taxonomic literature, and various rules of nomenclature. These subjects are also dealt with.

Phylogenetic systematics is an approach to systematics that accomplishes an ordering of organic diversity in such a way that our ideas concerning the inferred evolutionary relationships among organisms can be scientifically discussed and evaluated. Much reaction has been directed toward this approach from its critics. I believe that most of this reaction stems from a lack of understanding of phylogenetics. My major purpose in writing this book was to clearly and simply present phylogenetic systematics (to the best of may ability) in the hope that others will understand its goals and methods. Only through understanding can profitable criticism and subsequent improvement follow.

Lawrence, Kansas   E. O. WILEY
May 1981

1

INTRODUCTION

Comparative biology has experienced a kind of renaissance over the last 40 years. This renaissance is the result of the development of techniques that allow us to reconstruct the evolutionary relationships, or genealogies, among organisms. Dobzhansky made the famous statement that nothing makes sense in biology except in the light of evolution. Phylogenetics has provided a tool that allows investigators to place their observations within the historical context of descent with modification and ferret out historical and proximal factors that contribute to their observations. Methods that explicitly test hypotheses of the descent of species have resulted in rigorously tested phylogenetic trees. These trees form the base knowledge for scientists that range from investigating macroevolutionary dynamics of speciation and extinction to demonstrating that a dentist in Florida was guilty of spreading the AIDS virus to his patients (Chin-Yih et al., 1992; Hillis and Huelsenbeck, 1994; see also Metzker et al., 2002, for another case). The historical impetus of the renaissance was the work of a German entomologist, Willi Hennig (1913–1976).

Before World War II, Hennig began developing what would come to be known as phylogenetic systematics. Hennig did not develop his ideas in a vacuum nor did all of his principles emerge in a single work (Richter and Meier, 1994). Hennig absorbed the influence of such workers as Haeckel, Zimmerman, and Neaf, and in fact, he was not the first to advocate many of the ideas that now form the basis for this approach to systematics. According to the analysis of Richter and Meier (1994), strict monophyly was central to Hennig (1950), but a careful distinction between apomorphy and plesiomorphy, as used in Hennig (1966), appeared in 1952 while the term *paraphyly* was not adopted until a 1960 manuscript (providing at least part of the basis for Hennig, 1966). Willmann (2003) provides another analysis of the historical context of ideas that led to Hennig's development of what we now know as

phylogenetic systematics. Not all of Hennig's ideas play a central part in how the discipline is practiced today. For example, although we provide a basis for showing that Hennig (1966) used outgroup comparison, it is certainly not made explicit in Hennig (1966). He did, however, outline a coherent program of systematic philosophy and inquiry and his work was fundamental to the eventual success of the discipline. His first synthesis, *Grundzüeiner Theorie der Phylogenetischen Systematik* (Hennig, 1950), outlined the basic goals, and his later English-language *Phylogenetic Systematics* (Hennig, 1966) contained five basic ideas that began a major revolution in systematics:

1. The relationships that provide the cohesion of living and extinct organisms are genealogical ("descent") relationships.
2. Such relationships exist for individuals within populations, populations within species, and between species themselves.
3. All other types of relationships (e.g., similarity, ecology) have maximum relevance when understood within the context of genealogical descent.
4. The genealogical descent among species may be recovered by searching for particular characters (evolutionary innovations, synapomorphies) that document these relationships. Further, not all of the similarities that arise through descent are equally applicable to discovering particular relationships; some are applicable at one level of inquiry while others are applicable at different levels of inquiry.
5. Of the many possible ways of classifying organisms, the best general reference system is one that exactly reflects the genealogical relationships of the species classified.

Kiriakoff (1959) was one of the first to discuss Hennig's ideas in some depth in American literature. Wider discussion of these ideas among English-speaking scientists began after the publication of Hennig's (1965) summary of his philosophy, the publication of the revised English edition of *Phylogenetic Systematics* (Hennig, 1966), and Brundin's (1966) seminal work on chironomid midges. Early English-language applications of Hennig's methods include Koponen (1968: mosses) and Nelson (1969: fishes). In fact, Gareth Nelson's energy and enthusiasm for Hennig's ideas were the major factors leading to the success of phylogenetics and Nelson's (1969) classification of higher vertebrates was the first modern American attempt to classify vertebrates within a phylogenetic context.

Hennig (1950, 1965, 1966, 1969, 1975, 1981, 1983, and other works) had many ideas other than the five basic points listed above. Some of these ideas remain basic to the discipline (e.g., monophyly, apomorphy, and plesiomorphy), while others seem to have been discarded (e.g., rank of a monophyletic taxon based on absolute geological age). Others have been refined (e.g., character argumentation to determine relative apomorphy and plesiomorphy). Some current phylogenetic applications might have seemed foreign to Hennig. For example, Hennig (1966) neither employed nor discussed formal algorithms that deal with character conflict and minimum evolution (e.g., parsimony algorithms) much less more statistical and model-based approaches such as likelihood point estimates of phylogeny and Bayesian inference of phylogenetic trees. Phylogenetics is a dynamic discipline. It grows and changes

to take advantage of and explore new approaches to the task of discovering the tree of life. Regardless of how it has changed, phylogenetics stands in stark contrast to its competitors, evolutionary taxonomy (Mayr and Ashlock, 1991) and phenetics (Sokal and Sneath, 1963; Sneath and Sokal, 1973), as we shall elucidate more fully below.

PHYLOGENETIC PROPOSITIONS

This book is an introduction to phylogenetic philosophy and techniques. It is founded on five propositions:

1. There is a tree of life that links all living organisms in a genealogical nexus, and it is possible to reconstruct relationships among the species that populate the tree.
2. Relationships among organisms do not have to be invented and treated as some form of scenario; they only have to be discovered. Our hypotheses reflect our best efforts to discover these relationships.
3. All characters are potentially useful in discovering these relationships, but only some characters are useful at any particular and restricted level of analysis.
4. Phylogenetic classifications are logically consistent with the phylogenetic tree advocated by the investigator. Thus, they are candidates for being natural classifications superior to alternatives that are not logically consistent with the phylogenetic tree hypothesis.
5. The relationships between hypothesis, evidence, and summary must be transparent in the sense that one can examine the evidence used in arriving at each piece of the puzzle.

Phenetics occupies the opposite end of the spectrum from phylogenetics. Early pheneticists were hopeful that if they could arrive at a measure of overall similarity between species this would be useful in showing the evolutionary relationships of those species, or perhaps higher taxa (Sokal and Sneath, 1963). When this proved not to be the case, they largely abandoned the search for evolutionary relationships in favor of a system of grouping taxa by overall similarity. Evolutionary taxonomy occupied an intermediate position. Post-Hennigian evolutionary taxonomists largely adopted the methods of phylogenetic analysis advocated by Hennig (e.g., Mayr and Ashlock, 1991). However, they continued to assert that classifications could and should express a balance between overall similarity and genealogical relationships. While this sounds reasonable, we shall see that the methods of striking this balance were often arbitrary and result in illogical classifications if they contain nonmonophyletic groups.

Evolutionary taxonomy is the oldest of the three approaches we have discussed thus far. It is reflected in the work of some systematists to integrate classification and taxonomy into the Neo-Darwinian Synthesis that began in the 1920s, resulting in classic works by Ernst Mayr, George Gaylord Simpson, and Julian Huxley.

In essence, evolutionary taxonomists sometimes coupled Linnean rank (Order, Class, etc.) with some measure of how distinctive a group might appear. Perhaps the

ultimate expression of this practice was Julian Huxley's proposition that humans, as reasoning animals, should be accorded their own grade phylum (Psychozoa). One challenge to this arbitrary, hybrid system came from Hennig (1950), but his work, in German, was largely overlooked in the English-speaking world. Sokal and Sneath (1963) discussed Hennig's ideas, and Simpson (1961) commented on them, but Hennig's system was largely ignored by the majority of systematists. The second challenge to evolutionary taxonomy came from the pheneticists in the mid-1950s. Early pheneticists perceived a lack of rigor and scientific testability in evolutionary taxonomy and phenetics was an attempt to produce a more "operational" and repeatable form of systematic inquiry. The phylogeneticists entered the fray in earnest in the late 1960s, challenging both pheneticists and evolutionary taxonomists (e.g., Schlee, 1968, 1971; Nelson, 1971a, 1972a, b, c, 1974a, b; Kavanaugh, 1972; Cracraft, 1974; Wiley, 1975, 1976; Farris, 1977, 1980; Mickevich, 1978) with equal vigor.

Phenetics, as a systematic discipline, has largely disappeared from the playing field. It left a positive legacy in fostering the use of computers in systematic analyses and in the use of certain multivariate statistical techniques and the field of geometric morphometrics. Evolutionary taxonomy, as a program of systematic inquiry, has also largely disappeared. However, its legacy lives on in numerous textbooks in the form of classifications that contain groups whose existence is based on criteria other than common ancestry, and in this respect, its legacy is negative.

The major purpose of this book is to continue the work begun in the 1981 edition of *Phylogenetics* (Wiley, 1981a). Now, as then, we do not claim that all phylogeneticists will agree with our perceptions of phylogenetic research. The past 40 years have seen tremendous advances in both the theory and practice of phylogenetic systematics, but the basics have remained largely the same.

1. Biological diversity has been generated by microevolutionary processes and by speciation. Speciation includes a number of modes of lineage splitting as well as hybridization and (early in life's history) symbiosis. Character modification may be coupled with speciation, cause speciation, or proceed independently of speciation.

2. The historical course of evolution comprises both a continuum of genealogical descent at the level of individual organisms and a discontinuum caused by speciation and resulting in a hierarchy of species. In the absence of special creation or ongoing spontaneous generation, all organisms show a historical continuum through descent. Thus, species that appear to be very different from each other are related, given that life itself has a single origin. Discontinua (establishment of independently evolving lineages) at the level of species are the reasons that both species and higher taxa are parts of the natural world. That is, both species and higher taxa that are truly monophyletic groups are real, not nominal. We discover the relationship between the continuum and discontinuum when we can reconstruct parts of the tree of life and observe largely hierarchical relationships between species and clades.

3. A phylogenetic tree (Fig. 1.1a) is a graphic representation of the historical course of speciation. In the phylogenetic system, this is true even for phylogenetic trees populated only by higher taxa because every natural higher taxon is founded by a single species. Lines/edges are single lineages or a monophyletic group of lineages represented by their ancestor. Vertices/nodes are specia-
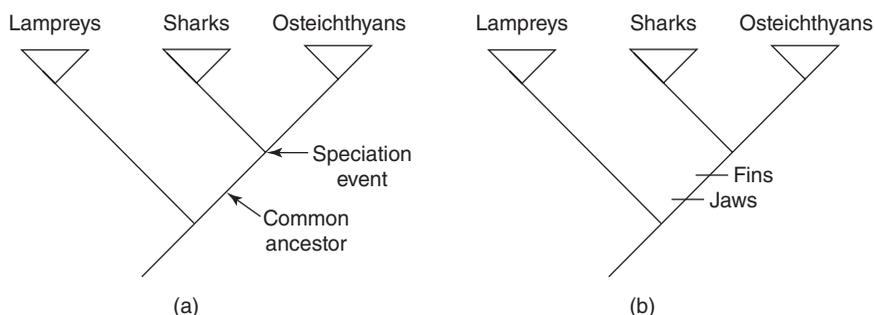
Figure 1.1. Two phylogenetic trees showing the relationships between lampreys, sharks, and osteichthyans (bony fishes and tetrapods). (a) The hypothesis of relationships. The node labeled "speciation event" is the speciation event that led to sharks (and kin) in one lineage and osteichthyans in the other lineage *relative to* lampreys. The edge labeled "common ancestor" represents at least one common ancestor shared by sharks (and kin) and osteichthyans not shared by lampreys. (b) Two evolutionary novelties (synapomorphies) that support the hypothesis that sharks and osteichthyans share a common ancestor not shared by lampreys. In both trees, the triangles denote that each clade is a group of two or more species, not a single species.

tion events. If we could discover it, a true phylogenetic tree of species is both necessary and sufficient to portray the history of evolution on both the specific and supraspecific levels of biological organization. On the empirical level, a hypothesis of relationship of species is necessary and sufficient to present the historical hypothesis of the investigator. Thus, confirmed trees are associated with confirming characters in the form of evolutionary novelties that are shared by the descendants of particular ancestral species (Fig. 1.1b). There are different ways to portray the tree, as we shall discuss in Chapter 4. Further, not all trees are phylogenetic trees; any acyclic graph is a tree, and many such graphs may portray phenomena such as gene evolution or even the relationship among geographic areas. Finally, some graphs are not trees at all, but cyclic graphs that may portray reticulate relationships.

4. Phylogeneticists attempt to recover parts of the tree of life through a comparative study of the similarities and differences of organisms.

5. The history of speciation may be recovered when speciation is accompanied by character change under certain conditions. In the simplest cases, such conditions obtain when the rate at which characters originate and are fixed keeps pace with lineage splitting and thus become candidates for documenting the lineage splits (Fig. 1.1b). The essence of the method is to search for characters that are indicative of unique common ancestry. These characters are the evolutionary innovations, or apomorphies, that are hypothesized to have evolved in that ancestor alone and to have passed on to the descendants of that ancestor where they act as historical markers, synapomorphies, of the common ancestor itself. In the phylogenetic system, the presence of these evolutionary innovations is considered *prima facie* evidence for the existence of the ancestor. The conditions under which character evolution will lead to erroneous histories is partly understood and will be discussed in appropriate sections.

The point is that phylogenetic systematics is not an infallible system of inquiry; it has its limits just as all research programs have limits.

6. Hypotheses about relationships among organisms are meant to estimate the true phylogenetic tree that exists in nature at an appropriate level of complexity. As such, tree hypotheses are not merely devices to efficiently explain the distribution of characters. Rather, they are meant to place character evolution in an explicit historical framework where the validity of the conclusions can be accepted or debated. In systematic studies, the appropriate level is usually the level represented by species or monophyletic groups of species. The fact that there is only one true tree at this level of complexity provides the basis for testing alternative hypotheses. If two hypotheses are generated for the same group of species, then we can conclude that at least one of these hypotheses is false. Of course, it is possible that both are false and some other tree is true.

7. Hypotheses of relationships convey only relative assertions about those taxa that are known to the investigator and analyzed by the investigator. For example, if we assert that chimpanzees are more closely related to humans than to gorillas, we are not claiming that there is only one ancestor shared by chimps and humans or that chimps are the only close relatives to humans, only that there is at least one ancestor shared by chimps and humans that is not shared with gorillas.

8. The major purpose of phylogenetic classification is to condense and summarize the inferred history of speciation as reflected by our best hypotheses of the history of speciation in a manner that is logically consistent with the phylogenetic tree. This summarization consists of a vocabulary of the names of species and monophyletic groups arranged in such a manner as to either reflect, or at least be consistent with, the underlying history of speciation.

TOPICS COVERED

The remaining part of this chapter is concerned with definitions of some basic terms, the relationship between phylogenetic systematics and other areas of science, and a brief introduction to the philosophy of systematics.

A major part of this book deals with ontological issues. Ontological issues are important because to not understand the ontological status of species, for example, is to not understand much about species at all. Thus, in Chapter 2, we develop the ontological concept that species are individuals (Ghiselin, 1966; Hennig, 1966), and we explore various species concepts, settling on the Evolutionary Species Concept as most useful in phylogenetic research. Supraspecific taxa are dealt with in Chapter 3 as both individuals and the natural units of phylogenetic classification. Although some have suggested that the concept of *natural higher taxon* has lost its meaning, we will suggest that phylogenetics provides a basis for just such a concept; it is the monophyletic taxon of Hennig (1966).

After developing concepts about the entities of phylogenetic research, we turn, in Chapter 4, to a consideration of phylogenetic trees. Hennig (1966) provided some fundamental insights into the nature of trees, and it is important to understand the

biological meaning that is contained in the very simplified trees that are the end product of phylogenetic research. A good part of the chapter is devoted to understanding the differences between different forms of phylogenetic trees. These differences are fundamental to understanding what we can infer from character analysis about evolutionary patterns.

Chapter 5 deals with characters. In that chapter we attempt to develop a concept of characters as properties of individual organisms and shared characters as properties of groups (groups both real and unreal in nature, which will correspond to homologies and homoplasies, respectively). We will also explore the concept of homology, reviewing some of the history of the concept and how current phylogenetic techniques are used to test propositions that character matches are homologs and how we connect different matches into transformation series.

Chapters 6 and 7 cover the basics of phylogenetic analysis. We begin with parsimony techniques (Chapter 6) and proceed to likelihood and Bayesian techniques (Chapter 7). Part of our agenda is to show that parsimony and likelihood are not so different and that it is possible to understand the relationship between these two seemingly different approaches to character analysis.

Chapter 8 is devoted to phylogenetic classification and the various issues of the meaning of taxonomic names. Included in this chapter are discussions of various approaches to phylogenetic classification, the logical relationship between classifications and phylogenetic trees, and the presentation of various conventions that may be used in the Linnean system. We then discuss the merits of the PhyloCode and contrast its claims and assumptions with those of the more traditional codes.

In the first edition, Wiley devoted an entire chapter to the alternative "schools" of evolutionary taxonomy and phenetics. But that was over 20 years ago, and there is little need for such a chapter. Instead, we devote Chapter 9 to biogeography. We consider the historical development of the field, while elucidating different biogeographic processes such as dispersal, vicariance, and geodispersal. Moreover, this chapter includes a discussion of various analytical methods in biogeography, their relative strengths, and how to implement them. Finally, we consider how extinction affects our ability to retrieve biogeographic patterns and the importance of biogeography for our understanding of past mass extinctions and the current biodiversity crisis.

The remaining two chapters are devoted to practical matters. Chapter 10 is devoted to specimen selection, field collecting, and curation, with an emphasis on modern data mining. The book ends in Chapter 11 with a consideration of systematic publication, the use of literature, the making of keys, a brief discussion of the Linnean code, and other issues that phylogeneticists must understand to practice taxonomy.

TERMS AND CONCEPTS

Phylogenetic systematics, like any other scientific discipline, has its own peculiar lexicon of terms and its own particular definitions that at times mean something different outside the discipline. Here, we introduce some basic terms and concepts as they are used in the book. Others will be introduced at various times when appropriate.

Disciplines

1. *Comparative Biology.* Nelson (1970) divided biology into two basic areas. He held that general biology was concerned with investigating biological processes while comparative biology was concerned with investigating biological patterns, and we concur with aspects of this definition. In general biology, the investigator picks organisms that are most likely to be amenable to studying a particular process of interest to them. In comparative biology, the investigator is interested in studying the characteristics of diverse organisms to infer the historical, evolutionary relationships between these organisms. For example, an ethnologist working in the realm of general biology is interested in the mechanistic explanation of a particular stimulus-response reaction. By contrast, the ethnologist working in the realm of comparative biology is interested in how common that stimulus-response reaction might be among organisms and how that stimulus-response reaction has evolved through time. In particular, he or she would be interested in determining if that response to stimulus evolved once or repeatedly. Phylogenetic systematics, like other systematic disciplines, is one comparative approach. The phylogeneticist is interested in estimating the pattern of organic diversity and thus the historical course of evolution. Any and all comparative data are potentially useful in this pursuit, and any and all comparative information can, in theory, be accommodated.

2. *Systematics.* Systematics is the study of organic diversity as that diversity is relevant to some specified pattern of evolutionary relationship thought to exist among the entities studied. This definition is somewhat narrower than others (e.g., Mayr, 1969; Nelson, 1970), which held systematics synonymous with comparative biology. From our perspective, not all comparative biologists practice systematics, even though all comparative data can be accommodated by systematics. For example, comparative physiologists may not analyze their data phylogenetically, but their data can be incorporated into a phylogenetic analysis or better understood by mapping it onto a well-confirmed phylogenetic tree.

3. *Taxonomy.* Taxonomy comprises the theory and practice of describing, naming, and ordering groups of organisms termed *taxa*. How the taxa are ordered into classifications defines the particular approach to taxonomic classification. The rules for naming are outlined in various Codes of Nomenclature, and these codes are now being challenged in new ways by those who seek to redefine taxonomy. This definition differs from some authors (e.g., Simpson, 1961) who equated taxonomy with systematics.

4. *Phylogenetic Systematics.* This is one approach to systematics and taxonomy that attempts to recover the phylogenetic relationships among taxa and in which formal biological classifications are consistent with these relationships. We refer to the discipline as phylogenetics and to those who practice it as phylogeneticists. Another common set of terms is *cladistics* and *cladists*. We do not object to these terms (first coined by an opponent, Mayr, 1969). However, we suggest that it originally implied a preoccupation with branching pattern and a de-emphasis on character evolution, neither of which is true. Indeed, recovering the pattern of character evolution reveals the pattern of branching

and speciation. The goal of phylogenetics is to give a complete account of speciation and character evolution.

Organisms and Grouping of Organisms

1. *Taxon.* This is a grouping of organisms at the level associated with the application of proper scientific names, or a grouping of such organisms that could be given such a name but is not named as a matter of convention. The plural is taxa. Some taxa (the natural ones) are considered to have an objective reality in nature apart from our ability to find and name them. Taxa in practice are groups named by systematists. As such, they are hypotheses about taxa in nature. As hypotheses, they may be accepted or rejected based on subsequent research, or even on logical grounds. For example, phylogenetic systematists reject paraphyletic taxa on logical grounds because such taxa result in classifications that are inconsistent with an accepted phylogenetic tree (Wiley, 1981b). Higher taxa are taxa that include more than one species. Species taxa are the lowest formally recognized taxa usually considered in phylogenetic analysis.

2. *Monophyletic Group.* A monophyletic group is a taxon comprised of two or more species that includes the ancestral species and all and only the descendants of that ancestral species (Fig. 1.2a). *Monophyletic group* is usually considered synonymous with the term *clade*, and the two terms are frequently used interchangeably. As used here, species are not monophyletic groups because they are self-referential entities of process while monophyletic groups are neither self-referential nor units of process, except the process of descent. Instead, they are entities of history. Monophyletic groups in nature are real, but again monophyletic groups named by systematists are hypotheses, and these hypotheses stand or fall on the empirical evidence.

3. *Para- and Polyphyletic Groups.* Paraphyletic groups are incomplete groups in which one or more of the descendants of the common ancestor are not included in the group (Fig. 1.2b). Invertebrata is an example, as are Reptilia (birds and mammals excluded) and Pongidae (*Homo* and allied fossil genera excluded). Polyphyletic groups are comprised of descendants of an ancestor not included in the group at all. Homothermia (birds + mammals) would be an example as
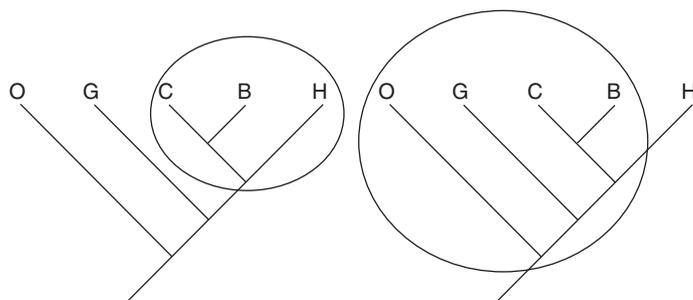
Figure 1.2. Concepts of monophyly and paraphyly. (a) A monophyletic Hominidae that includes humans (H), chimpanzees (C), and bonobos (B). (b) A paraphyletic Pongidae that includes orangutans (O), gorillas (G), chimpanzees, and bonobos but excludes humans.
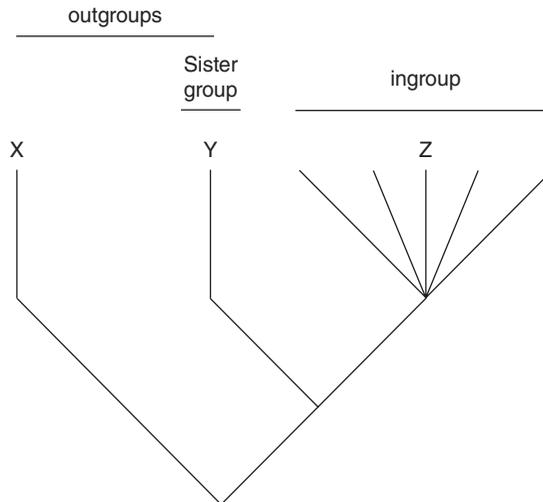
Figure 1.3. Some terms for groups used in a phylogenetic analysis. Relationships of outgroups to the ingroup are shown as "known" as a matter of prior knowledge, backed up with empirical data.

the ancestor of birds and mammals would presumably be included in Reptilia. Para- and polyphyletic groups are not real in nature. From the phylogenetic perspective, paraphyletic and polyphyletic groups named by systematists are illogical, either through ignorance (group named in the absence of a phylogeny) or practice (as in evolutionary taxonomic practice for naming paraphyletic groups).

4. *Sister Group.* In nature, a sister group is a single species or a monophyletic group that is the closest genealogical relative of another single species or monophyletic group of species (Fig. 1.3). True sister groups share a unique common ancestral species—an ancestral species not shared by any other species or monophyletic group. In phylogenetic analysis, a sister group is the hypothesized closest *known* relative of a group the investigator is analyzing, given current knowledge. Hypotheses of sister group relationship are fundamental to phylogenetic practice. In analyses, the sister group is the most influential outgroup for determining the relative merit of presumed homologies to indicate genealogical relationships within the group studied, as outlined in Chapter 6.

5. *Outgroup.* An outgroup is a species or higher taxon used in phylogenetic analysis to evaluate which presumed homologs indicate genealogical relationships within the group studied and which are simply primitive characters (Fig. 1.3). The outgroup is used to root the tree and determine character polarity. The sister group is a special-case outgroup. Critical analysis requires the investigator to consult both the sister group and at least one additional outgroup to make the determination about homologs.

6. *Ingroup.* The ingroup is the group that is being analyzed by the investigator. It is shown in Fig. 1.3 as a polytomy because relationships within the group

are unresolved before an analysis. Other graphic devices show the ingroup as a triangle.

Phylogenetic History and Evolution

1. *Relationship.* In the phylogenetic system, relationship means genealogical relationship. Justifications for hypothesizing relationships cannot be made by appeal to similarity alone, only by appeal to similarity as similarity relates to common ancestry. Does this similarity indicate that the taxa share a unique common ancestor relative to the other taxa studied? If so, then similarity is vital to the question at hand. If not, then the similarity is not vital to the question at hand (but might be to other questions). All entities (things that exist in the world) share properties and thus have relationships through these properties. The entities most relevant to phylogenetic systematics are organisms and groups of organisms. On the empirical level, this reduces to specimens examined and inferences (hypotheses) that these specimens and their properties represent entities of taxonomic interest, taxa. In the phylogenetic system, two taxa are related if they share a common ancestor. If life has a single origin, then all taxa are related, but this truism does not get us very far. Because all taxa share a common ancestor at some level, relationship is usually presented as a comparative statement involving at least three taxa. A is more closely related to B than to C if, and only if, A and B share a common ancestor not shared by C.

2. *Genealogy and Genealogical Descent.* Given evolution, genealogical descent exists in nature apart from our ability to discover it. Empirically, a genealogy proposed by a phylogeneticist is a graphic representation of a hypothesis of the descent relationships of one or more organisms from one or more ancestors. Pedigrees are genealogies on the level of individual organisms. Phylogenetic tree graphs are genealogies on the level of populations, species, and higher taxa. All trees graphs are divergent, as in the case of clonal organisms and most metazoan taxa. Cyclic graphs, frequently termed *reticulate trees* or *networks*, are not trees in the graph theoretical sense. They portray reticulate relationships, as in pedigrees of sexually reproducing organisms or species that originate via reticulate speciation. A graph with a single reticulation is not technically a "tree," although most systematists forgo the formalities of graph theory and call them trees.

3. *Cladogenesis.* Cladogenesis is branching, divergent evolution (Fig. 1.4). At the level of species, a cladogenetic event results from one of an array of speciation mechanisms that results in two or more species where only one species existed before the event. Populations within species may also diverge, creating geographic variation and a polytypic species. However, the local differentiated populations are not thought to represent independent evolutionary lineages because of ongoing (even if rare) gene flow.

4. *Anagenesis.* Anagenesis is a synonym of phyletic evolution, and these terms can be used interchangeably. Anagenesis refers to evolution within a lineage through population genetic phenomena (mutation, selection, drift, etc.). Over time, anagenesis leads to divergence between closely related species, whether
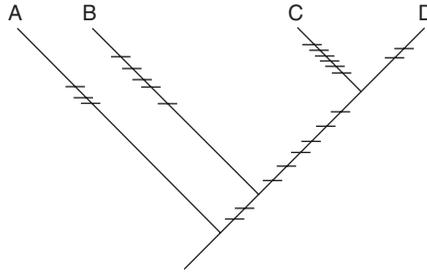
Figure 1.4. Cladogenesis and anagenesis. Each branching event (speciation event, node) is a cladogenetic event. Three such events are shown. Each tick mark represents "fixation" of an evolutionary novelty, and the number of such novelties is the mark of anagenesis, the evolution of characters along an evolving lineage. Note that in this diagram anagenesis proceeds at different rates along different lineages. For clarity, there has been no taxic extinction in this hypothetical clade and the number of novelties is proportional to all changes.

the time period is short or long, and evolution is episodic or continuing throughout the history of the lineage. The amount of anagenesis shown in the sample of characters and taxa on a tree may be graphically displayed by showing the number of changes that occurred between cladogenetic events, as in Fig. 1.4, or by making the edges longer in proportion to the number of such changes.

5. *Speciation.* This is an array of processes leading to the origin of one or more new species. Speciation may be cladogenetic (e.g., lineage splitting) or reticulate (e.g., speciation via hybridization), but it does not happen due to anagenesis alone.

6. *Speciation Event.* The historical result of speciation, a speciation event refers to a particular and historically unique event for the ancestral species in question. No particular time frame is associated with the term, thus speciation may be instantaneous or protracted. In the phylogenetic system, the origin in time of two sister species is considered to be identical regardless of the length of the speciation event. Thus, sister species and sister groups have the same time of origin.

7. *Vicariance Event.* This is a geographic separation of a once continuous biota such that the biota becomes two or more geographically separated biotas. For any particular species, a vicariance event may eventually result in complete speciation, semi-isolated populations that exhibit geographic variation, or may have no apparent evolutionary effect on the geographically separated populations. This is because the vagility of organisms is not uniform over all taxa in a biota. Further, the response to a vicariance event may differ among taxa because some taxa diverge more slowly than others. Thus, the long-term outcomes of vicariance events cannot fully and always be predicted for each and every species in the biota. However, in the long term it is expected that if a vicariance event truly divides the preexisting geographic range of a biota, eventually many of the component species affected will undergo differentiation and speciation.

Attributes of Organisms

1. *Character.* A character is a property of an organism. A character state is a feature, attribute, or observable part of an organism as interpreted by an investigator. Phylogenetically informative characters come in two or more states. Characters constitute those properties of organisms studied by systematists. Empirically, a character state is a part or attribute of a specimen that may be described, figured, measured, weighted, counted, scored, or otherwise communicated by one biologist to another. Characters gain their legitimacy through heritability, and character states gain legitimacy as other biologists see the character and the acceptance by others that the character state represents a legitimate "factorization" (decomposition into parts) of the specimen for purposes of description. Of particular interest to systematists is the question of whether two character states have different evolutionary origins and the extent to which they are free to vary independently (Wagner, 1996; Wagner and Stadler, 2003). Characters and character states are usually arrayed in a data matrix. The character constitutes a column of data, and the various states fill the cells (Fig. 1.5).

2. *Match or Character Match.* As used by Sober (1988), a match is a shared character state. More specifically, if two or more organisms are hypothesized to share a particular character state, the state is coded with the same symbol or assigned a common name. The presumption is that the shared state is a good candidate for being a shared homolog, although some matches turn out to be homoplasies or even analogies (each defined more fully below). *Match* roughly corresponds to the term *primary homolog* as introduced by de Pinna (1996). Empirically, character matches are coded with the same symbol and placed in the same data column (Fig. 1.5).

3. *Evolutionary Novelty.* An inherited change from a previously existing character state, the novelty is the transformational homolog of the preexisting character state. Phylogeneticists are most interested in novelties that become "fixed" (frequency near 100 percent excluding atavisms and back mutations),

	Paired Appendages	Branchial Arch	Internal Skeleton
Lamprey	None	Gill Arch	Cartilage
Shark	Fins	Jaws	Cartilage
Salmon	Fins	Jaws	Bone
Human	Legs	Jaws	Bone

(a)

	1	2	3
Lamprey	0	0	0
Shark	1	1	0
Salmon	1	1	1
Human	2	1	1

(b)

Figure 1.5. Two simple character matrices. The upper matrix expresses characters and their states in words. The lower matrix expresses the same characters and states as numerical codes.

although polymorphic characters can be easily analyzed with modern phylogenetic algorithms. All homologs begin their existence as evolutionary novelties. Further, the term is tied to a specific genealogical context. Independent origin of two highly similar character states results in two evolutionary novelties, not one. However, the conclusion that a match is actually two independent evolutionary novelties can only be a conclusion drawn from a phylogeny that is well corroborated by other characters leading to the conclusion of independent origins.

4. *Taxic and Transformational Homologies.* We will discuss the concept of homology in greater detail in Chapter 5. Taxic homologs are character states shared by two taxa and are the same state inherited from a common ancestral species. Empirically, taxic homologs are state matches that appear on a phylogenetic tree in the common ancestor of specimens (taxa) that have the character. Transformational homologs are different states, one state being the historical precursor of the other. Two (or more) homologs form a transformation series. One state is an evolutionary novelty that originated in an earlier common ancestor and diagnoses a larger monophyletic group. The other state(s) is a modification of the genetic and epigenetic information of the older homolog and diagnoses a monophyletic group included within the larger group. For example, in Fig. 1.5, pectoral fins are an evolutionary novelty of gnathostomes (jawed vertebrates) and front legs are an evolutionary novelty of tetrapods, a group nested within gnathostomes. Front legs are modified fins. Two (or more) homologies in a transformation series have relative relationships in the tree of life. The more ancient homology is termed a *plesiomorphy*. Two or more species that share this more ancient novelty share a symplesiomorphy. The other character state that is shared by members of a more restricted monophyletic group nested within the larger group is termed an *apomorphy*. Two or more taxa that have this character state share a synapomorphy. All symplesiomorphies at one restricted level of the entire tree of life are synapomorphies at one or more higher levels where they diagnose monophyletic groups that continue to exist at the time of the origin of the new, apomorphic homolog. Empirically, transformational matches are coded as different symbols in the same data column and transformational homologs confirm nested monophyletic groups. For example, states "gill arch" and "jaw" in Fig. 1.5 are hypothesized transformational homologs and thus a character pair comprised of hypothesized plesiomorphic and apomorphic homologs, with the evolution of one pair of gill arches to one pair of jaws occurring sometime between the origin of lampreys and the origin of the common ancestor of sharks and osteichthyans (Fig. 1.6). (Note that this is a relative hypothesis; there are other, fossil, taxa involved that are not shown.)

5. *Other Kinds of Homology.* Haszprunar (1992) has suggested a hierarchy of homologies, including iterative homology, ontogenetic homology, and polymorphic homology. We will discuss these distinctions in Chapter 5.

6. *Homoplasy.* Homoplasy is similarity achieved by independent evolution in different parts of the tree of life (Lankester, 1870). Homoplasies have different evolutionary origins and thus represent different (albeit similar) evolu-
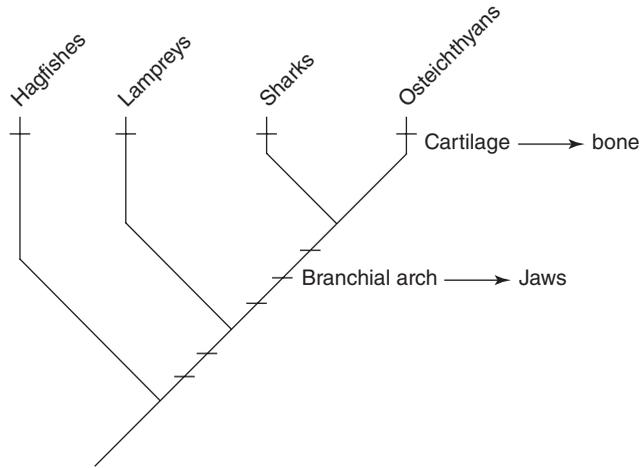
Figure 1.6. Relationships among some chordate animals. Note that the transformation of an anterior pair of gill arches to jaws is hypothesized to have been completed some time after the origin of lampreys but before the speciation event that gave rise to sharks and osteichthyans. Exactly when this happened in real time and whether the transformation occurred in a single ancestral species or over many species and speciation events cannot be determined using this tree. In other words, the amount of anagenesis and cladogenesis involved in the transformation of gill arches to jaws is not known.

tionary novelties. The terms *parallelism* and *convergence* are used frequently, although Eldredge and Cracraft (1980) refuted the notion that there was any concrete distinction between the two. Patterson (1982, 1988) provided a formal criteria for separating convergence from homoplasy, and we discuss this in Chapter 5.

7. *Analogs.* In its original context, *analogy* referred to organs that perform similar functions, whether they were homologous or not (Panchen, 1994). Today, analogous structures are usually taken to be those with very dissimilar structure but similar function, as in the wings of insects and birds.

8. *Holomorphology.* The holomorphology of an organism is the total spectrum of characters exhibited by that organism during its lifetime: its character properties. The holomorphology of a species is the sum of all the holomorphologies of its parts (organisms).

9. *Epiphenotype.* This is the characters of an organism at any particular time it is inspected during its life. This term is largely synonymous with the term *phenotype* for morphological characters, but includes the connotation that the epiphenotype is the result of an array of genetic and ontogenetic processes.

Classification

1. *Classification.* A series of words used to name and arrange organisms according to some principle of relationship thought to exist among the organisms.

Most formal taxonomic classifications are usually Linnean classifications formulated according to rules embodied in codes of nomenclature that have been adopted by international agreement.

2. *Phylogenetic Classification*. A classification that presents the genealogical relationships hypothesized to exist among a given array of organisms. Phylogenetic classifications have the property of being logically consistent with the hypothesized phylogeny of the organisms. As we shall see in Chapter 8, competing systems may not have this property.

3. *Category*. A category is any one of an array of rank nouns used to denote relative subordination of taxon names in a Linnean classification. Assigning a particular rank to a taxon has the effect of subordinating that taxon in a classification hierarchy. Particular ranks are a kind of category and may be used repeatedly. However, ranks have only relative and not absolute meanings in the phylogenetic system. Because they have only relative meaning, being a genus of rose plants does not have the same connotation of biological organization or characterization as being a genus of fishes. In the phylogenetic system, only sister groups are guaranteed to be comparable in terms of evolutionary history or biological meaning. The only exception to this principle are taxa ranked as species. Species, as units of process, may be compared directly. The following is an abbreviated list of categorical ranks used in this book:

> Kingdom
> Series
> Phylum (Zoology) or Division (Botany)
> Class
> Division (Zoology only)
> Cohort
> Order
> Family
> Tribe
> Genus
> Species

The Linnean Hierarchy is only one of several systems for translating a phylogenetic hypothesis into a phylogenetic classification. We will discuss the major alternatives, including unranked and numerically ranked classifications, and the newly proposed PhyloCode. Finally, it is important to understand that categorical ranks are kinds and not taxa. When we refer to a family, we are referring to a particular taxon ranked as a family and not to the categorical rank of family.

PHILOSOPHY AND SYSTEMATICS

Two broad areas of the philosophy of science impinge upon systematists. The first, ontology, is concerned with the meaning of concepts, things, entities, etc. The second,

epistemology, is concerned with how we acquire knowledge and justify hypotheses about these things and their relationships. For example, issues of whether the name of a species refers to an individual or a natural kind is an ontological issue while the issue of what constitutes evidence for hypothesizing that a particular collection of individual organisms comprises one or two species is an epistemological issue. One may depend on the other, as we shall see.

The first issue faced by systematists as an example of this dependency concerns the form of systematic hypotheses. Hull (1983), in response to the growing attachment of phylogenetic systematists to the philosophy of Karl Popper (e.g., Wiley, 1975), outlined the relationships between the ontological status of taxa and adopting a particular ontology in terms of the form of hypotheses we test. Hull recognized five sorts of hypotheses.

"All A are B." This hypothesis is in universal form. It is meant to apply universally over time and space. Such a statement has the potential to be easily falsified, but it cannot be completely verified. The proviso "potential" is important because the statement actually takes a more complicated form, as discussed below, and because there is always the possibility of experimental or observational error. Nevertheless, we can say that there is an asymmetry between evidence that confirms and evidence that disconfirms the hypothesis. In spite of hundreds or millions of confirming observations, only a single *valid* disconfirming observation can render the hypothesis false. For example, the hypothesis "all tetrapod adults have front legs" can be rendered false with the discovery of a single snake (or any tetrapod gastrula for that matter).

"Some A are B." This hypothesis is also in universal form. It simply states that of the many instances of B at least one A exists that is also B. This claim is easy to confirm; all one needs to do is show a single example. However, it is impossible to completely disconfirm in practice because one would have to find all Bs and show that none are As. For example, the hypothesis "some tetrapods lack limbs" could be easily verified by finding a snake, but it could not by completely falsified unless one could observe all tetrapods, living, dead, and future, to see that none lacked legs. There is an asymmetry between confirmation and disconfirmation, but this time it works in the opposite direction. Confirmation requires only a single valid observation, but hundreds and millions of disconfirming observations fail to render the hypothesis false.

"All A are B in 1970." This hypothesis is termed a numerical universal. It is in universal form but with a restriction: in this case the restriction is a time period (1970). Hull's example was "All justices of the Supreme Court of the Unites States of America in 1970 were males." Such numerical universals can, in principle, be as easily confirmed as disconfirmed, and the asymmetry between confirmation and disconfirmation is absent.

"Some A are B in 1970." This hypothesis is a numerical particular. Like the numerical universal, it is, in principle, as easy to confirm as to disconfirm because a single instance will confirm and a finite number of observations will disconfirm. Scientists (and philosophers) are not much interested in this form of numerical universal hypothesis.

Singular Hypotheses. There are also hypotheses in singular form. Hypotheses such as "Ed Wiley is a male" concern a particular entity and claim that the entity (Ed Wiley) has or lacks the properties of maleness. Given that we can

agree on the properties of maleness, the statement is as easily confirmed as disconfirmed.

It is exactly this problem, of establishing the properties of being a male, where ontology is important. What do we mean when we say that someone is a male? Is *male* a kind that is associated with properties and thus has an intentional meaning? Is *male* a set whose definition is extensional? Indeed, is *Ed Wiley* an entity or simply a set of cells? Such questions arise regularly in systematic philosophy, and we shall examine these controversies throughout the book.

Wiley (1989) suggested that the form of the hypotheses encountered in systematic research and the way they are tested is closely tied with the ontology of the things systematists study. Hull (1983) and Sober (1993) have reached similar conclusions. Hull (1983) points out that most scientists are seeking truly universal hypotheses, the kind where disconfirmation is more important than confirmation. Singular statements are important because they function in the tests applied to hypotheses in universal form. For example, if we are to test the proposition that most speciation involves the geographical subdivision of an ancestral species, we need singular examples of species pairs to test the proposition. If we can examine a sufficient number of speciation events, we might be able to extrapolate and reach the conclusion that the majority of species are formed through geographic subdivision. Or we might reject that hypothesis and conclude the opposite. Wiley (1989) suggested that the reason such hypotheses in universal form take a predominant role in science is that they are directed toward testing process theories where entities are important only to the extent that they have or lack the properties predicted of them by a process theory. These properties are embodied in the intentional definitions of kinds that are inherent in the theory. As Hull (1981:184) puts it:

> Many criteria have been suggested to mark the distinction between genuine natural kinds and mere aggregations, none of them totally successful. The criterion that I think holds out most promise is figuring in a genuine law of nature. Any kind term that appears in a law of nature is a genuine natural kind. Any putative term that does not is suspect.

Evolutionary theory predicts that monophyletic groups and only such groups emerge from various evolutionary processes termed *speciation*. They are composed of a common ancestral species and all of that species' descendants. Although *monophyly* is just a noun, the noun is associated with a prediction that we will find groups with the properties of monophyly if evolutionary descent is real. Groups given the adjective *monophyletic* should exist in the world because evolutionary theory predicts that common ancestry groups result from evolutionary processes termed *speciation*. Such groupings are sought because evolutionary theory predicts their existence. The assertion that a group is monophyletic is a hypothesis that a unique common ancestry relationship exists between the species of the group and does not exist with other species outside the group; but all such groups have similarly unique relationships. Thus, all truly monophyletic groups have the property of being composed of species, or higher taxa, who have exclusive, or unique, genealogical descent from a founder species. Each higher taxon we hypothesize to be monophyletic stands as a singular confirmation of macroevolutionary theory because macroevolutionary theory predicts that such groupings should exist.

We can say that Vertebrata or Angiospermae are hypothesized to be members of the natural kind "monophyletic group." The importance of monophyletic groups to the evolutionary process is considerable. If we fail to discover *any* monophyletic groups, then we will be forced to change our process theory in the face of a prediction (evolution results in monophyletic groups) that does not seem to be met in nature. To put it another way, we would reject the hypothesis that evolution results in a pattern of hierarchical descent. What evolutionary theory *does not* depend upon is the discovery of *particular* monophyletic groups. Macroevolutionary theory is not a theory of particular groups; it is a theory about groups in general. It is not affected in the least if we discover that a particular group thought to be monophyletic turns out to be fiction. It might be devastating for the investigator who proposed the group, but it does not cause the overthrow of a process theory. However, no current evolutionary theory postulates the origin of paraphyletic groups; they are one of Hull's "mere aggregations," or evolutionary theory as we now understand it is wrong. Paraphyletic groups, like polyphyletic groups, are created by systematists, not by nature. As such, they are arbitrary delineations regardless of the good intentions of the investigator.

What would a theory of evolution look like that does not predict the existence of monophyletic groups? Theories of spontaneous generation might result in a multitude of single lineages evolving up the scala naturae (Lamarckian evolution or evolution within the Aristotelian paradigm); or there might be pervasive horizontal gene transfer that overwhelms a signature of hierarchical descent. Finally, one could adopt the theory that evolution is a myth and that the world was created by a deity who organized diversity according to kinds and we are fooled into thinking that the kinds are groups with some historical significance ("God thinks cladistically;" Ridley, 1986:110). Empirical science has rejected the Lamarckian thesis, and science, in general, dismisses supernatural explanations from the purview of scientific inquiry (starting, so far as we know, about with Thales of Melitus).

The Form of Phylogenetic Hypotheses

Phylogenetics is a research program concerned with the relationships of organisms, species, and monophyletic groups of species. As such, it asserts that individual organisms are constituents of monophyletic groups and species that exist in nature. Some organisms, such as mules, form exceptions and might be thought to be constituents only of a monophyletic group and not to any one species. These assertions form part of the background knowledge or auxiliary assumptions that are taken for granted, relying on evolutionary theory to provide the justification for these natural kinds. Of course, the properties (and thus, definition) of the natural kind "species" is a contentious issue. Systematists, in general, and phylogeneticists, in particular, disagree among themselves as to what constitutes the natural kind "species" and even if there might be more than one kind. But most do not disagree that there must be at least some kind of species.

The ontology of taxa hypothesized to have the properties of monophyletic groups and species (of whatever sort) is important precisely because their ontological status affects the manner that hypotheses are tested. If natural taxa, in general, are entities (and thus particulars or individuals in the philosophical sense), then hypotheses

concerning their existence or their relationships, or their status, are singular in form and confirmation and disconfirmation are symmetrical. That is, each instance of disconfirmation may be countered by a single instance of confirmation and the hypothesis is accepted if confirmation is greater than disconfirmation. If, however, natural taxa are natural kinds, then disconfirmation counts more than confirmation. One reason Sober (1993) was suspicious of the idea that Popper's falsifiability was appropriate for phylogenetics is that single instances of disconfirmation do not and should not lead systematists to reject phylogenetic hypotheses (see also Sober, 2008).

Hull (1981) concluded that hypotheses in systematics are largely singular hypotheses. Systematic hypotheses usually assert that particular entities (for example, *Pinus ponderosa*) are parts of other particular entities (Pinaeacea), or that they are members of natural kinds (the assertion that *Pinus ponderosa* is a member of the kind "species"), or that they are byproducts of empirical mistakes (for instance, that a systematist made a mistake in naming *P. ponderosa*). As singular hypotheses, these three alternatives are hypotheses in which confirmation and disconfirmation (i.e., verification and falsification or confirmation and refutation) are coequals. The discovery of a character that validly disconfirms a particular hypothesis can be countered by the discovery of a character that validly confirms a particular hypothesis. (Of course, one can argue as to what constitutes a valid confirmation!) In the end, one counts up the number of confirmations and disconfirmations and picks the hypothesis that best meets the criterion that has been selected for accepting one hypothesis over another.

Hull's reasoning refutes much of the systematic literature devoted to the applicability of the philosophy of Karl Popper (1965) to phylogenetics (a literature that begins with one of our own attempts to show that Popper fit phylogenetics better than evolutionary taxonomy; Wiley, 1975). A scientific arena where hypotheses are singular and verification and refutation are symmetrical is not the Popperian Arena, regardless of what inspiration one might gain from reading Popper's works (which in Wiley's case was considerable).

Popper was interested in falsification because he wished to discover a clear demarcation between scientific statements and nonscientific statements and, at the same time, solve the problem of induction. This is important, of course, but Popper was never really successful in his quest for reasons discussed by Sober (1993:46–54). Sober (1993) suggested a more modest goal: scientific hypotheses should be vulnerable to observation. For our hypotheses to be supported by observational evidence, they must be vulnerable to disconfirmation. In systematics, disconfirmation comes in the form of patterns of characters that imply a different relationship from the current hypothesis. Sober (1993) derived the principle of vulnerability from the Likelihood Principle, and we advocate that this principle can usefully be applied within a phylogenetic framework: If an observation (O) favors one hypothesis (H1) over another (H2), then "not-O" would favor H2 over H1 because if the probability of O given H1 is greater than the probability of O given H2, then the probability of not-O given H1 must be less than the probability of not-O given H2. Or:

$$P(O\,|\,H1) > P(O\,|\,H2), \text{ then } P(\text{not-O}\,|\,H1) < P(\text{not-O}\,|\,H2)$$

In a traditional parsimony framework, the emphasis would not be on probability. In "simple" parsimony all observations would be treated as equally likely (and of equal weight) such that number of observations becomes the arbiter of hypotheses. In particular, the hypothesis with the greatest number of observations in its favor would be endorsed. But in weighted parsimony, likelihood of transformation differs among and within different characters. And in likelihood, the emphasis would be the likelihood of observing the data given a particular tree topology and set of branch lengths.

Interestingly, Sober (1993) chose to discuss the issue of vulnerability in his treatment of creationism. In doing so, he exposed another important component of scientific theories. When discussing the idea of falsifiability as it relates to Popper's distinction between science and nonscience, Sober pointed out that for a Popperian theory to be tested in a strictly deductive manner, we must assume that any and all auxiliary assumptions are true. Because we can never verify that the auxiliary assumptions are true, then it is not strictly possible to falsify a theory in a deductive framework. This suggests that subscribing to rigid Popperian falsificationism is not a tenable strategy. A way out of this dilemma is simply to reject strict deductivism and embrace the concept of vulnerability derived from Sober's likelihood reasoning. In terms of the creationism debate, Sober suggested that it was the inability to discriminate between auxiliary assumptions (Biblical literalism, or intelligent design, or Zuni or Hindi theological assumptions, etc.) that rendered creationism untestable, not vulnerable, and thus not science. In passing, Sober (2008) discusses many of these issues as well as issues concerning such topics as parsimony, likelihood, and Bayesian analyses. We recommend this particular book as an updated account of Sober's philosophical approach to evidence in science.

In summary, the philosophy of systematics is a philosophy of testing alternative singular hypotheses within a framework of hypothesis vulnerability. Hypotheses must be vulnerable to disconfirmation. If they are not, then they are not testable. Strict Popperians obviously will not agree with every aspect of this philosophy. Still, the strength of the phylogenetic research program is two-fold. First, hypotheses must be transparent in that conclusions must be drawn based on empirical evidence thought by the investigator to be valid. Second, hypotheses must be vulnerable in that the evidence presented as confirmation for any particular hypothesis can be challenged by new evidence or the reinterpretation of old evidence. Ideas cannot stand on authority or experience; they must stand on evidence.

CHAPTER SUMMARY

- Phylogenetic systematists reconstruct the evolutionary relationships among organisms, species, and taxa using homologies that are hypothesized to indicate unique genealogical relationships.
- Phylogenetic systematists classify species and higher taxa in such a manner that the resulting classification is logically consistent with the recovered phylogeny.
- In phylogenetic systematics, the term *relationship* refers to genealogical relationship rather than overall similarity.

- Only monophyletic groups are considered natural in the phylogenetic system.
- Hypothesis testing in the phylogenetic system consists of the analysis of the characters of organisms, and the "best" hypothesis of genealogical relationship is deduced from the weight of confirming over disconfirming character evidence.

2

SPECIES AND SPECIATION

A basic task of the systematist is to estimate and describe the diversity of species in the group or faunal/floral study. From the realist perspective, one assumes that there is some true number of species in any particular group or in a particular region at any time. We strive to approach that number, but our efforts are always reduced to estimates. How accurate an estimate might be will depend on many factors. One important factor is how closely the species concept applied by the investigator approaches species as they exist in nature. Other factors also play important roles. These include the biology of the organisms, the experience of the investigator, the amount and critical nature of previous investigations, the quantity and quality of specimens available, and the analytic techniques employed.

The word *species* is applied in three distinct ways that are of interest to phylogeneticists. First, we assert that organisms living in nature, studied through examining specimens, are organized into taxa to which a binominal, or species name, is applied. These are assertions about particular species that are hypothesized to be a naturally occurring nexus of individual organisms comprising one of the basic units of evolutionary organization. Second, we seek concepts that are associated with the properties that we think all species possess. That a particular species (*Pinus ponderosa*) is an example of a particular species concept (e.g., it is an example of an evolutionary species) is an assertion that this particular species has the properties associated with the concept. Third, we have rules that govern the forming of species names and specify where in a hierarchy of classification species taxa belong. This third application is covered in Chapter 11 when we discuss taxonomy.

Four major discussions of species follow in this chapter. First, we will discuss the ontological status of species-as-taxa. This is a largely philosophical discourse. However, we think it is an important one. To not know whether binominals refer to

kinds, sets, or individuals is to admit that we know nothing of species at all (Coleman and Wiley, 2001). Second, we will discuss species concepts, the concepts that attempt to describe the properties that particular species have that make them species. Third, we will discuss modes of speciation, with particular attention paid to how these modes may affect patterns of descent. Finally, we will present some recommendations for particular decisions concerning whether a particular group of specimens should or should not be considered distinct species.

WHAT IS IT TO BE A SPECIES?

There are three basic ways of thinking about the nature of particular species: as kinds, as sets, and as individuals. As it turns out, the question has practical importance. If species are kinds, then empirical tests applied to the question of whether a group of specimens can be taken to represent a species take a different form than if species are sets and yet another form if species are individuals. As pointed out by Coleman and Wiley (2001), these questions are different from the question of whether species are real or nominal. We take the realist perspective: species exist in nature apart from our ability to perceive them and at any one time there are only so many species, not more or less depending on fancy. However, one could be a nominalist and still be faced with whether nominal species are kinds, set, or individuals.

Species as Kinds

Building on his original thesis (Ghiselin, 1966) that species are individuals, Ghiselin (1974) points out that philosophers, ever since Aristotle, have treated species as kinds/classes, while biologists, ever since Buffon, have treated species as particulars/individuals. To Ghiselin, this represents a disjunction between the philosophy of science and the practice of science. In support of his argument, Ghiselin cites (then) recent philosophers who treat species as classes: Greg (1950), Buck and Hull (1966), Lehman (1967), and Ruse (1969, 1971). Some philosophers continue to reject the thesis, searching for kind concepts that might fit. For example, Boyd (1991) suggested that species were "homeostatic cluster kinds," a concept supported by Griffiths (1999) and Ereshefsky (2001). Others have searched for concepts of kinds that might fit species (Ruse, 1987; Mahner and Bunge, 1997; Boyd, 1999; Wilson, 1999a).

Without changing the traditional Western philosophical concept of kind, we can think of kinds as concepts associated with defining properties such that individuals (particulars) are either members of one kind or another. These defining properties are both necessary and sufficient. If a particular kind "falls out" of a scientific theory (Quine, 1969) because the theory posits that certain entities should have certain properties if the theory is true, then it is termed a *natural kind* and the properties are said to be predicted by the theory. To put it another way, scientific theories address particulars/entities through natural kinds, or kinds that are thought at the time the theory is formulated to be natural kinds such that particular examples can be found in the real world. For example, theories about population genetics contain many kinds; "Mendelian gene" and "Mendelian population" are examples. If one

can find examples of, say, a Mendelian population, then this gives a clue that population genetics theory is on the right track. However, if examples of the kind "Mendelian population" are repeatedly searched for and not found, then one would doubt that the theory of population genetics had anything interesting to say about the real world. Likewise, if evolutionary theory, broadly conceived, posits the existence of species in general, then this suggests that there is at least one natural kind, "species," that has properties that are manifested by particular species as part of their "speciesness." Never mind that we do not fully understand the true nature and properties of all species (or all atoms or all planets), but it would be difficult to see how we could find a single particular species without some notion of what it is to be a species that is gained from theories about the world. And we should always keep in mind the fact that if our theories about the world are wrong, then the kinds of species we think are present in the world may be an illusion.

Natural kinds, as kinds associated with general theories about processes that occur in the real world, are eternal and immutable, unbounded by either time or space, at least so long as the theory behind the kinds is considered valid. If the theory is proven false, then the kinds associated with the theory may be false as well. For example, the natural kind "helium" falls out of, or is integral to, theories of atomic physics as the natural kind of atom that has two protons. Theories of atomic physics explain why particular atoms have two protons, how having two protons confers properties such as chemical reactivity (or lack thereof), and under what circumstances one expects individual helium atoms to originate in nature.

In contrast to natural kinds, nominal kinds are simply nouns with definitions. The definitions contain certain properties, and when a thing fits the definition, the name is applied. For example, the noun *bicycle* refers to a certain kind of two-wheeled vehicle. Motorcycle is another kind of two-wheeled vehicle and automobile is a kind of four-wheeled vehicle. Motorcycle and bicycle belong to the kind "two-wheeled vehicle" and automobiles and motorcycles belong to the kind "vehicle with a motor." No one pretends that such kinds are an integral part of some scientific theory, although all might agree that they are descriptive nouns with meaning. So the difference between natural kinds and nominal kinds is not the difference between sense and nonsense, but how such kinds relate to, or are irrelevant for, scientific theories of the world.

Wiley (1989) stressed that process theories are tested by seeing if the natural kinds they predict are fulfilled by particular entities behaving in the way predicted by the properties of the kinds. For example, atomic theory predicted that after the decay of a particular uranium atom, the result is an atom of lead and two atoms of helium. Failure to observe such a result under appropriate conditions would cause the entire theory to be suspect because predictions of the theory would not be met in nature by observing the entities (lead and helium atoms). Indeed, entire scientific theories can be overturned or extensively modified by finding one critical instance of the failure of the theory to meet the circumstances observed in nature, so long as scientists can agree that the test is valid. This, in turn, may lead to disposing of kinds once thought to be natural kinds and their replacement with kinds that are canonical to the new theory. So, natural kinds are extremely important to science as well as philosophy.

So, the question is: are particular species natural kinds? Ghiselin (2002 and earlier works from 1966) and others (e.g., Hull, 1978) argue that particular species (e.g.,

Homo sapiens and *Pinus ponderosa*) are not natural kinds at all. Species are mutable and have particular places in the fabric of history. If they have necessary and sufficient definitions, these definitions are not associated with process theories but with their time and place of origin, and in the post-Darwinian world, their ancestry. Particular species may be diagnosed, but not defined, by their character properties (Ghiselin, 1984; Wiley, 1989). Coleman and Wiley (2001) examined the issue from an analytical point of view and rejected the thesis of species-as-kinds by pointing out, among other things, that the relationship between nested kinds with transient relationships is not the relationship between particular species and the kinds with which they are associated. The line of reasoning used by them warrants some discussion.

Consider the kind "noble gas." This kind is predicted from atomic physics as the natural kind whose member atoms have the necessary and sufficient property of having their orbitals filled with electrons. Thus these atoms are relatively unreactive (cores of old stars being an exception). Members would include *individual* atoms of helium and neon. Nested within the natural kind "noble gas" are other natural kinds, the kinds "helium," "argon," "neon," and so on. The relationships of nested kinds are what we term transitive: an *individual* helium atom, being a member of the natural kind "helium," is also a member of the natural kind "noble gas."

Now, consider species. Evolutionary theory posits that there is at least one natural kind "species," associated with natural processes we term speciation. Taxonomists name particular species; *Homo sapiens* is an example. Coleman and Wiley ask: does the same logical relationship exist between the kind "evolutionary species" (or "biological species" or any other kind "species") and *Homo sapiens* that exists between noble gas and helium? To preclude arguments over words, they adopted the neutral term *constituent* rather than member (implying membership in a kind) or part (implying part of a whole). The answer is that the same logical relationship does not obtain, and considering the reasoning behind this is instructive. While an individual atom with two, and only two, protons is a constituent of both the kind "helium" and the kind "noble gas," and an individual human being is an constituent of the taxon *Homo sapiens*, an individual human being is certainly not a constituent of the kind "evolutionary species" or any other kind of species, nor is it a monophyletic group or a natural taxon. Thus, *Homo sapiens* is not acting as a natural kind nested within the kind "evolutionary species," any other species kind, or taxon. It is, however, a part of Hominidae, just as Ed Wiley and Bruce Lieberman are parts of *Homo sapiens* and parts of Hominidae.

Species as Sets

We must begin with the recognition that sets are not kinds. Sets are defined by extension, that is, by their membership. In contrast, kinds are defined by intension, that is, by the properties of their members. Second, sets are treated by philosophers as individuals, not as classes or kinds. Systematists are drawn to set theory because it is a well-worked system of inclusive relationships that seems to fit what systematists do in their research, especially in classification. So, it would seem natural to think of individual organisms as members of a set to which a binominal was applied. However, there are serious problems with this concept. Sets with different memberships are different sets. Owing to birth and death, the set of all humans today is different from the set of all humans yesterday. If species are sets, then there would

be many species of humans; for instance, the species that exists today, the one that existed yesterday, the one that existed two minutes ago, etc.

There is also a problem with sets and evolution. Hull (1981) argued against species-as-sets, pointing out that species, but not sets, are capable of undergoing evolutionary change. However, Kitcher (1984) argued that species were sets composed of all their present, past, and future members and thus a union of a series of subsets that exist during short time slices. He also argued that under this concept, species could be understood to change in the sense that different frequencies of traits could be observed between members of different species' subsets. To him, species change their membership over time because membership of one subset could be different from membership of a later subset. Further, one species could be understood to give rise to other species, because members of one subset of an ancestral species could be parents of members of one subset of a descendant species. Coleman and Wiley (2001) point out that Kitcher's concept of change transforms customary notions of species arising, evolving, and going extinct into one that can be explained within set theory, suggesting that the very transformation precludes species from being Kitcherean sets. He fails to show that a set-theoretical notion of species changing is an improvement upon the common, but supposedly confused notion of species changing. In addition, he does not demonstrate that the set-theoretical notion of species changing is the same as the traditional notion of species changing. Finally, the thesis of species as sets needs some notion of historical sets, a concept not yet articulated. Thus, although species can perhaps be shoehorned into Kitcherian sets, more seems to be lost than gained.

Species as Individuals

As stated above, Ghiselin (1966) and Hull (1976), as well as Hennig (1966), were among the first to seriously entertain the idea that species were individuals and not kinds. Probably because systematists had treated species as individuals for over a century (Ghiselin, 1974), biologists were quick to adopt this proposition regardless of systematic approach (cf. Wiley, 1981a; Mayr, 1982). Ghiselin has written extensively on the thesis (e.g., Ghiselin, 1997, 2002, and references cited therein) and a considerable body of literature has grown up around the notion (e.g., Coleman and Wiley, 2001). Species have many of the characteristics we look for in particular entities; they have a birth through speciation and a death through extinction. They can change their parts without changing their names, just as an individual organism can turn over its cells. They can also be named and diagnosed, but never defined, and their boundaries are fuzzy. However, there is more to the thesis. The concept is important when one considers the relationship between process theory, kinds, and entities. If species are individuals, then hypotheses about species are singular statements, and historical singular statements at that. If so, then hypotheses about species and about species' relationships are tested by weight of evidence, as outlined in Chapter 1.

SPECIES CONCEPTS

Although species-as-taxa are individuals, species concepts are kind concepts (Ghiselin, 2002 and earlier works; Wiley, 1989). As kind concepts, they are defined

by intention, each concept having properties that should provide necessary and sufficient conditions for "speciesness." There are many species concepts, most of them reviewed by Mayden and Wood (1995) and Mayden (1997). Entire books are devoted to the subject (e.g., Mayr, 1957, 1963, 1970; Vrba, 1985; Claridge et al., 1997; Wheeler and Meier, 2000) as well as papers too numerous to cite. As kinds, we might expect that some species concepts are candidates for being natural kinds. Such concepts would "fall out" of evolutionary process theories (Quine, 1969). We might also expect some concepts to be nominal kinds, kinds that do not have direct connections with evolutionary theory but which are thought to be useful in some manner by those who invented them. Both Mayden (1997) and Wiley (2002) have suggested this distinction, and we will explore this possibility as we review some of the more prominent concepts. The goal is to see if a species concept can be found that is a natural kind concept and one that is suited for general application in phylogenetics.

Ghiselin (1974:539) began the process of sorting thorough species concepts under the paradigm of species-as-individuals. He concluded that concepts such as the Morphological Species Concept or the Phenetic Species Concept treat individual species (e.g., *Pinus ponderosa*) as "classes defined in terms of the traits of organisms rather than as individuals having the properties necessary and sufficient for membership in the species category."

Wiley and Mayden (2000a–c) suggest that systematists form species concepts in a manner that reflects their ideas of how these concepts function in systematic and evolutionary theory. Some systematists form concepts that allow them to discover what they think are species. Others form concepts based on how they think species function in the evolutionary process. We suggest that this is nothing but the familiar debate on operationalism that surfaces from time to time in science (Wiley, 2002). Some systematists think that it is important to form "operational" concepts. Now, much depends on what one means by "operational." If what is meant is a concept that leads to testable consequences, then, all science should be operational. However, if what is meant is the philosophy of operationalism (Bridgman, 1927), the approach is problematic. In the philosophy of operationalism, kinds and concepts are not drawn from higher level "covering laws," but rather, they are defined by the observations by which they are measured or applied. An operational concept *sensu* Bridgman includes a discovery criterion that permits the systematist to apply some definite criteria to the question of whether a particular sample of specimens is drawn from a species or is simply a variation within a species. This approach seems seductive; it has the aroma of hard science by carefully providing specific criteria to be applied. But as pointed out by Hull (1968), operationalism carries a heavy load. Within the operational epistemology is embedded a particular ontology. The entities or properties recognized are defined by the operation employed. So if there are two operational criteria for species recognition, then there are two kinds of species. Just as there are two kinds of operational criteria for determining weight, there are two kinds of weight (English and metric come to mind, but there are more). Further, one can hardly conceive how such kinds are a consequence of process theories because they are invented by systematists.

It is exactly this distinction that led Frost and Kluge (1994), Mayden and Wood (1995), Mayden (1997), and de Querioz (1998) to suggest that a distinction can be drawn between what might be termed "general" concepts and "operational" con-

cepts. General concepts provide an ontology from which "operational" concepts (=testable concepts) may be applied. We will return to this thesis in a later section.

Process-Based Concepts

Process-based concepts attempt to characterize species in a manner predicted by "covering laws" thought to explain processes occurring in nature. Evolutionary theory (specifically that part of evolutionary theory concerned with the origin of species) provides such a "covering law," which posits that there are two kinds of entities one might expect to find in nature that are of primary interest to systematists, species and monophyletic groups. Given that evolution appears to result in a hierarchy of entities, the monophyletic group seems to be a natural kind that falls out of the general theory of descent with modification, coupled with the general theory of speciation. Indeed, the pattern of biological diversity, the observation that it does not form a *scala naturae*, and the observation that it is largely hierarchical, seems to argue that descent with modification is not sufficient to explain biological diversity unless it is coupled with some theory of lineage splitting. Even the counter-example of the divergent hierarchy, reticulate speciation, proves instructive. Reticulate speciation is not possible unless there is something to "reticulate." In other words, reticulate speciation depends on the prior existence of a divergent hierarchy. Otherwise, there would be no reticulate events apart from sexual reproduction.

Because all particular examples of the kind "monophyletic group" have a beginner, a common ancestral species, this suggests that species are also a necessary kind. In other words, if "monophyletic group" is a natural kind, then that from which particular monophyletic groups are derived ("species") might also have the status of being a natural kind (Wiley and Mayden, 2000a). The results form the basis for expecting, as Hennig (1966) suggested, that organisms have two sorts of relationships. Tokogenetic relationships are formed on the basis of reproduction and obtain among individual organisms. Phylogenetic relationships are formed by severing reproductive ties to the extent that two tokogenetic systems are formed out of what was once a single tokogenetic system (or, in the case of speciation via hybridization, two tokogenetic systems form a third through tokogenesis between the two systems). For sexual organisms, tokogenesis is nonhierarchical while phylogenesis is hierarchical (speciation via hybridization is the exception). The trick is to understand how we might discover the hierarchy. Organisms access information from previous generations during their growth and development. In other words, evolution seems to be an information-conserving process in an entropic informational environment (Brooks and Wiley, 1986), and a vast literature on inheritance and ontogeny suggest that this is the case. The history of organismal evolution is rather unusual among natural processes in that we can recover at least parts of this history through studying the characters of organisms. In other words, phylogenetic descent seems an unusual natural process as it conserves some information over large time spans, permitting us to reconstruct evolutionary relationships.

What general concept of species might we develop given these observations? We suggest that the most general concept is a concept of species-as-lineages (Simpson, 1961; Hennig, 1966; Wiley, 1978, 1981a). Species are the largest tokogenetic systems of biological organization (Frost and Kluge, 1994). In Ghiselin's (1974:538) words,

they are "the most extensive units in the natural economy such that reproductive competition occurs among their parts." Above this level of organization are monophyletic groups and ecological communities, and each belongs to a different "scalar hierarchy" (Eldredge and Salthe, 1984; Eldredge, 1985). Scalar hierarchies are hierarchies of levels of organization, with each higher level having emergent properties not fully explained by lower levels. The "economic hierarchy" would be one familiar to ecologists (organism, population, community, ecosystem, etc.) while the "historical hierarchy" would be one familiar to phylogeneticists (organism, population, species, monophyletic group). Below the level of organization represented by species, and within the scalar hierarchy of history, are several entities including populations and individual organisms. This general concept of species is met in the Evolutionary Species Concept (ESC) discussed below. After introducing this concept, we will briefly discuss other process-based concepts, as well as several discovery-based concepts.

The Evolutionary Species Concept

The ESC, originally proposed by Simpson (e.g., 1961) is built around the notion of a singular descent community, a concept Hennig (1966) claimed could be traced at least as far back as Buffon (1749). Although Hennig (1966) favored the Biological Species Concept (BSC) of Mayr (1963), he treated species in a fashion similar to Zimmermann (1937, 1943), as individualized lineages. It is ironic that the concept most amenable to phylogenetic systematics would have been championed by an opponent of the discipline, George Gaylord Simpson (1961). The common goal for both Hennig and Simpson was a species concept with a temporal dimension. In particular, Hennig (1966:29) asserted:

> We have defined the phylogenetic relationships we are trying to present as those segments of the stream of genealogical relationships that lie between two processes of speciation. Thus, by definition phylogenetic relationships exist only between species; they arise through the process of species cleavage. The key position of the species category in the phylogenetic system corresponds to the following: the species are, in the sense of class theory, the elements of the phylogenetic system. The higher categories of this system are groupings of species according to their phylogenetic relationships.

Wiley (1978) saw the similarity of Simpson's (1961) evolutionary species concept and Hennig's (1966) characterization of species as basic elements in phylogenetic systematics. Others, such as Ax (1987), Kluge (1990), Funk and Brooks (1990), Frost and Hillis (1990), Brooks and McLennan (1991, 2002), Frost et al. (1992), Frost and Kluge (1994), Mayden and Wood (1995), and Mayden (1997, 2002) have integrated the concept into their approaches to the phylogenetic system. The ESC has been used in studies of speciation where biologically comparable units are necessary for analysis (Wiley, 1981a; Mayden, 1985; Wiley and Mayden, 1985; Lynch, 1989; Brooks and McLennan, 1991, 2002; Frey, 1993). The most recent definition is provided below.

> An evolutionary species is an entity comprised of organisms that maintains its identity from other such entities through time and over space and which has its own independent evolutionary fate and historical tendencies (Wiley and Mayden, 2000a).

Wiley and Mayden (2000a) provided an amplification of the concept in an effort to further clarify what it is to be an evolutionary species.

1. Particular evolutionary species (*Homo sapiens*, *Fundulus nottii*) are individuals with origins, existences, and ends.

2. Evolutionary species are entities within which tokogenesis prevails (Hennig, 1966).

3. A phylogenetic tree is a cartographic device (O'Hara, 1993), or graph, that portrays the hypotheses that particular species are the result of historical processes (Frost and Hillis, 1990; Kluge, 1990) and that we have discovered them during the course of research. They are, as Hennig (1966) asserted, entities that result from various processes of speciation.

4. As such, the ESC is synonymous with the Lineage Concept of Hennig (1966, not named by him but see Figures 15 and 16), the Cohesion Concept of Templeton (1989; according to Endler, 1989), the Cladistic Concept of Ridley (1986), the Concept of Population Lineages (O'Hara, 1993; de Queiroz and Gauthier, 1994), the General Lineage Concept of de Queiroz (1998), and the Hennigian Species Concept of Meier and Willmann (2000).

5. As such, the ESC is not synonymous with the BSC of Dobzhansky (1935, 1937) and Mayr (1942, 1963), the Recognition Concept of Paterson (1984), or various forms of the Phylogenetic Species Concept of Mishler and Brandon (1987), Mishler and Theriot (2000), de Queiroz and Donoghue (1988, 1990), Nixon and Wheeler (1990), or Wheeler and Platnick (2000).

6. Sexually reproducing species may show tokogenetic cohesion patterns (reproductive ties among individuals) that do not correlate with any hierarchical relationships that might exist among their parts. This is due to gene flow that might erase, for example, a time hierarchy of colonization. Contrary to the concerns of Donoghue (1985), this is to be expected based on first principles (Hennig, 1966; Kluge, 1990).

7. Completely asexual species (where they exist) have relationships among their parts that are identical to their descent relationships and are, therefore, dissimilar to sexual species and similar to clades. However, asexual species are similar to sexual species in that tokogenetic relationships predominate. In Ghiselin's (1974) terms, reproductive competition occurs among their parts. To the extent to which such species are strictly asexual, the hierarchy of such clone vectors is driven by character change and not changes in cohesion patterns. Wiley and Mayden (2000a) make a case for the inclusion of asexual species within the ESC but acknowledge that the ontology of asexual species is different from sexual species (Hennig, 1966:82–83), just as the ontology of sexual and asexual species are not identical to that of higher taxa (Wiley, 1980).

8. The ESC has been criticized for including phrases such as "maintaining identity" and "has its own independent evolutionary fate and historical tendencies" (e.g., Mayr, 2000). Similar criticisms regarding phrases such as "unitary role" have been criticized as well by Mayr (1982) who considered them unmeasurable. However, all such phrases have ontological not epistemological meaning (Wiley and Mayden, 2000a). They refer to the individuality of species. They also have testable consequences, as discussed below.

Justifications for the ESC   Wiley and Mayden (2000a–c) suggest reasons for advocating the ESC in phylogenetic research. The first is a justification for the reality of species-as-lineages. The second is a justification for the assertion that species can be viewed as independently evolving lineages. The remaining justifications are specific for application of the ESC to phylogenetic systematics.

If monophyletic groups are real and the kind "monophyletic group" is a natural kind, then lineage-species must necessarily exist. Monophyletic groups are diagnosed as groups descended from a common ancestral species that was, in retrospect, the first member of the group. If the monophyletic group has reality in nature, then the ancestral lineage from which all subsequent species in the group arose must also have objective reality in nature. Discounting spontaneous generation and special creation, the only alternatives are the idea that species do not give rise to other species (Mishler and Theriot, 2000: only populations are ancestors) or that species are nominal (which amounts to the same thing because they would lack reality).

The question is: what evidence do we have that monophyletic groups exist? Empirically, the "proof" lies in the presence of synapomorphies. Synapomorphies are the lingering effects of common ancestry, just as sunlight is a "lingering effect" of Sol's atomic chemistry. If this is true, if synapomorphies are valid evidence that monophyletic groups exist, then we can ask another question. What is the origin of a synapomorphy? At the highest level of descent, all synapomorphies begin as autapomorphies of single lineages, evolutionary innovations that became fixed subsequent to lineage splitting. How such autapomorphies became fixed is an issue that can only rarely be addressed because we cannot examine the population dynamics of lineages that existed in the past, only their descendants. We may be able to form hypotheses of why such characters persist and we may understand how they originate in the organism through studying development, but chances are slim that we will ever be able to study why they came to characterize one lineage while their homolog remained unchanged in another lineage. In fact, we may conclude that such characters were fixed in one particular ancestor only to later find evidence that it was carried as a polymorphism through several speciation events before fixation. This uncertainty seems to be the reality of historical research. Regardless of these difficulties, if at least some evolutionary innovations are real apomorphies of some monophyletic groups, then they originated as autapomorphies in single lineages, those entities termed evolutionary species. Indeed, as asserted by Hennig (1966), such an ancestral lineage is equivalent to the monophyletic group at the time of the group's origin. Thus a monophyletic group ranked as a family was, at the time of its origin, composed of a single evolutionary species (Hennig, 1966; Wiley 1977a, 1978).

At least some ancestral species must be independently evolving entities with their own evolutionary fate and historical tendencies because synapomorphies can be used to diagnose monophyletic groups. The very fact that we are successful in reconstructing the relationships of many groups is evidence that lineages have the capacity to evolve independently from other lineages. If the origin and spread of evolutionary novelties were not highly constrained by the independence of lineages from each other and if lineages freely exchanged heritable information, then we could not reconstruct phylogenetic relationships and the history of evolutionary descent would not appear as it appears. Again, the exception proves the rule. Horizontal (lateral) gene transfer can cause havoc in phylogenetic analyses if not recognized (Daubin et al., 2003). In cases where speciation occurs via hybridization,

we may see the opposite, a confusing pattern of synapomorphies (Funk, 1985; Smith 1992) and a struggle to reconstruct evolutionary history. Of course, even in the absence of reticulation, phylogenetic relationships are not always easy to reconstruct due to unusually slow or rapid evolutionary rates in the characters examined. For instance, when rates of speciation are high relative to rates of change, as in the case of the morphology of some African rift lake cichlids (Greenwood, 1984), or the case in some diatoms (Theriot, 1992) when certain modes of speciation prevail, we may lose the trace of hierarchy.

The ESC is the logical analog of the kind "monophyletic group." Neither is "operational," and this is a strength of the concept, not a weakness (Wiley and Mayden, 2000a; Wiley, 2002). As we shall see, a concept can embody testable consequences without being operational *sensu* Gilmour (1940).

Application of the ESC affords the possibility that binominals (names applied to particular lineages) are comparable (in their knowledge claims if not their validity) across all groups. All evolutionary species are theoretically comparable because they all represent the largest tokogenetic systems of the organisms included in the tree of life (Frost and Kluge, 1994). They are neither parts of larger nor smaller tokogenetic systems (although they are certainly parts of monophyletic groups). A deme of *Fundulus nottii* is probably not comparable with a deme of *Pinus ponderosa*, the family Fundulidae (killifishes) certainly is not comparable to the family Pinaecae (pine trees) except that they might be monophyletic. But *F. nottii* is comparable, in theory, to *P. ponderosa* to the extent that each represents a lineage whose origin lies in a speciation event. Thus, the general characteristics of species and phenomena associated with, or causal to, speciation can be studied with more than zero degrees of freedom. We may ask questions about species in general (the properties of the kind species as it functions in evolutionary theory); we are not restricted to asking questions only about particular species. Species as basic units of biogeography become comparable, and their relationship to Earth history can be investigated. Thus, the ESC provides a biologically meaningful concept for comparing species in phylogenetics (Hennig, 1966), coevolution (Brooks and McLennan, 2002), historical ecology (Brooks, 1985), biogeography (Brundin, 1966; Brooks, 1981), speciation (Wiley, 1981a; Wiley and Mayden, 1985; Endler, 1989; Frey, 1993), and paleontology (Eldredge, 1993; Krishtalka, 1993).

The ESC provides an ontological basis for a logically consistent relationship between species and phylogenetic trees that is comparable to the relationship provided by the concept of monophyletic group *sensu* Hennig (1966). As we shall see in Chapter 4, applying the ESC to such basic questions as the number of alternative phylogenetic trees for a particular group of species or higher taxa results in far fewer alternative hypotheses than some investigators assert. Further, some of these "alternative" hypotheses turn out to be impossible alternatives.

Variations on the ESC   There are a number of process-based concepts largely or entirely synonymous with the ESC. They include the following:

The Biological Species Concept of Ghiselin (1974, 1997). Ghiselin redefined the BSC as "the most extensive units in the natural economy such that reproductive competition occurs among their parts." Frost and Kluge (1994) considered that this definition rendered the BSC "coincidental" with the ESC "at least in their application to biparentals."

The Cohesion Species Concept (CoSC). The cohesion species concept is the most inclusive population of individuals having the potential for phenotypic cohesion through intrinsic cohesion mechanisms (Templeton, 1989:12). As stated by Templeton (1989:4), the CoSC is closely related to the ESC, differing primarily in perspective. The CoSC emphasizes mechanisms promoting cohesion while the ESC emphasizes the "manifestation of cohesion over evolutionary time." Endler (1989) finds the concept largely synonymous with the ESC.

The Cladistic Species Concept (ClSC) of Ridley (1986) is the ESC: "A species is then that set of organisms between two speciation events, or between one speciation event and one extinction event, or that are descended from a speciation event" (Ridley, 1986:3). Ridley attempted to capture what Hennig (1966) actually meant. He emphasized the individuality of species-as-taxa and the subordinate role other species concepts have to the ESC/ClSC. Ridley's (1986) interpretations of the ClSC relative to the ESC does differ in some respects, specifically, the persistence of ancestral species. However, we consider these differences minor. Note that de Queiroz (1998:59) classified Ridley's concept as a variant of the Phylogenetic Species Concept, discussed more fully below. We view Ridley's concept as an ESC process-based concept because it explicitly defines species based on the results of process and thus attempts to capture species as they are thought to exist in nature. The ClSC is closely related to the *Internodal Species Concept* of Kornet (1993).

The General Lineage Concept (GLC) of de Queiroz (1998) is the ESC. De Queiroz (1998) was attempting to show that there was a general concept that subsumes other concepts. His thesis, like that of Mayden (1997) was that all modern concepts could be grouped into concepts that were reflections on a more general concept. The difference between de Queiroz (1998) and Mayden (1997) is that Mayden made the decision that all other concepts were reflections on the ESC while de Queiroz (1998) made the decision that a new name for the ESC was needed. We do not agree.

Process-Based Concepts Emphasizing Reproductive Isolation

The most popular process-based concept emphasizing reproductive isolation is the *Biological Species Concept (BSC)* of Dobzhansky (1937) and championed by Mayr (1942, 1963, 1969, and numerous papers summarized in Mayr, 2000): A species is a group of interbreeding natural populations that is reproductively isolated from other such groups (Mayr and Ashlock, 1991:26).

Templeton (1989) has termed the BSC the "isolation concept" based on its reliance on isolating mechanisms. The BSC captures part of the ESC (Mayden, 1997; Mayden and Wood, 1995) for those evolutionary species that have achieved reproductive isolation and that are living in sympatry. Of course, sympatry is common. The question is: what is the pattern of newly formed species? The overwhelming answer is that most newly formed species live in allopatry, not sympatry (Wallace, 1855; Mayr, 1963; Lynch, 1989; Mayden and Wood, 1995; Wiley, 2002). In fact, Ghiselin (1989) has even coined "Mayr's Law" that states "under ordinary conditions, speciation without antecedent geographic isolation does not occur" (Ghiselin, 2002:154). Most criticisms of the BSC concern either the testability or the concept in practice. Wiley and Mayden (2000a) point out that if speciation is usually allopatric (Wiley and Mayden, 1985; Lynch, 1989; Grady and LeGrande, 1992; Chesser and Zink, 1994; Mayden and Wood, 1995; Mayden, 1997; Ghiselin, 2002; Wiley, 2002; Coyne and Orr,

2004), then speciation does not require that species achieve reproductive isolation. Presumably, many newly evolved sister species live in adjacent geographic regions and in the same ecological context as the ancestral species. They never meet each other. This situation might exist for millions of years and subsequent speciation might take place in the interim. If so, and if reproductive isolation has not been achieved (as in, for example, fishes of the genus *Xiphophorus* studied by Rosen, 1979), then are we to claim that new species can arise from entities that are not, themselves, species? It seems a high price to pay when a lineage concept such as the ESC is available. It was for this reason that Paterson (1978, 1984) developed the *Recognition Species Concept (RSC)*: A species is that most inclusive population of individual, biparental organisms that share a common fertilization system. Paterson realized that organisms in populations are recognizing other members for the purposes of interbreeding rather than trying to avoid reproducing with other organisms that they could never encounter. The RSC emphasizes how isolating mechanisms keep individual species together (Templeton, 1989; de Queiroz, 1998). The concept is built around a specific mate-recognition system (SMRS) that consists of specific coadapted signals and releasing factors that are exchanged between partners and that are maintained by stabilizing selection. Such factors might operate from the level of behavior to the level of genetic recognition on the cellular level. Lieberman (1992) discussed how to potentially extend the RSC and SMRS into a phylogenetic context.

Mayden and Wood (1995) summarize other criticisms of the BSC, two of which are critical for our purposes. The BSC lacks a lineage perspective; it is nondimensional. It is difficult to translate the concept into a concept suitable for phylogenetics or paleontology, research programs that inherently contain a time dimension. Further, the BSC (and its variants) does not embrace asexual species, which are left in limbo as pseudospecies (Dobzhansky, 1970).

The *Hennigian Species Concept (HSP)* of Willmann (1986) and Meier and Willmann (2000) is difficult to characterize because it seems a mix of the BSC and the ESC. It is like an extreme form of the BSC in that complete genetic isolation relative to other species is required for species. It is like the ESC in that species should be delimited by speciation events, given that speciation involves a transition from tokogenetic to phylogenetic relationships, and the idea that phyletic speciation is an artifact and not a process (Wiley and Mayden, 2000a). Other points raised by Meier and Willmann (2000) also focused on the persistence of ancestral species. Wiley and Mayden (2000b) diverged with them on this issue, considering the persistence of ancestral species to be an open question, more in the realm of ontology than epistemology. Further, they argued that there were empirical hurtles that must be overcome before such issues could be addressed in the real world (Wiley and Mayden, 2000b, c). In short, persistence of ancestral species should, in theory, result in unresolved polytomies of descendants. The extent to which unresolved polytomies imply persistence of ancestral species is unknown and likely to remain unknown in most cases. Thus, the question is of little immediate empirical importance in actual practice, although it remains an interesting ontological issue. We also point out that while Frost and Kluge (1994) embrace the necessary extinction of ancestral species, they also embrace the ESC, implying that the issue of ancestral persistence can be considered quite apart from the issue of the validity of the ESC.

Wiley and Mayden (2000b) felt that the issue of total reproductive closure as a feature of the HSC was much more interesting. As ichthyologists, they knew of no recently evolved and closely related species of North American freshwater fishes that are 100 percent reproductively isolated (except through allopatry) and they knew of many cases where hybridization occurs among species that are not each other's closest relatives. They suggested that the demand that species be totally closed reproductive communities was extreme and not met in nature for many groups. Further, it would require sinking many species, thereby potentially masking hierarchies that have already been corroborated with synapomorphies. They suggested that occasional tokogenetic events (rare hybridization, for example) were not synonymous with defining the extent of tokogenetic systems. In this, they agreed with Mayr (1963, 1969, and other works) who has long acknowledged that some species recognized under the BSC may hybridize with other species recognized as valid under the same concept.

Phylogenetic Species Concepts

There are several versions of concepts known as the *phylogenetic species concept*. It is understandable why there are so many of these concepts. As phylogeneticists, we naturally want to distinguish ourselves by adopting phylogenetic species concepts. Some of these concepts are purely operational, but most are a mix of process-based and operational criteria. We list some of these concepts below.

Phylogenetic Species Concept I, Species as Monophyletic Taxa. A species is "a geographically constrained group of individuals with some unique apomorphous characters" (Rosen, 1978:176). To Donn Rosen, this was a relatively simple concept that makes intuitive sense and is easy to operationalize. There are many species that can be diagnosed by discovering autapomorphies (see also Nelson, 1989). A variant of this concept was articulated by Mishler and Donoghue (1982), Donoghue (1985), Mishler (1985), and de Queiroz and Donoghue (1988), who argued that one way of defining species is as monophyletic lineages that are diagnosed ("defined") by apomorphy. Transferring the concept of monophyly from the traditional level of characterizing groups of species (Hennig, 1966) to the population level carries its own problems. Under this concept, all ancestral species are rendered paraphyletic because they would be characterized by plesiomorphies and would not include all of their descendants. Wiley and Mayden (2000a) posed the question: if the phylogenetic system allows for "paraphyletic species" (all ancestral species), then how can the phylogenetic system reject paraphyletic groups of species? One interesting response to this conundrum is to deny species ancestral status and to refer ancestry to the level of populations within species (Mishler and Theriot, 2000). Another possibility is to allow species to be defined in different ways (for example, as reproductive communities or evolutionary lineages), that is to represent different things (BSC or ESC), depending on the interests of the investigator and the scientific questions being addressed.

Phylogenetic Species Concept II, Species as Diagnosable Clusters. Eldredge and Cracraft (1980) and Nelson and Platnick (1981) suggested that it was not apomorphies but diagnosability that was the major criterion. Eldredge and Cracraft (1980:92) suggested that a species was "a diagnosable cluster of individuals within which there is a parental pattern of ancestry and descent, beyond which there is not, and which

exhibits a pattern of phylogenetic ancestry and descent among units of like kind." Nelson and Platnick (1981:12) presented a similar concept; species are "simply the smallest detected samples of self-perpetuating organisms that have unique sets of characters." Similar definitions were formulated by Cracraft (1983) and Nixon and Wheeler (1990). Wheeler and Platnick (2000:59) carried this line of reasoning to one logical conclusion:

> Speciation is marked by character transformation. In turn, character transformation occurs through the "extinction" of ancestral polymorphism (see Nixon and Wheeler, 1992). The moment of speciation is, in theory, precise and corresponds to the death of the last individual that maintained polymorphisms within a population.

The concept has serious problems when we consider its implications to process theories. Wiley and Mayden (2000b) point out that adopting this version of the PSC amounts to divorcing cladogenesis from species entirely. For example, unless two sister species fix their respective apomorphies at the same time they will have different times of origin. This would mean that sister groups, in general, do not necessarily have the same time of origin. This would destroy one of the comparative bases of phylogenetic systematics. Adoption of such a concept would also lead to dividing tokogenetically cohesive lineages into chronospecies if more than one apomorphy is fixed at more than one time. Our reasons for rejecting chronospecies are outlined below.

Some Additional Species Concepts

The oldest and most commonly applied species concept is the *Morphological Species Concept (MSC)* (e.g., Cronquist, 1978), a concept that has been in play at least since Aristotle. Morphological species are those that have morphological differences typical of what we think of as species; this is an entirely epistemological view of species. This can reduce to the "cynical" view (Kitcher, 1984) that species are what taxonomists say they are (the *Taxonomic Species Concept*; Blackwelder, 1967). More modern incarnations of this concept are found in the *Phenetic Species Concept (PSC)* of Sokal and Sneath (1963), Sokal and Crovello (1970), Sneath and Sokal (1973), and Sneath (1976). The PSC I and II are also related, although they provide an ontological flavor to this solely epistemologically constrained concept.

The *Ecological Species Concept* of Van Valen (1976) is closely related to the ESC in some respects in that it stresses lineage independence and defines a species as a lineage that occupies a different adaptive zone than other lineages in sympatry but is evolving independently from lineages outside its range.

The rise of modern genetics and the ability to investigate genetic variation in natural populations has led to several genetics-based concepts, most of which are closely allied to the BSC. These include the *Genetic Concordance Concept* of Avise and Ball (1990) that posits that species may be identified by the concordance of multiple independent genetic markers or the closely related idea that species form genetic clusters (Mallet, 1995) or are those groups of populations fixed for unique isolating mechanisms (Wu, 2001a, b). *Compilospecies* are species that appropriate the genes of other species via interspecific hybridization (Harlan and De Wet, 1963; Aguilar et al., 1999). No doubt there could be as many species concepts of this nature

as there are interesting genetic phenomena although they would thus be limited in their relevance and entities defined as such would offer little opportunity for scientific comparisons among equivalent entities.

SORTING THROUGH SPECIES CONCEPTS

The number of papers and books written on the "species question" is large. More recent compilations include papers in Otte and Endler (1989), Claridge et al. (1997), Howard and Berlocher (1998), Wilson (1999b), and Wheeler and Meier (2000). In a series of papers, Mayden and Woods (1995) and Mayden (1997) have catalogued some 22 species concepts and others have appeared since these publications. Mayden (1997, 1999) suggests that one of the confusing aspects of the "species debate" has been the notion that the 22+ concepts are somehow equivalent in their ontological and epistemological merits, and that this is not the case. Mayden outlined four criteria (first formulated in Mayden and Woods, 1995) that a primary concept must have to function as the most general concept of species (Mayden, 1999:97).

1. The concept must have a time dimension such that the concept can be applied directly to phylogenetic hypotheses.
2. The concept must view species as individuals and not kinds so that species are considered entities.
3. The concept must be unbiased "as to the type of organism, data, or sexual tendencies of the organisms" (i.e., to asexual as well as sexual organisms).
4. The concept must be nonoperational in the sense of Bridgman operationalism (defined more fully below), but it must also have testable consequences.

Criterion 1 ties the primary species concept to phylogenetics. Criterion 2 treats species as real entities existing in nature; each of these are meritorious goals in our view. Interestingly, Criterion 2 makes species analyzable. Sets are, of course, analyzable, but we have asserted that species are not sets at all. Kinds are analyzable, but only through examining entities for regularities relative to theory prediction. If species are not entities, how can we use them to test theory? Criterion 3 suggests that the primary concept should cover asexual as well as sexual species. It also suggests that no one kind of data is inherently better (or worse) than other kinds of data that might be applied to recognizing species (for an amplification see Mayden, 2002).

   Criterion 4 might sound unusual and is worth a short discussion. When Mayden (1997) states "nonoperational," he is asserting that it should not be operational in the sense of Bridgman. That is, it should not contain a prescribed set of operations. The problem with species concepts that detail a prescribed set of operations is that they allow us to find only those species that fit the operations, thus limiting ourselves to only a part of the species-level diversity that we might discover. This is quite different from requiring a primary species concept to have testable implications (Wiley and Mayden, 2000a). In other words, a species concept might be quite "nonoperational" in the sense of Bridgman, but quite testable by examining the consequences of its definition.

   Mayden and Wood (1995) found that some species concepts underestimate species diversity while others might overestimate species diversity. Mayden (1997,

1999) is interested in a primary concept that maximizes our ability to discover the actual number of species that exist, or have existed, in the world. That is, both Mayden and Wood (1995) and Mayden (1997, 1999) take a realist perspective; the metaphysical position that because species are real there must only be a finite number of them over time. (Any alternative is equally metaphysical.) Mayden (1999) groups all concepts into three categories. Most species concepts have some operational criterion embedded within the definition, rendering them strongly epistemological in nature. Some of these concepts are essentialistic (Adams, 1998). Mayden finds one concept, the ESC, to be the most robust and theoretical concept of the lot. But he also sees considerable consilience between this concept and most other concepts. In this, he agrees with de Queiroz (1998) who views most species as a manifestation of a single more general concept.

However, there is at least one concept that is not consilient with the ESC, the *Chronospecies* (Mayr, 1942) or *Successional Species* (Simpson, 1961) concept (Wiley, 1981a; Mayden, 1997). A successional species is a binominal applied to one segment of a single lineage that has, in the opinion of the describer, undergone sufficient anagenetic change to be recognizable. This concept is particularly popular among paleontologists (e.g., Gingerich, 1979; but see Krishtalka, 1993, who provides a strong criticism). Some versions of PSCII allow such species (Wheeler and Platnick, 2000). Simpson (1961) saw the need to recognize successional species because he thought that there was no nonarbitrary way to subdivide the reproductive continuum of the tree of life. Wiley (1978), citing Hennig (1966), countered that the tree was naturally subdivided by speciation events, and Wiley (1981a) rejected the concept of chronospecies because it ran counter to the thesis that species were individuals (Ghiselin, 1966) and lineages, not partial lineage segments. The ESC provides a clear alternative, but we can understand Simpson's quandary if we assume that he did not have a hierarchical perspective nor an appreciation of the difference between tokogeny and phylogeny.

Hennig (1966) saw the relationship between tokogeny and phylogeny as nontransitive. Consider this: in the absence of spontaneous generation, all life is tokogenetically connected. The nontransitive nature of the difference is apparent when we consider that if we had all of the tokogenetic relationships before us, we would be able to reconstruct the tree of life, but if all we had were the species-level relationships, we could not reconstruct the tokogeny relationships (Coleman and Wiley, 2001). Phylogeny plays by the rules of reproduction, but adds a layer of natural processes on top that we label as the various processes of speciation. Recognizing speciation events is not an arbitrary exercise, even if we miss a few and get some wrong. If we take Hennig's perspective and cleave nature at its joints, we will always separate some parent from some offspring; it is the nature of a world in which species evolve rather than a world in which species are created. If, however, we subdivide a unitary lineage, we are imposing on nature; there is no speciation event, only tokogeny.

SPECIATION: MODES AND PATTERNS

There are many avenues to choose if one wishes to study speciation. In a perspective essay, Endler (1989) asked the question: "What are we trying to explain?" Evolutionary biologists are likely to be interested in intrinsic mechanisms such as

gene changes that promote reproductive isolation or mate recognition. Much of the research in this approach has been summarized recently by Coyne and Orr (2004) and need not be repeated here. Systematists are likely to be interested in discovering how patterns of phylogeny and biogeography can illuminate our understanding of various modes of speciation. This dichotomy between those who study mechanism and those who study pattern was discussed by Mayr (1963), who noted that most of the "fathers" of the neo-Darwinisn synthesis had little appreciation for the role of speciation (Dobzhansky was the exception). Further, Mayr (1963) championed the ideas of Karl Jordan and David Starr Jordan that geographic isolation played the key role in the formation of species. In some sense, the dichotomy continues. Evolutionary biologists tend to emphasize mechanism over pattern, systematists pattern over mechanism. For example, the bulk of Coyne and Orr's (2004) treatment of speciation concerns mechanisms that promote reproductive isolation and they have a preference for the Biological Species Concept. By contrast, Brooks and McLennan (2002) emphasize pattern and the bulk of their treatment concerns the effects of vicariance on emerging phylogenetic patterns of descent. They emphasize the ESC. Neither entirely ignores the other, but both emphasize that side of the coin for which they have the tools to study.

The origin of new evolutionary lineages can be achieved in many ways through different modes of speciation. The question for systematists is: can we understand something about speciation modes by examining phylogenetic and distributional patterns? The question for evolutionary biologists is: can the fruits of systematic analysis assist us in studying the mechanisms of speciation? It is quite obvious to us, if Coyne and Orr (2004) are typical of the evolutionary community, that the evolutionary biologists are paying attention. The fruits of systematics are of increasing value to the evolutionary community.

Wiley (1981a), Wiley and Mayden (1985), Brooks and McLennan (1991, 2002), and Lieberman (2000a) suggested that patterns of phylogeny and biogeography can be used to study various modes of speciation given certain assumptions; the major assumption being that (1) the biogeographic distribution of the ancestral species can be inferred by the distribution of descendants and (2) the biogeographic patterns have not been significantly altered by postspeciation dispersal or extinction. Wiley's (1981a) reasoning was largely a result of considering the implications of the work of Donn Rosen (1978, 1979) on speciation patterns in Middle American fishes and Platnick and Nelson's (1978) introduction of vicariance biogeographic methods, both of which stressed the discovery of common patterns of biogeography between clades. About the same time, Brooks (1981) introduced the first of his quantative methods of biogeographic analysis (later to be refined, as outlined in Chapter 9). Most systematists have a deep interest in the role that biogeography played in the formation of species. But the ability to find common patterns using testable methods of matching phylogenies with distributions was possible only with the development of Hennig's phylogenetics and the ability to match biogeographic areas with taxon relationships, first formulated by Dan Brooks (1981). As examples, Cracraft (1982) studied speciation in various Australian bird clades and Lynch (1989) developed methods for testing various modes of speciation. Lynch (1989) found allopatric speciation to be the dominant mode in the clades he studied and found a surprisingly high percentage of sympatric speciation. Grady and Legrande (1992) applied Lynch's methods on North American freshwater catfishes of the genus *Noturus* and

also found allopatry to dominate. Chesser and Zink (1994) argued that allopatry also seemed to be the rule for birds, although they were critical of Lynch's criteria for identifying sympatric speciation. They suggested that the level of sympatry Lynch recovered might be explained by differing dispersal capabilities of clades analyzed rather than differing modes of speciation.

As pointed out by Losos and Glor (2003), the assumptions associated with methods based on discovery of common biogeographic patterns, especially when proceeding from terminal nodes (more recent events) to deeper nodes (more ancient events) should be carefully considered. Not all groups of species will be amenable to detailed analysis of speciation or the identification of speciation modes using biogeographic patterns and phylogenetic hypothesis (see for example, Wiley and Mayden, 1985). However, if we understand how some of the basic patterns might arise, we will be equipped to deal with clades for which the assumptions are reasonable.

Most authors recognize three basic modes of speciation. Allopatric speciation involves the physical subdivision of an ancestral gene pool into two or more descendant gene pools. Sympatric speciation is speciation without geographic subdivision. Parapatric speciation is speciation that involves partial geographic subdivision and differentiation in spite of limited gene flow across a hybrid zone. Below we first begin our discussion of each of these modes from the neontological perspective. In essence, the focus will be on how we can use phylogenetics and biogeography to study speciation in recent faunas given various modes of speciation. After discussing the neontological perspective, we will discuss speciation in deep time and have a look at the tempo of speciation as well as its mode. It is in deep time that we can get at phenomena such as punctuated equilibria, phenomena that are harder to study in a neontological context.

Allopartic Speciation

The role of geographic subdivision and subsequent differentiation has long been recognized as a major mode for the origin of new species. Darwin recognized it in his notebooks and in the *Origin*, even if he did not accord it as high a place as sympatric speciation. Wallace (1855) recognized that closest relatives are more likely to be found in adjacent regions than together in sympatry. Mayr (1942, 1963, et seq.) accorded allopatry the primary role in species formation. More recent discussions are found in Eldredge and Cracraft (1980), Wiley (1981a, 2002), Wiley and Mayden (1985), Funk and Brooks (1990), Lieberman (2000a), Brooks and McLeanan (2002), and Coyne and Orr (2004). One of the reasons that allopartic speciation seems so common is simply because it is biologically less restrictive than those that obtain with parapatric and sympatric speciation: any force causing divergence can yield speciation if two populations are allopatric (Coyne and Orr, 2004:85).

Different authors parse out different forms of allopatric speciation in different ways. Lieberman (2000a) and Brooks and McLeanan (2002) make distinctions based on whether subpopulations experience vicariance *in situ* (no dispersal) or one subpopulation establishes a new range by dispersing over an existing barrier. Thus, they characterize allopatric speciation as either "passive" or "active." Coyne and Orr (2004) make the cut based on population size. Establishment of new species by tiny subpopulations (founders of 1–100 members) characterize peripatric speciation

whether or not the population is split from the ancestor by a new barrier ("passive") or the founders disperse over an existing barrier ("active"). Lynch (1989) considered peripheral isolates to be species that occupy a range of 5 percent or less than their sister species. But 5 percent of the range of a fairly widespread ancestral species can still be a large range relative to the number of demes and individuals that are separated from the ancestral gene pool and may contain more than 1–100 individuals (even if that number represents effective population size). In fact, Brooks and McLennan would character this as "microvicariance." Coyne and Orr (2004) would call it simply vicariance because the usual population genetic phenomena associated with small population size would not apply.

When we consider large time scales, we also have to be careful about using the term dispersal. In a later section, we will deal with larger time scales and the question of punctuated equilibria (Eldredge and Gould, 1972). Dispersal could take the form of geodispersal (described in Chapter 9), which promotes subsequent vicariance.

Allopatric Mode I: Vicariance   Vicariant speciation obtains when an ancestral species is geographically subdivided by a physical or climatic barrier and when both subpopulations of the ancestral gene pool are too large for genetic drift to effectively overwhelm other forces (gene flow, selection) in shaping the short-term evolution of one or both daughter species. Both "dumbbell" (when divided populations are roughly equal in size, e.g., Bush, 1975) and "micro-vicariant" (when the daughter population is much smaller than the parent population) models apply (Fig. 2.1). In fact, both can operate at the same time in continental situations. For example, Wiley and Mayden (1985) document vicariance along the Gulf Coastal Plain of North
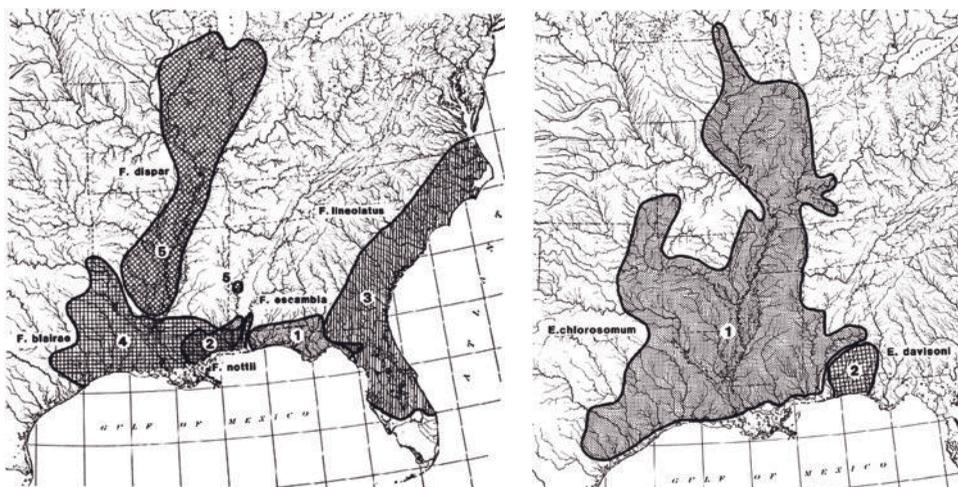
Figure 2.1. Allopatric speciation. (a) Dumbbell speciation between sister species of topminnows *Fundulus nottii* (1) and *Fundulus escambiae* (2) and sisters *F. blairae* (4) and *F. dispar* (5). (b) "Microvicariant" speciation between sisters *Etheostoma chlorosomum* (1) and *E. davidsoni* (2). Note that Wiley and Mayden (1985) ascribe the same vicariant event to both clades and that the difference between "dumbbell" and "microvicariance" resides in the relative amount of range that is subdivided. From Wiley and Mayden (1985); used with permission, The Missouri Botanical Gardens.

America for a number of aquatic clades where the most recent speciation event resulted in coincident speciation of sister species east and west of the Mobile River basin. In some cases, such as *Fundulus nottii* and *Fundulus escambiae*, the ancestral range was divided equally ("dumbbell"). In other cases, such as *Etheostoma chlorosomum* and *E. davidsoni*, the ancestral range was divided very unequally, with *E. davidsoni* occupying very little of the original ancestral range (assuming that the range is accurately estimated by summing descendant ranges). With paleontological data, it is sometimes hard to distinguish between "dumbbell" and "micro-vicariant" versions of allopatric mode I speciation. However, several clades of trilobites show frequent instances of vicariant differentiation (Lieberman and Eldredge, 1996; Lieberman, 1997; Meert and Lieberman, 2004). This is true for other fossil clades as well (e.g., Rode and Lieberman, 2004, 2005).

There are several criteria that can be used to identify the occurrence of allopatric speciation. These have been identified by several authors in a phylogenetic context including Croizat et al. (1974), Brooks et al. (1981), Nelson and Platnick (1981), Wiley (1981a, 1988a, b), Cracraft (1982), Brooks (1985), Funk and Brooks (1990), Brooks and McLennan (1991, 2002), Lieberman and Eldredge (1996), Lieberman (2000a, 2003a), and Rode and Lieberman (2004, 2005). These criteria in some form or another are even manifested as far back as the works of Jordan (1905). They were codified in numerical order in a review of speciation by Coyne and Orr (2004). To summarize, they are as follows:

1. Species borders correlated with existing geographic or climatic barriers (which presumably are younger than the range of the ancestral species).
2. Allopatry of young sister species.
3. Geographic congruence between species borders (or secondary hybrid zones) among species of different clades.
4. Absence of sister species in areas where geographic isolation is unlikely.

Each of these criteria require three components to assess the probability that allopartic speciation occurred: (1) a phylogenetic hypothesis of relationships, (2) a detailed knowledge of the distribution of the species analyzed and their closest relatives, and (3) a knowledge of the correlation of geographic history or climatic history and the biogeographic ranges of species. Wiley (1981a) suggested that Criterion 3 should obtain if vicariant speciation affected entire biotas, resulting in replicate speciation patterns among clades. Wiley and Mayden (1985) demonstrated that the areas of endemism can be identified through the use of phylogenies and Criteria 1–3. They consider the identification of replicate speciation events to be one of the major contributions of systematics to evolutionary biology because study of replicate events affords the opportunity to see if there are common patterns in vicariance speciation between clades that are relatively unrelated; the occurrence of such common patterns of evolution across geographic space in independent clades would suggest that common earth history factors played the primary role in influencing speciation. Thus, climate or geological change may have played the primary role in influencing patterns of evolution in the regional biota of interest.

The phylogenetic pattern that might emerge from repeated vicariant speciation in a clade would depend on the history of geographic subdivision. Although we

might expect dichotomous patterns, with a single ancestral population divided into two populations by a geographic barrier (Fig. 2.2a), polytomies are certainly possible, for instance when numerous populations are isolated effectively simultaneously by the coeval emergence of several geographic barriers (Fig. 2.3a). The challenge is that Mode II patterns may mimic Mode I patterns (Fig. 2.2b, 2.3b), as discussed below.

Allopatric Speciation, Mode II Peripatric Speciation   Peripatric speciation occurs when a small number of individuals become isolated from the central population and differentiate. Peripatric is another term for peripheral isolate speciation (Mayr, 1963), except that it represents the general case rather than referring only to situations where a small population colonizes new ranges or becomes isolated around the periphery of the ancestral range. However, the chances that a small isolate within the range of an ancestral species will remain isolated long enough to differentiate is considered by Coyne and Orr (2004) unlikely, unless there is peripheral isolation. The manner in which patterns resulting from peripheral isolation can mimic Mode I is discussed below.

Distinguishing between Allopatric Modes of Speciation   Except for the possibility that genetic drift will be more likely to drive a small population to differentiate, evolutionary mechanisms affecting both modes of speciation are similar (Coyne and Orr, 2004). The community and ecological contexts of the two modes can be quite different. For example, in Mode I speciation, we might not expect the ecological context of the descendants to be any different from that of the ancestor. Entire biotas are affected, not single clades. Thus, we would expect a greater amount of ecological conservatism, especially as that conservatism related to biotic and abiotic parameters of larger landscapes. This kind of conservatism was investigated by Peterson et al. (1999) where there is a correlation between Grinellian niche parameters and degree of evolutionary relationship. More closely related species share similar niche parameters. It can also be observed using phylogenetic trees. Sister species share niches and the geographic range of one species can be predicted using its sister species niche when niche parameters are computed and then projected on geographic space (McNyset, 2009).

Mode II allopatric speciation can mimic Mode I if peripheral isolates later disperse from their original biogeographic areas to occupy a larger area (Fig. 2.2b). However, in such a case we might not expect to see a common pattern of replicate speciation among constituents of different areas that are defined as areas of endemism. In other situations, Modes I and II allopatric speciation can produce the same kind of phylogenetic pattern, complete with replicate patterns of speciation. For example, the same kind of pattern might emerge from multiple microvicariant events and peripatric speciation, resulting in soft polytomies due to ancestral conservatism (Fig. 2.3a, b). Colonization of newly available habitat may produce replicate patterns among clades as discussed in the Hawaiian example below and shown diagrammatically in Figure 2.4a, b. In addition to phylogenetic analysis, other types of information can help tease apart the modes of speciation involved. Brooks and McLennan (2002) suggest three additional types of information, to which we add a fourth.
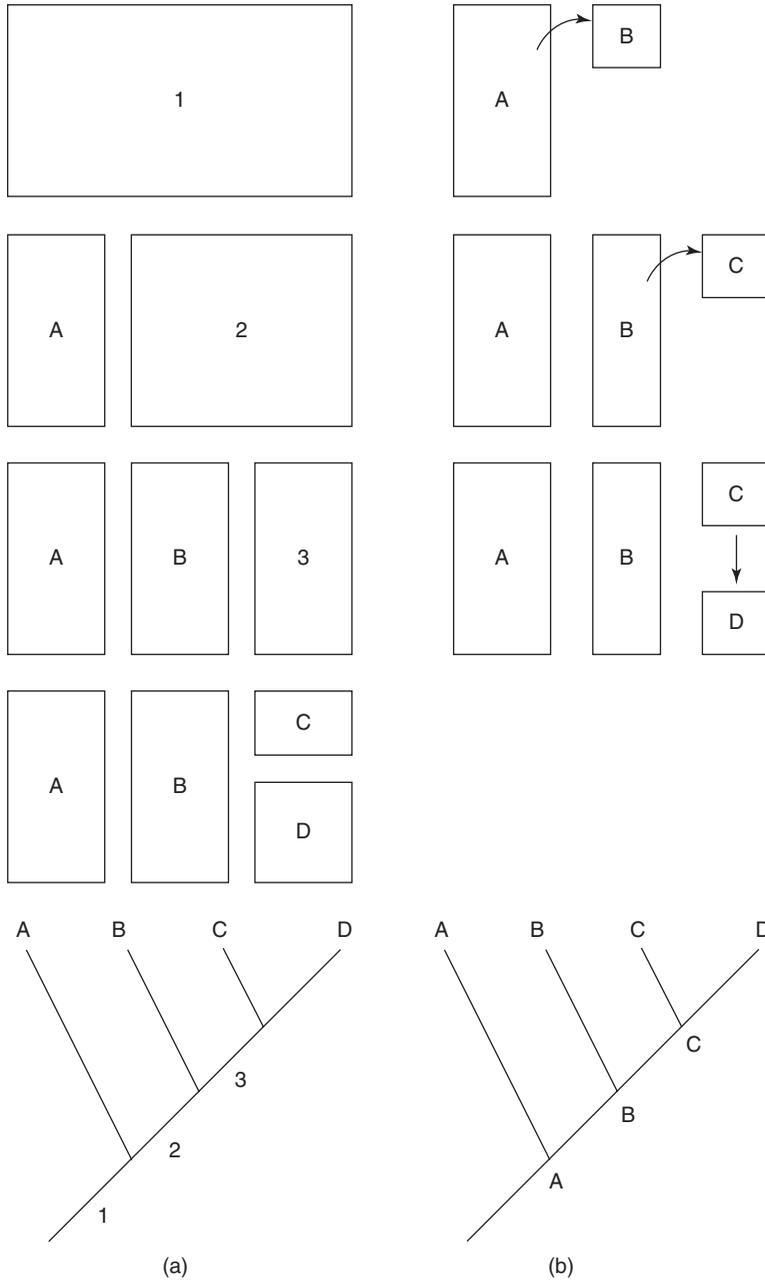
Figure 2.2. Modes I and II speciation can result in the same dichotomous pattern of descent: (a) the vicariance of a widespread ancestral species into four descendants and (b) dispersal, speciation, and range extension can result in the same distributional and descent pattern.
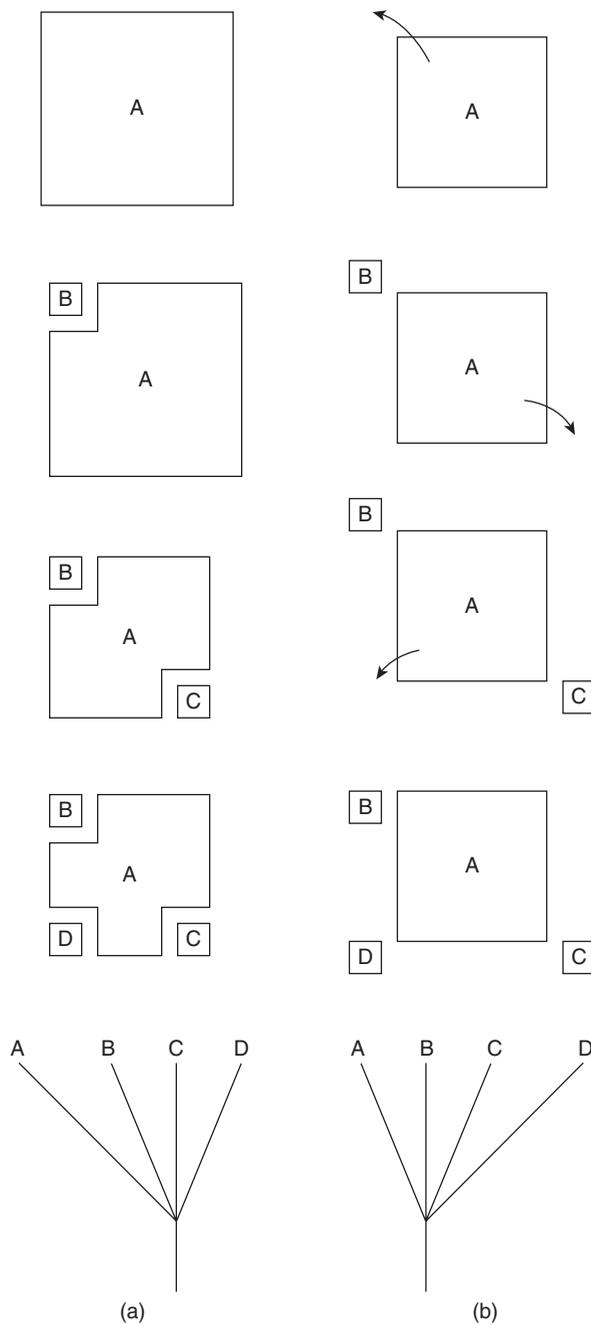
Figure 2.3. Mode I microvicariant and peripatric speciation can result in the same pattern: (a) sequential microvicariant events beginning with a widespread ancestral species and (b) dispersal and peripatric speciation establishing the same biogeographic and descent pattern. Note that the polytomies in both cases are "soft polytomies" and that the actual events occurred in sequence.
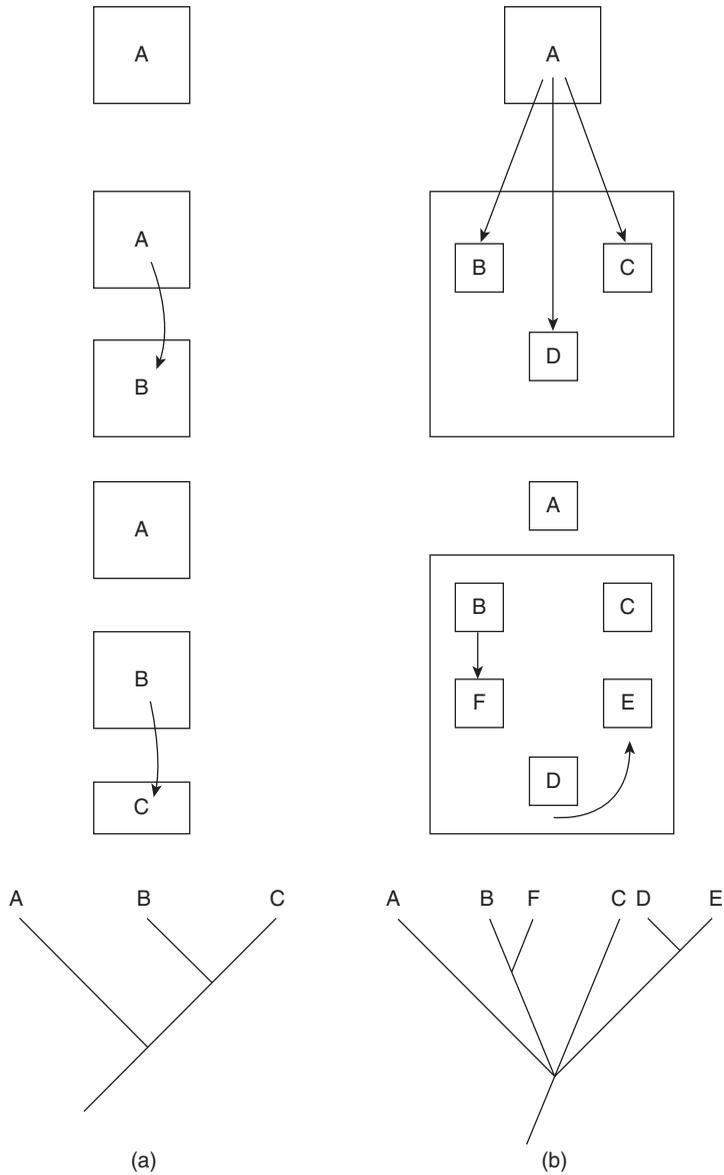
Figure 2.4. Two speciation scenarios, the Hawaiian scenario (a) and (b) the New Lake scenario. In the Hawaiian scenario, a dichotomous pattern of descent is the result of sequential colonization of new land as it emerges from the sea floor. In the New Lake scenario, suitable habitats in a newly formed lake are colonized more or less simultaneously as the ancestral species establishes itself in the lake, with subsequent colonization in other parts of the lake.

1. Can we correlate disjunction with geological processes that might have caused the disjunction, implying the passive disjunction of the ancestral range? Alternatively, it may be possible to identify barriers, with available habitat across barriers that predate speciation events, implying the active type of Mode II speciation.

2. It is also worth determining, by study of descendants, whether ancestral species were characterized by high or low gene flow. High gene flow species spread evolutionary novelties, promoting differentiation more quickly than low gene flow species. While high gene flow inhibits differentiation in the absence of isolation, it promotes differentiation once isolated.

3. The geography of the present distribution of the descendant species is also of interest. Given no subsequent dispersal, does it appear that one species (or several) are peripheral isolates?

4. What of the other species inhabiting the same areas? Do they show distributions and phylogenetic relationships similar to the clade studied?

Of the four types of information, 1 and 4 are most accessible to systematists. In the case of the dichotomous phylogeny, we can contrast two examples that show how knowledge of the geology can discriminate between vicariance (Mode I) and peripheral isolates speciation (Mode II).

North American Freshwater Fishes. The fish faunas of North America are a particularly good laboratory for studies of speciation patterns, because there are so many systematic ichthyologists investigating the phylogenies of these fishes and because the ranges of these species are well understood. The predominant pattern of recent speciation is Mode I based on the fact that most terminals on the phylogeny are distributed allopatrically and their clades seem to predate the most recent isolation events (for a review see Wiley, 2002). Detailed studies by Wiley and Mayden (1985) and Mayden (1985, 1987, 1988b, 1992) portray a picture of clades whose ancestral distributions are older than the present barriers separating living species. Indeed, Mayden (1988b) found that present-day distributions of clades he studied were more highly congruent with pre-Pleistocene drainage patterns than the drainage patterns produced by Pleistocene glaciation. In terms of point 4, Mayden (1988b) found the phylogenies of seven clades of fishes from the Central Highlands of the United States congruent with the known history of that area. In contrast, Wiley and Mayden (1985) found very little similarity among the clades inhabiting the northern coast of the Gulf Mexico, but did find that most clades contained sister species correlated with a shift in drainage pattern of the freshwater streams of the area that apparently affected clades of both freshwater and coastal marine groups.

Hawaiian terrestrial floras and faunas. Mayden's (1988b) analysis showed the power of congruent phylogenies and geography to make the case for Mode I allopatric speciation. Wiley (1981a) thought that it was this congruence that made the case for what Brooks and McLennan (2002) later termed the *passive mode* of allopatric speciation. But the examples from Hawaii in the book edited by Warren Wagner and Vicki Funk (1995) demonstrate that Wiley (1981a) was not entirely right: active mode allopatric speciation can sometimes also result in geographic congruence between different clades of organisms. Hawaii shows the classic Hennigian progression rule of phylogenies that map island hopping from the oldest to the youngest island in the histories of many clades. Some of the patterns, however, are complex (back island hopping and intra-island speciation), and even incongruent in some cases; still, typically the most basal member of a clade is on the oldest island (frequently Kauai) and the crown species is on the youngest island (Hawaii) (Wagner and Funk, 1995). Thus it is a combination of knowledge of the geology of the area (islands arise sequentially and the timing is fairly well known) and the

congruence of the phylogenies that leads to the opposite conclusion that one might reach when studying continental or oceanic patterns.

Parapatric Speciation

Parapatric speciation obtains when there is partial separation of the gene pool of an ancestral species resulting in a narrow primary zone of hybridization. Divergence occurs because of selection against hybrids within the hybrid zone, creating a genetic sink. Such divergence is usually said to occur along a steep selection gradient (Endler, 1977). For the systematist, two species that have undergone a parapatric speciation event will appear as sister species and parapatric in their distribution, with ranges abutting and a hybrid zone present. The major contribution of the systematist is to set the preconditions for further study of the system. The first precondition is that the species involved must be sisters. Secondary contact between nonsister species is evidence that the contact zone is a secondary contact zone, not a primary contact zone. However, there can also be secondary contact between sister species. Determining the parameters of the contact zone, whether it is primary or secondary, narrow and a genetic sink, or broad and introgressing, requires the tools of phylogeography and population geneticists (Avise, 2000; Coyne and Orr, 2004). It could be the case that secondary contact is implicated in reinforcement of reproductive isolation between species (e.g., Hoskin et al., 2005).

Sympatric Speciation

Sympatric speciation is speciation within the range of an ancestral species where geography does not play a causal role in species formation. Two broad classes of sympatric speciation can be identified—reticulate and divergent. Reticulate sympatric speciation results in the origin of a new species via a hybridization event between two existing species, creating allopolyploid (increased in chromosome number) and recombinate (no increased in chromosome number) species. Divergent sympatric speciation results in the origin of a new species in response to a number of mechanisms, including disruptive selection, sexual selection, and autopolyploidy.

Reticulate speciation may be homoploid or polyploidy and is rare in animals but relatively common in plants (see Rieseberg and Willis [2007] for a review of plant speciation). Homoploid speciation is relatively rare and is characterized by rapid karyologic evolution without a change in chromosome numbers.

Recombinational speciation (Grant, 1981) obtains when two species give rise to a new species that is fertile, true breeding and reproductively isolated from its parents without a change in the number of chromosomes. The mechanism is chromosomal rearrangement, and the daughter species has a restricted range within the sympatric ranges of the parental species.

The "trick" to reproductive isolation is the presence in the hybrid of several chromosomal rearrangements that are different, in combination, from the parental species. Coyne and Orr (2004) provide a nice discussion of recombinational speciation in the wild involving three species of *Helianthus* that largely results from work by Loren Rieseberg and his research group. The details of the genetic mechanisms involved are not of direct interest here, and we refer you to Coyne and Orr's (2004) discussion and Rieseberg and Willis (2007) for a recent discussion of plant speciation

in general. Of more interest to systematists is the expected pattern of speciation. The *Helianthus* story involves two widespread North American species of outcrossing, annual, diploid plants with similar chromosome numbers—*Helianthus annuus* and *H. petiolaris*. They are not sister species (e.g., Rieseberg, 1991). A third species, *H. anomalus*, is found in Utah and Arizona well within the ranges of both parental species. It has some chromosome rearrangements of each parent, but some unique chromosomal rearrangements unique to it (Rieseberg et al., 1995).

Heteropatric speciation (Smith, 1966; Getz and Kaitala, 1989). If members of an ancestral gene pool use different resources, then it is possible that they will diverge if fidelity to something like a food source or a microhabitat is coupled with assortative mating. Although the frequency (and even reality) of heteropathic speciation is debated, it is easy to predict the phylogenetic and geographic pattern that would obtain if heteropatric speciation is a possible explanation for any two species: they must be sympatric and sister species (Wiley, 1981a).

Reticulate speciation, or speciation via hybridization. This mode leads to the origin of a new species via hybridization between individuals of two parental species. Speciation is thought to occur either instantaneously or over a restricted number of generations. As we will see in Chapter 4, phylogenetic analysis that includes taxa of hybrid origin doesn't necessarily lead to a hypothesis of hybrid origin. Thus, there is no clear phylogenetic pattern to guide the investigator and other biological criteria must be used. In the case of cichlids from Lake Barombi Mbo in Cameroon, Schliewen and Klee (2004) suggested that hybridization of ancestral allopatric lineages that became sympatric upon invasion of the new lake explains part of the origin of the species flock. This conclusion was reached based on discord between mitochondrial and nuclear gene phylogenies.

IDENTIFYING MODES OF SPECIATION IN THE FOSSIL RECORD

Phylogenetic approaches. There are aspects of paleontological data that make it challenging at times to precisely determine the mode of speciation responsible for the diversification of a pair of sister taxa. This is because some of the information described earlier that is helpful for elucidating modes of speciation is not available with paleontological data, due to the incompleteness of the fossil record. Still, even with fossil taxa, inferences can be made and hypotheses can be tested about the primary mode of speciation responsible for the diversification of a clade, and individual species within that clade. The manner that modes of speciation governing the diversification of clades (fossil and extant) can be considered in a broader biotic context will be outlined in Chapter 9; here our focus is on elucidating the single clade case. Stigall Rode (2005a, b) focused on speciation patterns in Devonian brachiopods. She first performed a phylogenetic analysis on all available species in the closely related genera *Floweria* and *Schuchertella*. She then substituted the species names with their area of geographic occurrence and used a parsimony algorithm (described more fully in Chapter 7) to infer the geographic distributions of the ancestral nodes (Fig. 2.5). Then the patterns of character change in geographic distribution can be traced on the tree. Instances of contraction in geographic range associated with cladogenesis represent potential vicariance; episodes of expansion in geographic range associated with cladogenesis represent potential examples of
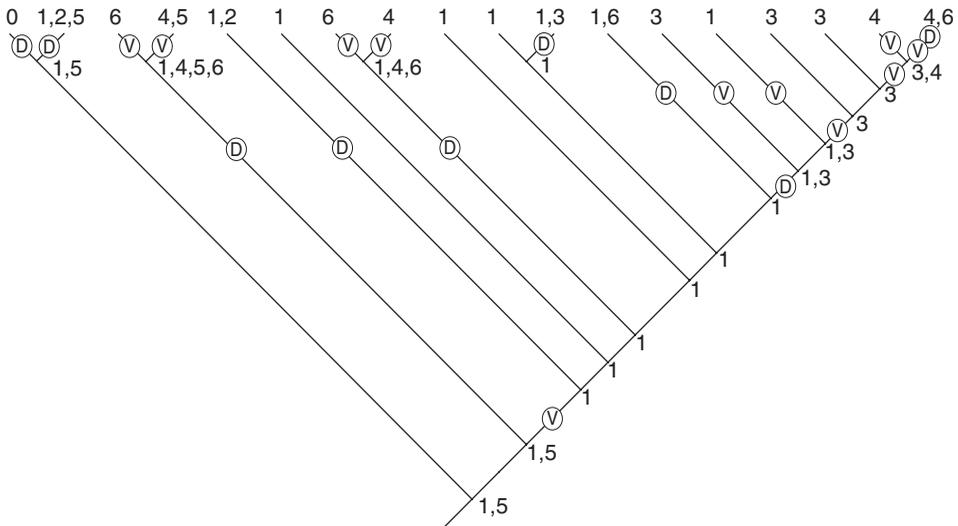
Figure 2.5. Area cladogram of species of Devonian brachiopods from Stigall Rode (2005b) with inferred episodes of speciation by vicariance denoted by "*V*" and inferred episodes of speciation by dispersal denoted by "*D*." Areas are shown mapped to nodes and terminals where "*0*" is Europe, "*1*" is the Northern Appalachian Basin, "*2*" is the Southern Appalachian Basin, "*3*" is the Michigan Basin, "*4*" is Iowa and the Illinois Basin, "*5*" is Missouri, and "*6*" is the western United States. These were all areas of marine endemism during the Devonian period. Used with permission of A. Stigall, Ohio University, and *Journal of Systematic Palaeontology*, Taylor & Francis.

peripatric differentiation; episodes of range conservatism at cladogenesis represent potential examples of sympatric or within area differentiation. Abe and Lieberman (2009) analyzed speciation modes and patterns in a clade of Devonian trilobites undergoing a dramatic evolutionary radiation (Fig. 2.6). They found that much of the radiation was occurring within a single biogeographic region, Bolivia, a biodiversity hotspot. In the Devonian, Bolivia was a high latitude, rich, marine habitat for trilobites (and other taxa). The episodes of speciation in Bolivia comprised either sympatric differentiation or smaller scale allopatric differentiation. The latter is perhaps more likely given that Bolivian geology provides evidence for three distinct marine basins with semi-endemic faunas, but a more definitive answer will require additional explorations of the geology of that region and additional range data. Congreve and Lieberman (2010) focused on modes of speciation not in a radiating clade but instead in one passing through the end Ordovician mass extinction, perhaps the second most severe mass extinction, as measured by percentage diversity loss (upward of 70 percent) in the history of life. Most cladogenesis in this clade of trilobites seems associated with range expansion, and thus potentially peripatry (Fig. 2.7) although there is some more limited episodes of vicariance. In turn, each of these studies provides evidence for shifting modes of speciation throughout a given clade's phylogenetic history. For example, clades as a whole seem to show episodes of vicariance followed by within area differentiation followed by range expansion at times followed by subsequent vicariance. Moreover, these transitions appear to be associated with episodes of climatic change, principally global warming

Famennian

Frasnian

Givetian

Eifelian

Emsian

Pragian

Lochkovian
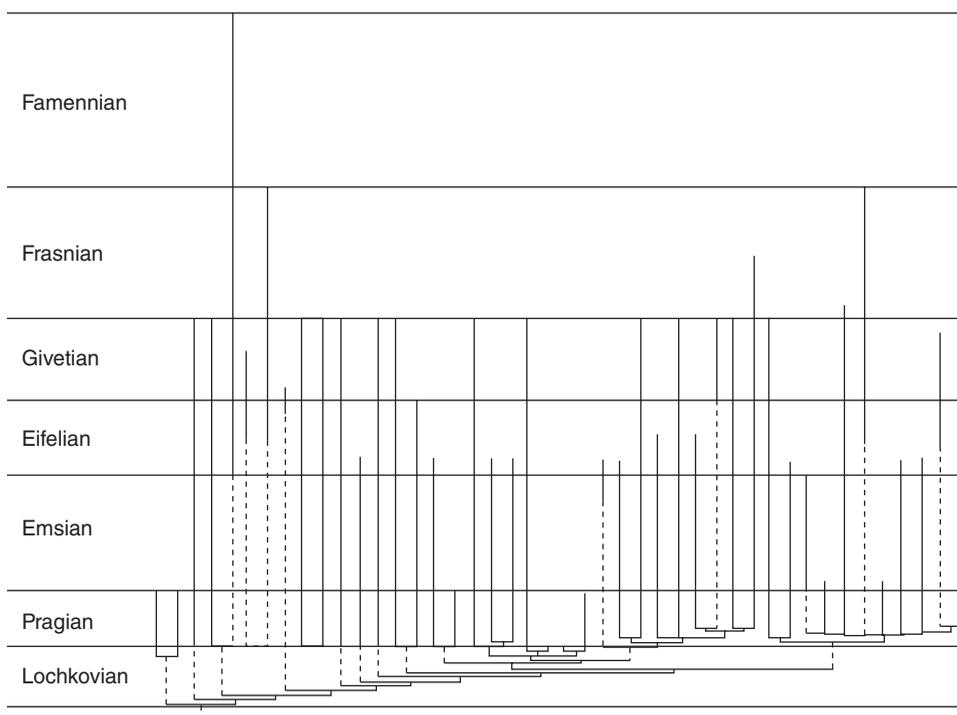
Figure 2.6. Phylogeny of *Metacryphaeus* group calmoniid trilobites from Abe and Lieberman (2009) showing the evolutionary radiation of the group and the pattern of speciation during the different stages of the Devonian period. Dotted lines represent inferred origination based on time of sister taxon divergence. Used with permission of *Evolutionary Biology* and Springer.

and cooling, which triggered rising and falling sea-level (due to the waxing and waning of ice sheets). The intervals of sea-level fall often correspond with episodes of vicariance, whereas the intervals of increasing sea-level instead correspond to episodes of range expansion and possible peripatry. Such patterns might be expected with marine invertebrates (and vertebrates). By contrast, different scenarios might obtain with terrestrial taxa, in terms of the relative effects global warming and cooling might have on facilitating isolation and range expansion; still, with these it is possible to consider speciation modes in the context of tempo of evolution and climate change. For instance, Maguire and Stigall (2008) considered modes of speciation within fossil Equinae, an important fossil group that has figured in several macroevolutionary studies. They found episodes of repeated range expansion and dispersal at speciation followed by vicariance followed by subsequent dispersal. Additional examples considering speciation mode in fossil taxa can be found in Lieberman and Eldredge (1996), Lieberman (2000a, 2003a), Stigall Rode (2005a, b), and Hendricks and Lieberman (2008).

Punctuated equilibria. Probably the paradigm example of an analysis of speciation mode in the fossil record comes from Eldredge's (1971) study of allopatric speciation in trilobites, which forms the hallmark example of the punctuated equilibria hypothesis (Eldredge and Gould, 1972). Eldredge's (1971) analysis was based explicitly on phylogenetic information, described in greater detail in Eldredge
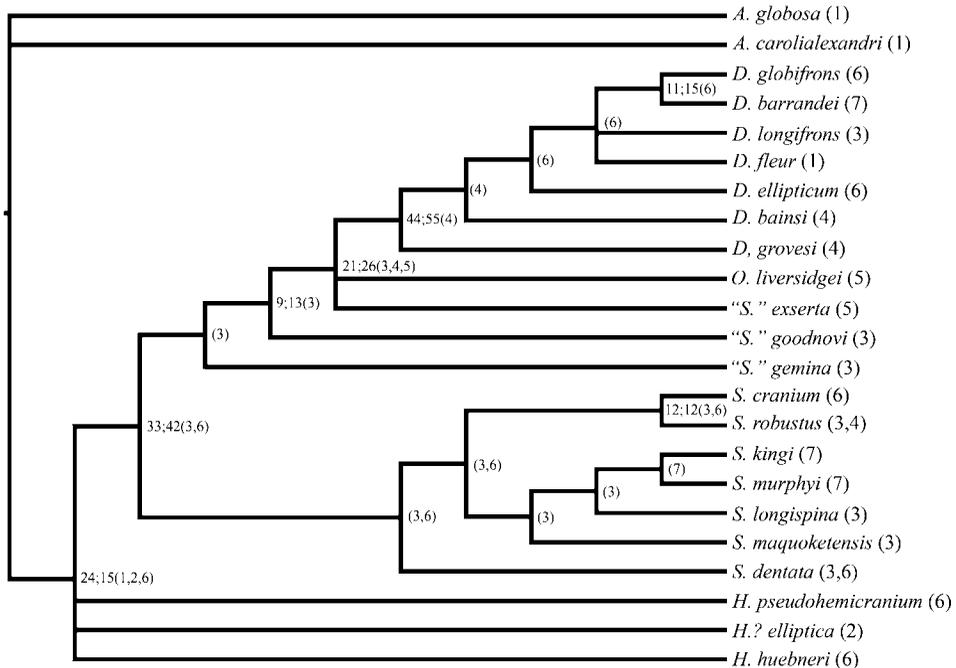
Figure 2.7. Results from phylogenetic and biogeographic analysis of Deiphonine trilobites from Congreve and Lieberman (2010). Strict consensus of six most parsimonious trees shown and genera are labeled. Values at nodes in plain text are Bootstrap and Jackknife values. Values that are bracketed, i.e., (1, 2), are the biogeographic areas used. 1, Bohemia; 2, Tarim Plate; 3, Eastern Laurentia; 4, Northwestern Laurentia; 5, Australia; 6, Baltica; 7, Avalonia.

(1973), and showed that peripatry was the dominant mode of speciation in the clade of Devonian trilobites studied. Although as outlined in the original formulation of punctuated equilibria speciation was assumed to primarily involve peripatry, i.e., allopatric mode II speciation, there is no need that punctuated equilibria could not also involve vicariant differentiation, i.e., allopatric mode I speciation (Lieberman, 2000a). Indeed, Vrba (1980, 1985) partly reformulated the theory of punctuated equilibria primarily to involve vicariant differentiation, rather than peripatry, that was triggered by episodes of climatic change that fostered geographic isolation of populations of species.

Punctuated equilibria was also significant, not only for its focus on species origins, primarily by allopatric speciation, but also because it asserts that in the long term, over millions of years, most species were stable throughout their evolutionary history. This assertion was in direct contrast to the Neo-Darwinian view (e.g., Dobzhansky, 1937; Mayr, 1942; Simpson, 1944) that posited that species were lineages that were gradually transformed into other lineages over deep time, without a distinct breakpoint separating one species from another (referred to as phyletic gradualism). Studies conducted since Eldredge (1971) and Eldredge and Gould (1972), for instance, Lieberman et al. (1995) and Eldredge et al. (2005), continue to reiterate the notion that stasis is the rule for most, though not all, species. The

long-term stability of species over millions of years provides additional support for the notion that species are individuals that possess a unique birth and death point (although the species as individual formulation does not require stasis). Still, it is easier to view and identify species as individuals in light of long-term stasis as opposed to ephemerality or evanescence. Given that punctuated equilibria emphasizes stasis and allopatric differentiation, it also posits that species evolution occurs by cladogenesis. Considering all of these points, the evolutionary species concept that we have endorsed here fits very comfortably ontologically and epistemologically with the notion that punctuated equilibria is a, or the, primary descriptor of species evolution.

THE EVOLUTIONARY SPECIES CONCEPT, SPECIATION, AND ECOLOGY

Phylogenetic systematics has much to offer to the disciplines of conservation biology and ecology. One of the basic tasks in the study of biodiversity and conservation is counting the number of species in an area. The number of species actually counted will depend on the number of species recognized by taxonomists, and the number of species recognized by taxonomists will depend on the species concept employed, so species concepts make a difference.

The diversity of fishes and birds offer striking examples of how diversity calculations differ under different concepts. During the first half of the twentieth century the Biological Species Concept (BSC) was extensively applied to both fishes and birds. The end products of this application were the publication of various checklists purporting to furnish fisheries biologists and ornithologists with those species thought valid. In ichthyology, the number of species has increased with time, and this increase is not due to the discovery of new populations of species unknown to science but instead to the increasing acceptance of various forms of the ESC (Wiley, 2007).

Ornithology, where Ernst Mayr's influence was best felt, is also turning away from the polytypic species of the BSC and toward the recognition that species are those entities that share phylogenetic rather than tokogenetic relationships. This is usually expressed through application of the Phylogenetic Species Concept (Zink and McKitrick, 1995; Peterson and Navarro-Siqüenza, 1999), although it is evident to us that it is, in fact, little different "in principle" (Zink and McKitrick, 1995:710) from the ESC. Under the BSC, there are 110 bird species endemic to Mexico while under the ESC there are 249 endemic species. By 2004, Navarro-Siqüenza and Peterson (2004) had refined the number from 135 BSC species to 323 "evolutionary/ phylogenetic" species. The new view of avian diversity in Mexico changes our understanding of regional endemism and calls for new assessments of areas in need of conservation.

EMPIRICAL METHODS FOR DETERMINING SPECIES LIMITS

Wiens (1999) in his review article on polymorphic characters in phylogenetics commented that although much attention had been paid to species concepts, little atten-

tion had been paid in recent literature to developing empirical methods of delimiting particular species (papers by Davis and Nixon, 1992; and Sites and Crandall, 1997, were cited as exceptions). However, since 1992, there has been an increase in the number of papers detailing methods that aid in such delimitations. Many of these have been critically reviewed by Sites and Marshall (2004), and their analyses form the basis for this section.

Sites and Marshall (2004) analyze 12 methods proposed to delimit species. They divide these methods into two categories: tree-based and nontree-based approaches. Tree-based approaches involve a posteriori delimitation, after conducting a phylogenetic analysis. Such methods may use individual organisms or some geographically constrained a priori populations as basic units of analysis. Nontree-based methods use various grouping strategies in an attempt to delimit the units (either largest or smallest) that can be diagnosed. Some of these methods seek to make the common practices of taxonomists more repeatable; others seek to use the knowledge that has been gained from phylogenetic analysis, biogeography, molecular, and population genetics.

Nontree-Based Methods

Species are obviously important as the basic entities of phylogenetic analysis (Hennig, 1966, and many others), but resolving the tree of life will take many generations, and in the meantime, there are many areas of biology that require us to delimit species in the absence of an explicit phylogeny. Species are the basic units of analysis in, among other fields, macroecology (the missing species problem; Blackburn and Gaston, 1998), ecological forecasting of native and invasive species (Peterson, 2003), global biodiversity assessments (Wilson and Peter, 1988; Krishtalka et al., 2002), macroevolution (Eldredge and Cracraft, 1980), and conservation biology (Agapow et al., 2004). The complete phylogeneticist not only reconstructs phylogenies, but also practices all aspects of systematics, including alpha taxonomy and the description of new species. Thus, there are many circumstances in which species are delimited in the absence of explicit phylogenetic hypotheses and nontree-based methods are practiced. Sites and Marshall (2004) review seven methods to delimit species that do not depend on prior knowledge of the phylogeny. Most of these methods require molecular data while one can be applied exclusively to morphological data in the absence of molecular data. Because the vast majority of taxonomic work is performed in the absence of genetic data, the traditional method of finding morphological discontinuities, with due concern to distributional data and mode of reproduction will be discussed first, followed by refinements of this traditional method.

1. *Morphological/Genetic Discontinuities (M/GD).* The most venerable of all methods, M/GD is based on the assumption that variability within species is less than variation between species and that morphological or genetic discontinuity is the mark of lineage isolation. M/GD is suitable for either morphological or genetic data.

Sites and Marshall (2004) list the criterion of "morphological discontinuities" (e.g., Cronquist, 1978) separately from population aggregation analysis (PAA, see

below). We consider PAA as a refinement of the traditional method of simply searching for such discontinuities, and adding Mallet's (1995) concept of genetic clustering simply extends the number of potential diagnostic characters to the molecular realm. However, use of multivariate analysis to discover shape differences (e.g., Macleod, 2002; many techniques reviewed in Zelditch et al., 2004) can lead to the discovery of complex diagnostic differences between species and reveal subtle discontinuities not readily apparent in traditional qualitative character analysis.

The use of morphological discontinuities can sometimes be confounded if populations are drawn from a wide geographic area. A widely distributed species might show geographic variation while maintaining cohesion due to local selection pressures or chance. Or, they may appear to show such a phenomenon for certain characters but not for others. An example is furnished by various analyses of the North American killifishes *Fundulus zebrinus* and *F. kansae*. Populations of these putative species are widely distributed from west-central Missouri south to the Rio Grande and west to the eastern slope of the Rocky Mountains (with populations widely introduced west of the Rockies). They have been variously considered species (e.g., Minckley, 1973; Collier, 1979) or subspecies (e.g., Echelle et al., 1972). Poss and Miller (1983) used principle components analysis and autocorrelation techniques to demonstrate that certain diagnostic characters thought to diagnose the two species were highly autocorrelated with distance between samples. Based on this, they concluded that a single species was present, *F. zebrinus*. The reasoning is somewhat analogous to that presented in the next section. However, subsequent work by Kreiser (2001), using mitochondrial sequences, shows a distinct break at the Canadian River (Fig. 2.8), and these authors support the hypothesis that there are two species, not one. Apparently the variation observed in morphology is the "ghost of geographic variation past" or maintained by selection if indeed there are two independent lineages.

Approaches based on morphological discontinuities are of course the way that paleontological species identification proceeds, too. Trilobites, for instance, have a roughly 300-million-year history in the fossil record. Numerous cases of species definitions involving Cambrian trilobites roughly 525–510 million years old (Lieberman, 1998, 1999, 2001, 2002a) or Devonian trilobites roughly 390–370 million years old (Lieberman et al., 1991; Lieberman, 1993, 1994; Lieberman and Kloc, 1997; Abe and Lieberman, 2009; Congreve and Lieberman, 2010) proceed in parallel to neontological studies. In particular, narrowly ranging species are easiest to diagnose, broadly ranging species may show more geographic variation, and a series of qualitative, meristic, and morphometric character data sets can be used to define clusters of organisms representing putative species. Further, typically trilobite species vary within distinct limits and do not overlap one another.

2. *Population Aggregation Analysis (PAA)*. PAA (Davis and Nixon, 1992) is a formal protocol for the traditional approach of grouping samples into discrete groupings and, as such, codifies the traditional practices advocated by those who seek to identify morphological or molecular discontinuities that can be correlated with species limits. Davis and Nixon distinguish heritable "attributes" as being either "traits" or "characters." Traits are attributes that vary among members of a single local population while characters are fixed among individuals of a single population. For each local population, an attribute profile is prepared by tabulating all traits
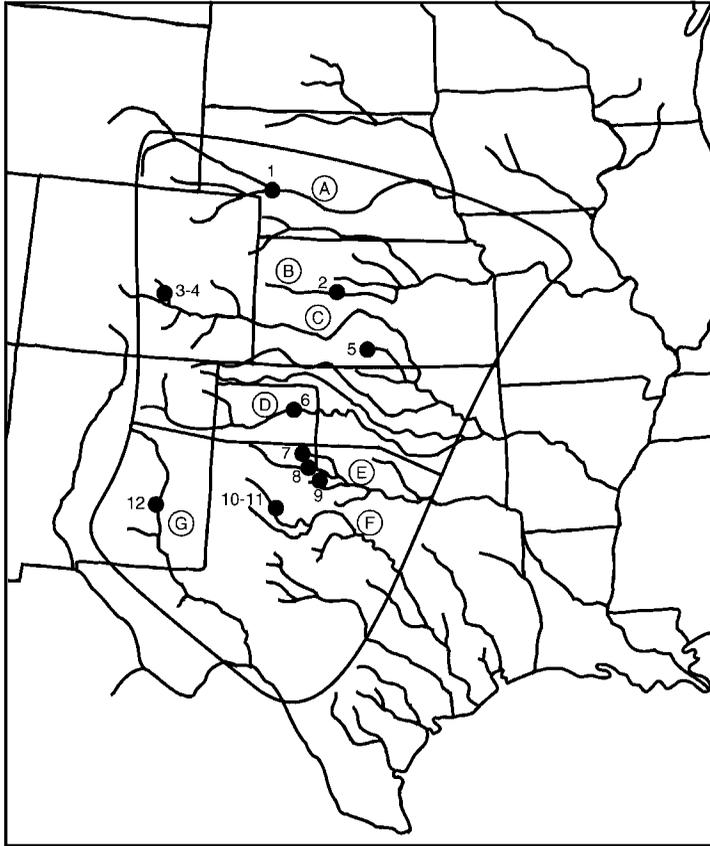
Figure 2.8. Range map of two sister species of killifishes, *Fundulus kansae* (north) and *F. zebrinus* (south). Dots are population samples studied by Kreiser (2001), and letters are drainage systems. The sharp break between the species occurs between the Canadian (D) and Red (E) river drainages. From Kreiser (2001); used with permission, The American Midland Naturalist.

and characters. Profiles of each population are then compared, and attributes are characterized as either character attributes (fixed differences among populations) or trait attributes (those that vary among populations). Species are formed by aggregating populations that share characters. If there are no fixed differences, then all populations belong to the same species. This is using the phylogenetic species concept that defines *species* as the smallest aggregation of populations that can be diagnosed by characters.

Fixed differences equates to crisp and clear diagnostic characters, exactly the kinds of characters that convince colleagues that two or more species are represented. There is, however, the issue of what *fixed character* means exactly. In addition, Davis and Nixon (1992) point out that PAA can underestimate the number of diagnosable units if a sufficient number of characters is not sampled, and it can overestimate the number of diagnosable units if the number of specimens is not sufficient

(i.e., mistaking polymorphisms for fixed differences through examining few individuals). Of course, these are general problems; they are not confined to PAA. So, there is a more general issue: how many characters and how many specimens are needed? Brower (1999) also points out that there is a question of what constitutes a character when it comes to DNA data, the entire string of sequence (a haplotype, for example) or each base pair. See CHA below for a discussion on this point.

The "Fixed Character" Issue. Sites and Marshall (2004:202) call attention to a perceived deficiency in PAA "character fixation (is) difficult to show at conventional levels." Claiming that a species truly has a fixed character, one with a frequency of 100 percent, would require that all individuals of a species be examined, clearly an impossible goal. Examining a large number of specimens would ensure that polymorphisms occurring at intermediate levels are detected, but rare polymorphisms (on the order of $p = 0.01$) require a prohibitive number of samples for the average empirical study, as shown by Wiens and Servedio (2000:632) based on methods developed by Swofford and Berlocker (1987) to detect polymorphisms for polymorphic alleles. Even if the statistical threshold that there is only a 5 percent probability that a particular trait is polymorphic is adopted, the samples sizes needed to reach this conclusion are large and, for most studies, claiming fixation amounts to an assertion, not a conclusion.

Wiens and Servedio (2000) suggest that a more modest, but valid, claim can be made. Consider what we are attempting to demonstrate. If the sample we are working with is a sample drawn from an independently evolving lineage, then we would expect that gene flow between this lineage and its closest relative is negligible, if not zero. Such a finding would corroborate the hypothesis that the sample represented an independently evolving lineage. Thus, Wiens and Servedio (2000) suggest that diagnostic characters present at a frequency of 95 percent or higher in one population and 5 percent or lower in another population are "close enough" and indicate that little or no gene flow is occurring between the two presumed species. They frame the question in the following manner: what is the probability that at least one of a number of apparently "fixed" diagnostic characters meets the criterion of being a diagnostic character, a character "fixed" at some frequency determined by the investigator to be indicative of the hypothesis of lineage independence? The null hypothesis states that a rare homolog to the presumed diagnostic character is present at a frequency greater than "p = the selected cut-off frequency." The investigator can select any cut-off frequency, 5 percent being reasonable. The data needed to accept or reject the null hypothesis are the number of individuals sampled (n), the number of characters surveyed for potential diagnostic differences (c), and the number of characters found to be "fixed" diagnostic characters after the survey (k). They provide a modified version of the binomial test designed to test the null hypothesis that any rare characters are actually present at a predetermined frequency (p, the frequency cut-off; e.g., 5 percent or that set by the investigator). The probability can be calculated from a formula. A number of approaches mostly use molecular methods and are built around the idea that separate lineages are characterized by one or another form of genetic isolation. Each requires different assumptions.

The *Field for Recombination (FFR)* method of Doyle (1995) is a variant of PAA used for nuclear genes that show codominance. Carson (1957) asserted that sexually reproducing species with populations interconnected by gene flow should

be characterized by a field of recombination. Doyle (1995) reasoned that if this is true, then the distribution of alleles should define species' boundaries better than the gene trees of these alleles. For example, a neighbor-joining tree of the Class II major histocompatability complex DQB-1 alleles from Gaur et al. (1992) shows that no species of primates analyzed has a monophyletic set of DQB-1 alleles. "Clades" include *Homo + Pan*, *Homo + Gorilla* and various combinations of *Pan*, *Homo*, *Gorilla*, and various tailed apes. This result obtains in spite of the observation that humans do not share alleles with gorillas or chimps, rather, the clusters represent related alleles, not shared alleles. The incongruent nature of the results reflects the hypothesis that these families of alleles have incongruent coalescence times. Doyle points out that such patterns are also found in flowering plant self-incompatibility loci and ADH loci of grasses (see Doyle, 1995, for literature).

Doyle defines an FFR as alleles within the same allele pool. Alleles within the same allele pool are potentially able to recombine, forming heterozygotes, and these heterozygotes define a field of recombination. Doyle presents a simple example for a single locus (Fig. 2.9). Of the eight individuals observed, there are five heterozygotes. The heterozygotes ab, bc, and cf are indicative of an allele pool composed of alleles a, b, c, and f. The heterozygotes ed and eg are indicative of an allele pool e, d, and g. Thus, there are two FFRs, one composed of individuals 1–5, the other composed of individuals 6–8. Of course, single alleles would rarely be expected to define an FFR for an entire species, so Doyle (1995) suggests a multilocus approach and presents his techniques of Multi-Locus or ML-FFR, which would be expected to yield larger FFRs and delineate species boundaries. Note that although the obvious application of FFR analysis is allozymes, it can be extended to loci identified via DNA sequencing. Informally, we can also see its application in morphology, where morphological intermediates are identified as belonging to the same gene pool and, thus, to the same lineage.

The *Genetic Distance Good and Wake (GenD$_{GW}$) Method* (Sites and Marshall, 2004) is a variant of the M/GD that can be used with multilocus allelic frequency data under the assumption that gene flow and genetic drift are in equilibrium. Widely distributed species might differ in gene frequencies due to local adaptation or drift but still maintain cohesion through gene flow. If so, then there should be a
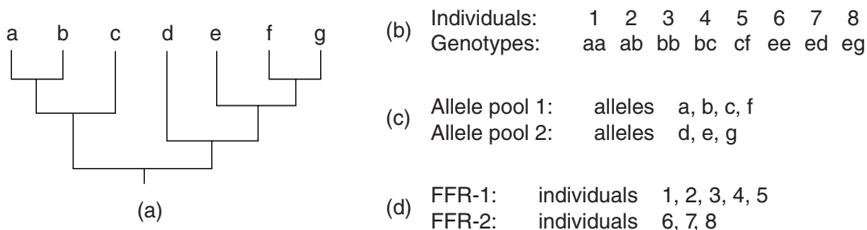
Figure 2.9. Doyle's (1995) one-locus example of defining the allele pool and field for recombination (FFR) using nuclear allele data. (a) An allele tree showing the relationships among alleles. (b) Observed individual genotypes of eight individual plants. (c) The gene pools based on the observed genotypes. (d) The recombination fields. Note that individuals belonging to field 1 can produce heterozygotes like individual 2 but individuals in field 2 cannot. From Doyle (1995), with permission from the American Society of Plant Taxonomists.

correlation between the genetic and geographic distances between populations. Good and Wake (1992) proposed a method for testing isolation by distance. The method consists of regressing a measure of genetic distance between all pairs of populations against geographic distance. The populations are identified a priori and can be based on geographic locality (e.g., within a basin) or taxonomically. Within-species regressions are expected to have slopes that intersect the origin while between-species regressions are expected to have slopes that do not intersect the origin (Fig. 2.10). This technique has been used for a number of salamander groups (e.g., Jackman and Wake, 1994; Tilley and Mahoney, 1996).

Another approach using distances was used by Highton (1998, 2000, and references therein) to infer reproductive isolation among *Plethodon* salamanders by comparing genetic distance between a priori samples to those genetic distances that
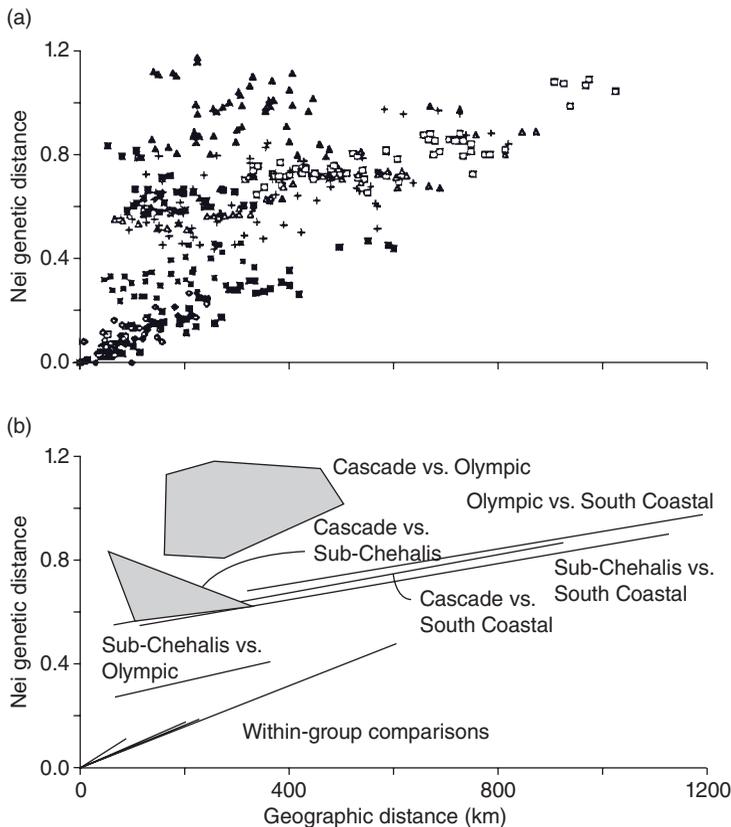
Figure 2.10. The genetic distance method of Good and Wake (1992). (a) A scatter plot of pair-wise comparisons of populations within and between four species of the salamander genus *Rhyacotriton* over geographic distance. Within species pair-wise distances among populations are solid squares and between species pair-wise distances are other symbols. (b) Regression of pair-wise distances over distance between populations within species intersect the origin while those between species do not. From Good and Wake (1992); copyright University of California Press, with permission.

commonly characterize other vertebrate species. Sites and Marshall (2004) term this approach the *Genetic Distance Highton (GenD$_H$)* approach. It assumes a molecular clock yielding a time dependent "emergence" of reproductive isolation. If conspecific, a histogram of pair-wise distances among populations should have a unimodal distribution (the null hypothesis), but if bimodal, then two species might be present in the analysis. If a bimodal distribution of sufficient magnitude is found, and if this can be correlated with morphological and coherent distributional data, then inferences regarding species boundaries may be strong (see Highton and Peabody 2000; for example, the bimodal peaks correspond with coherent distributional patterns that are indicative of a coherent range rather than scattered through the landscape).

Another distance measure variant is *Hybrid Zone Barrier Analysis (HZB)*, which is built around the assumptions that genetic drift and gene flow are at equilibrium and uses an isolation-by-distance model (Sites and Marshall, 2004). For example, Porter (1990) investigated hybrid zones between nominal species of admiral butterflies based on Wright's (1931, 1968–1978) hierarchical F-statistics under an island population model.

Puorto et al. (2001) present a different use of distance data, combining multivariate morphological analyses and genetic distances. This method, characterized by Sites and Marshall (2004) as the *Correlated Distance Matrix method (Coor-D)* is built on the assumption that independent lineages should be characterized by congruent patterns of genetic variation (mtDNA clusters) and morphological variation. Morphological variation is summarized by computing distances between specimens. This matrix is used as a dependent variable that can be evaluated for potential alternative causal factors using Mantel tests, including sex, geographic distances, and patristic distances of (in the case presented) mtDNA haplotypes. Given that factors such as sex and geographic distance between samples can be factored out, the question is: are patterns of mtDNA clusters congruent with patterns of morphological variation? The specific application sought to test the hypothesis that two species of *Bothrops* were present in eastern Brazil, as evidenced by two clusters of mtDNA haplotypes that show partial geographic overlap. However, the Mantel tests applied did not show any significant association between morphological variation and mtDNA variation, indicating that the samples were drawn from a single lineage.

Tree-Based Methods

Tree-based methods treat species as emergent hierarchical entities. Basically, species are recognized in reference to a specified phylogenetic hypothesis. Failure to find hierarchical relationships is evidence that gene flow is causing a breakdown of the potential emergence of a hierarchical pattern. Many of the methods use individuals as terminal units of analysis and rely on a criterion of exclusivity (sometimes called monophyly) to identify potential species.

Phylogenetic/Composite Tree-Based (PCT) methods as outlined by Brooks and McLennan (2002). Sites and Marshall (2004:208) rightly point out that the number of species recognized depends on the species concept held by the author of the hypothesis. If an evolutionary species concept is embraced, then there will be as many species lineages (exclusive groups) as emerge from the analysis (including

the unsampled ancestral species). If various forms of the phylogenetic species concept are employed, the result might be the same number of species or a greater number of species. For example, the PSC embraced by Wheeler and Platnick (2000) might call for new species with each fixation of an apomorphy, even in anagenetically evolving lineages. As mentioned earlier, much depends on the species concept, which is why concepts are important.

Closely related to PCT is what Sikes and Marshall term the *Wiens-Penkrot (WP)* method (Wiens and Penkrot, 2002). It is based on the proposition that gene flow, as evidenced by either morphology (population samples) or molecular data (individual haplotypes), will break up potential hierarchical patterns that might emerge if gene flow was not occurring. The species concept underlying the assumptions is the ESC. The absence of strong hierarchical signal (weakly supported alternative trees uncorrelated with biogeographic patterning) rejects the hypothesis that two or more species are present in the analysis.

Wiens and Penkrot (2002) present their method as a series of bifurcating decision trees based on the concordance of phylogenetic analysis, the locality of the samples examined, and the species assignments of the samples. The focal species is a series of populations of interest; the relevant question is: do populations from the same region group together exclusively in a phylogenetic analysis, thereby comprising a species? A reference species is a closely related species assumed to be an exclusive group for purposes of the analysis. Two possible outcomes of such an analysis are shown in Fig. 2.11.

If one is working with haplotype data, then the phylogenetic analysis is performed on individual haplotypes. If one is working with morphology, then entire local populations are the units of analysis. Decisions are made on the basis of the phylogenetic positions of the individuals (haplotypes) or populations (morphology) of the focal species on the resulting phylogenetic tree in reference to two or more species that are also analyzed. For example, if all populations or haplotypes of the analyzed species appear as an exclusive group in the resulting tree relative to the reference species, then the decision leads to corroboration of the focal species as a species, with the assumption that the lineage is acting independently. If the focal species is resolved into two exclusive groups, then two species are present. If the focal species does not resolve as an exclusive group but is intermingled with the reference species, then the focal species is not a species at all as it is not acting as an independent lineage relative to the reference species (or series of species). Wiens and Penkrot (2002) emphasize that robust results depend on the quality of the analysis. Weakly supported clades are of little use in the decision-making process. Strong inferences can only be made if highly corroborated results obtain; that is, high bootstrap, likelihood or a posteriori probabilities.

Wiens and Penkrot (2002) call attention to similarities in their approach and that of Brower (1999; CHA, see below) but emphasize that because sexually reproducing species are characterized by gene glow between populations, their approach is superior in providing a criterion for dividing the haplotype tree into species. W-P is also similar to Nested Clade Analysis (NCA, discussed below) but can be performed when sample sizes are low, as is frequently the case in real-world systematic studies.

Baum and Shaw (1995) approach delimiting species using coalescent theory (Hudson, 1990). This method is termed the *Genealogical Exclusivity Method (EXCL)*
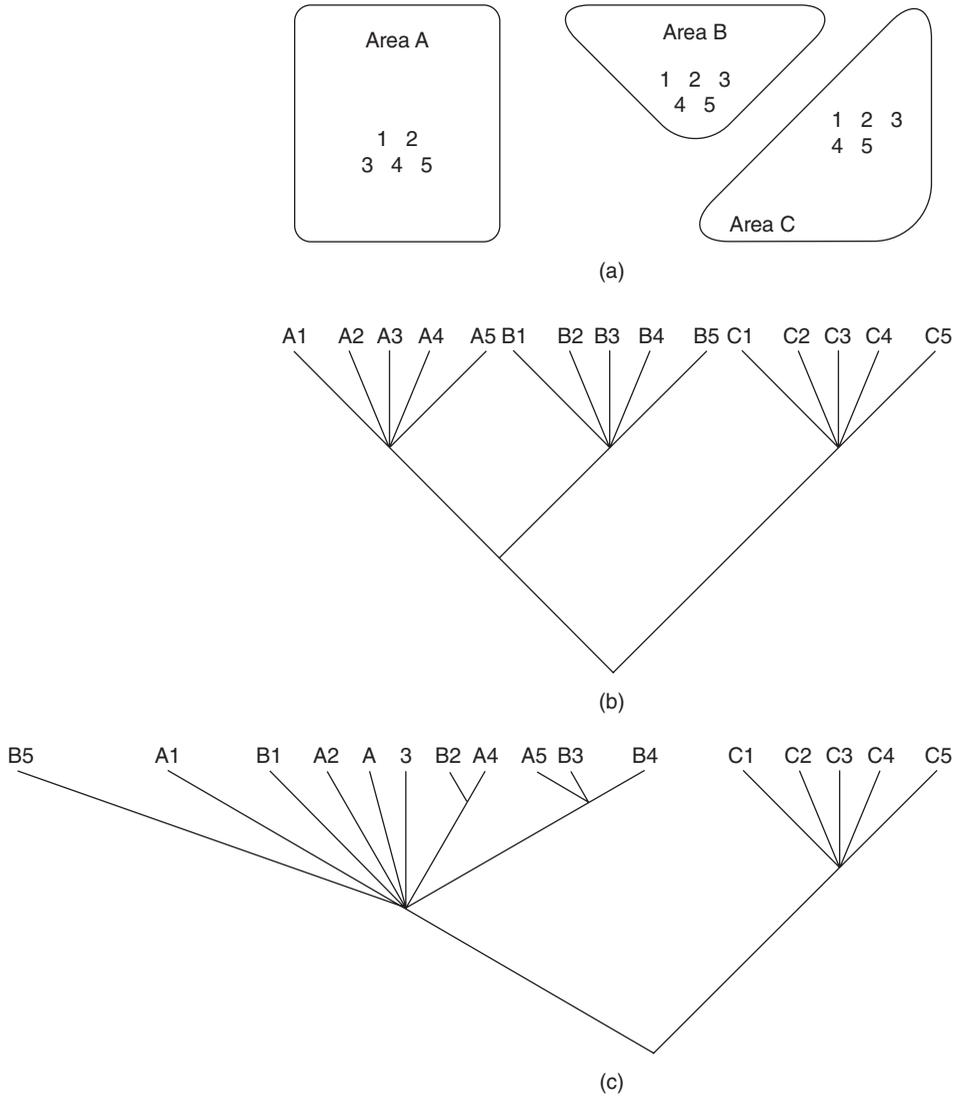
Figure 2.11. Hypothetical examples of outcomes applying the Wiens-Penkrot method. (a) Three geographic regions each with a set of five representative populations. (b) The haplotypes sort to exclusive groups on a strongly confirmed tree that is congruent with geography, providing evidence that species are involved. (c) The haplotypes do not form exclusive groups correlated with geography providing evidence that a single species is involved.

by Sites and Marshall (2004) and is based on suggestions by Avise and Ball (1990). In essence, species are delimited by the phylogenetic concordance of neutral, unlinked gene phylogenies. The underlying assumption is that lineages have been independent for a sufficient period of time for their gene phylogenies to coalesce, demonstrating that they are on the hierarchical side of the tokogenetic-phylogenetic interface. The method consists of separate phylogenetic analysis of unlinked genes. A strict consensus of calculated and consensus nodes are taken as species markers.

The species delimited appear exclusive in the resulting consensus tree. The exclusive groups mark coalescence.

Cladistic Haplotype Aggregation (CHA) (Brower, 1999) is a modified approach to PAA that departs from the original method in two respects. First, it is specifically applied to DNA sequence data; Brower recognized that such data can be analyzed in two ways. Second, it uses phylogenetic analysis as a vehicle to examine the distribution of populations on a rooted or unrooted tree as part of the decision-making process.

Brower (1999) recognized two forms of PAA character analysis applied to sequence data. PAA1 treats the entire sequence as the attribute and is similar to morphology in this respect. PAA2 treats each base pair as the attribute. CHA differs from the two forms of PAA in that a phylogenetic analysis is performed on all haplotypes using base pairs as attributes. If groups of haplotypes map on the tree to a priori populations, then species boundaries are drawn such that each species appears exclusive on the tree. That is, a diagnosable species would be one whose haplotypes are joined by a contiguous section of an unrooted tree. This may appear to create paraphyletic species, but the unrooted tree is not a phylogeny per se, but a "grouping diagram" that accepts or rejects a particular a priori hypothesis of species boundaries. (Another way of putting it is that the characters used are not polarized.) Rejection of a priori groups obtains if they do not map in this manner; then the hypothesis that two or more species are present is rejected.

Nested Clade Analysis (Templeton et al., 1995; Templeton, 2001, 2004). Nested Clade Analysis (NCA of Sites and Marshall, 2002) is built around statistical tests designed to test the null hypotheses that (1) all organisms are sampled from a single population and (2) separate lineages found through rejection of hypothesis (1) are genetically or ecologically interchangeable.

The first step is to produce a haplotype map and nest the haplotypes to produce a hypothesis of haplotype relationships. The null hypothesis that all organisms are sampled from a single population uses permutation tests to determine if haplotypes are distributed randomly on a haplotype tree relative to the geographic location of the samples. For example, the test might show that there is no association between the geographic position of haplotypes and their place on the tree produced by nesting haplotypes. In such a case, the null hypothesis is accepted and the inference is that all halpotypes are representatives of populations that are connected by gene flow. Alternatively, one might find that all haplotypes are clustered geographically and phylogenetically. In this case, the number of species recognized would equal the number of exclusive groups. Templeton (2001) discusses his methods in detail, and Templeton (2004) presents a decision tree to guide the investigator.

The second step is to fully implement Templeton's cohesion concept by testing whether the exclusive groups are genetically or ecologically exchangeable. Such tests call for considerable knowledge of the biology of the organisms. Templeton (2001) discusses some of the limited tests that have been performed. There are potential pitfalls. For example, many closely related species have similar to identical ecologies and lack pre- or post-mating isolating mechanisms. Exactly what constitutes a relevant genetic or ecological trait is questionable. However, the concept of genetic or ecological interchangeability may be an important criterion where it can be applied, especially in asexual organisms.

Potential uses for NCA are plentiful, especially when the investigator is not simply performing systematic revisions but is attempting to study speciation. One major drawback of NCA analysis in many systematic studies is that it requires dense sampling throughout the range of the potential distribution of the species (or species group). This drawback can be overcome by careful research design before the project begins. Wiens and Penkrot (2002) suggest a hierarchical approach, with an analysis using W-P protocols to ferret out the obvious species and a follow-up study designed a priori to implement NCA for those populations that might contain additional species not discovered during the W-P study.

CHAPTER SUMMARY

- Species concepts are kind concepts, but particular species are individuals.
- Species form lineages and are the largest tokogenetic arrays within which reproduction predominates.
- Of the many species concepts, the ESC provides the best concept for integrating systematics with allied disciplines.
- Patterns of phylogenetic relationship coupled with biogeographic analysis can yield insights into the processes of speciation.
- Estimates of the number of species in an area can affect other disciplines including community ecology and conservation biology.
- A number of empirical methods have been developed that assist the phylogeneticist in reaching decisions about the number of species in a sample of organisms.

3

SUPRASPECIFIC TAXA

In Chapter 2, we explored the nature of species. We concluded that species-in-nature were individuals and that one species concept, the Evolutionary Species Concept, was the best candidate concept for the natural kind "species." In this chapter, we will discuss the nature of groups of two or more species—supraspecific taxa. Literally, any named assemblage of two or more species comprises a supraspecific taxon. Most of the mathematical permutations of the possible array of supraspecific names are quite unacceptable to any taxonomist (the number of possible taxonomic combinations is vast: Felsenstein, 1978a). We can conclude that either all possible combinations of species are perfectly acceptable or some combinations are better than others. In this chapter, we will suggest that supraspecific taxa of a particular kind, monophyletic taxa *sensu* Hennig (1966), have objective reality and thus form the basis for a natural classification. Further, we will suggest that "monophyletic taxon" is a natural kind based on evolutionary principles. Finally, we will suggest that such taxa have an objective basis for existence apart from our ability to find them and that only such taxa are to be preferred over the astronomically high number of possible taxa that we could name. *One of the objectives of phylogenetic systematists is to attempt to discover monophyletic taxa and to either name them formally or make their presence known by other means.* Of course, as with all science, we are constrained to proposing hypotheses that particular monophyletic groups exist. So, although we can conceive of the kind "monophyletic group" as a natural kind, we must understand that our conjectures based on empirical evidence are hypotheses subject to confirmation or disconfirmation as new evidence is discovered.

CONCEPTS OF NATURALNESS AND SUPRASPECIFIC TAXA

Natural is considered by many systematists to be a loaded term. To claim that one taxon is natural leads to the conclusion that an alternative taxon is "unnatural." But this is exactly what we wish to achieve, in spite of the fact that some (e.g., Mayr, 1969) eschew the term because it is historically burdened. Indeed, there are three concepts of "natural" commonly encountered in the systematic literature.

Taxonomists generally define Aristotelian naturalness as that quality a taxon has when the things placed in the taxon agree in characters that embody the essence of the group (Crowson, 1970). The properties are both necessary and sufficient in that having the properties demonstrates that an entity belongs to the group and lacking the properties excludes an entity from the group. Phenetic naturalness is another concept. A taxon may be considered natural in the phenetic sense if all of the members of the group are more similar to each other (by some measure of similarity) than to any entity placed outside the group (Davis and Heywood, 1965; Crowson, 1970; Sneath and Sokal, 1973). Of course, one must agree about what constitutes a measure of similarity.

Phylogenetic naturalness is the concept we adopt. A supraspecific taxon may be considered natural if the members of the taxon include an ancestral species and all descendants of that ancestor. Monophyletic groups *sensu* Hennig (1966) are natural groups and paraphyletic groups as well as polyphyletic groups are unnatural. We shall see why in later sections.

Wiley (1981a:71–72) briefly outlines some history for these concepts. As Western science emerged during the eighteenth and nineteenth centuries, it became harder to determine exactly what comprised the "essence" of a taxon and, thus, the basis for a "natural system." Linneaus (1753) used reproductive morphology for plant classification, but workers such as Adamson (1763), de Jussieu (1789), and de Candolle (1813) moved toward plant classifications that encompassed diverse morphology. Such systems, based on shared similarity, were considered more natural because they did not depend on one's opinion as to what characters were "essential," only on agreements as to what plants shared more characters. Overlain on this move toward similarity were various systems of classification, some Linnean and some distinctly idiosyncratic. The rise of evolutionary thinking (Darwin, 1859) suggested that there might be a cause behind classification. As Mayr (1942) states:

> [T]he puzzle of the high degree of perfection of the natural system (was solved) in a manner that was as simple as it was satisfactory: The organisms of a "natural" systematic category agree with one another in so many characteristics because they are descendants of one common ancestor! The natural system became a "phylogenetic" system. The natural system is based on similarity, the phylogenetic system on the degree of relationship.

While this overlooks many of the controversies that raged during the nineteenth century, Mayr's statement may reflect what the winners thought as they converted to the evolutionary paradigm. It also had a salutary political benefit; one could join the winning team without abandoning one's favorite classification scheme by simply

switching the interpretation one placed on the classification. However, it took almost one hundred years from Darwin's (1859) statement that classification should reflect phylogeny to understanding the implications of this statement, which were realized fully with the ascendancy of the phylogenetic paradigm.

The Linnean system recognizes supraspecific taxa as a function of rank rather than biology. We will treat all Linnean taxa above the species level as "higher" taxa and restrict the term *supraspecific taxon* to a biological entity that contains two or more species. The issue comes up in discussing so-called monotypic taxa, which are Linnean constructs containing single species but not clades containing an ancestor and all of its descendants. This "higher Linnean taxon" may contain one to many species as a byproduct of ranking.

THE NATURAL TAXON

Natural taxa are those taxa that have a real existence in nature, being neither artificial nor manmade. Such a taxon exists independent of human perception and requires discovery. This concept carries specific connotations:

1. Natural taxa exist whether or not there are any taxonomists around to perceive and name them.
2. Because they exist in nature independent of our ability to perceive them, natural taxa require discovery; they cannot be invented.
3. Natural taxa originate via natural processes, and thus, any taxon that is natural must be composed of parts that make the whole consistent with the natural processes that caused their existence, as we currently understand these processes.

The idea that natural taxa are groupings of organisms that exist in nature is neither new nor novel. Aristotle asserted that all real things have a cause. Evolutionary taxonomists had similar views. For instance, Simpson (1961:55) stated: "The taxa of natural classifications must have some relationship … with groups of whole organisms really existing in nature." Crowson (1970:275) stated: "[A] perfectly natural classification of plants and animals might even be considered as objectively existing, thus requiring to be discovered rather than invented." Hennig (1966:77–83) characterized phylogenetic groups as groups that have the qualities of individuality and reality.

We may now turn our attention to the question of what exists in nature. We have to be careful, because what we think exists in nature is what we search for, and if we have the wrong concept or idea, we will be led astray. If we examine history, we can see that many things once thought to have existed in nature, in fact did not exist at all. (And certainly some things we have no idea exist are all around us.) In science, our process theories profoundly affect what we see and what we see affects our theories.

Consider disease theory. At one time, miasmas were considered to be the cause of cholera and the theory and the concept was central to forming health policies in England during the mid-nineteenth century. Miasmas were supposedly caused by air charged by an "epidemic influence" interacting with organic decomposition.

What smelled bad *was* bad. Those who held to this theory of disease held that miasmas existed in nature. Those who held to germ theory rejected miasmas and held that living organisms were the cause of cholera. Germ theory was vindicated because repeated experimentation and observation (by Robert Koch in 1884) established *Vibrio cholerae* as the disease-causing organism. We should be rightly suspicious of any claims that "thus and such" exist in nature, and we should look carefully at the evidence for such claims.

Given current scientific theories, we may ask: what exists in nature that we might study? There are those individual things that we can discover, describe, and explain in reference to our perceptions of the natural laws, and processes we posit shape the world around us. For example, physicists posit that various atoms and molecules exist in nature and that much of their behavior can be explained because their interactions can be understood (albeit incompletely) by certain physical laws, and that the basic individuals (the atoms) can be organized into kinds of elements that make these interactions predictable. The Periodic Table serves to organize this knowledge. The individual atoms are entities, and the kinds found in the Periodic Table are defined by necessary and sufficient properties that serve both to identify an atom as to its kind membership and to predict some of that atom's chemical properties in interactions with other atoms. Yes, it's true: atoms are themselves composed of even smaller individuals. But atoms are not the passive sum of their subatomic particles; atoms have emergent properties, and these properties are predicted by "higher-level" theories that cover the behavior of individual atoms. At this level of theory, the subatomic particles are "noise." Yet at the next lower hierarchical level, these same subatomic particles are the relevant individuals to be explained and form the boundary conditions for the behavior of the atoms themselves.

Biology is characterized by many levels of complexity and thus with an array of individual entities that may function at one to many levels. These levels of organization have been said to form a scalar hierarchy (Salthe, 1985) in which higher levels of organization provide the boundary conditions for lower levels of organization and each higher level is characterized by emergent properties not entirely predictable by, but related to, the lower levels. We believe that at least some of these levels are natural because the properties of individuals we observe match the properties predicted by theory. For example, a particular population of flies may be said to be a Mendelian population because the population of flies exhibits those properties of a Mendelian population as predicted by population genetic theory. Or, a group of species may be said to be a monophyletic group because the relationship properties existing between these species are those properties predicted for the natural kind "monophyletic group," a kind that emerges from the general theory of descent with modification and speciation. (To wit: if species speciate, then we can reasonably deduce that evolution will result in monophyletic groups of species.) Levels in the scalar hierarchy that are not typically of direct interest to systematists include cells (although these may provide the foundation for entities or the characters of entities used by systematists) and ecological communities. Levels systematists are typically interested in include individual organisms, populations, species, and clades. Much of our interest in individuals and populations lies in figuring out how they relate as parts of species and clades, the major levels of interest to phylogenetic systematists.

MONOPHYLY, PARAPHYLY AND POLYPHYLY

Hennig equated "monophyletic group" with "natural taxon" and considered both species and monophyletic groups to be individuals with objective reality in nature (for example: Hennig, 1966:146). He provided two definitions of the mono-phyletic group, one referencing ancestral species and the other referencing known taxa.

1. A group is monophyletic "if it can be shown that all species (or individuals) included in it actually descended from a single stem species, but also that no species derived from this stem species are allocated outside the group in question" (Hennig, 1966:73).
2. "A monophyletic group is a group of species in which every species is more closely related to every other species than to any species that is classified outside the group" (Hennig, 1966:73).

The second definition does not escape reference to common ancestors because the term "related" is defined in strictly genealogical terms, with reference to common ancestors (Hennig, 1966:74): "A species x is more closely related to another species y than to a third species z if, and only if, it has at least one stem species in common with y that is not also a stem species of z." The same concept holds for higher taxa (Fig. 3.1a). Miidae is more closely related to Xidae than to Yidae if and only if Miidae and Xidae share a stem species (ancestral species) that is not shared by Yidae. Hennig (1966:71) made it clear that a monophyletic group must include the ancestor and all descendants of the ancestor, and we support his contention and not the differing views of Tuomikoski (1967) and Ashlock (1971).

A subtle but important point can be made about Hennig's concept of monophyly. It is a relative concept in the sense that one always makes claim about the mono-phyly of a group relative to another group. This has relevance when we consider the
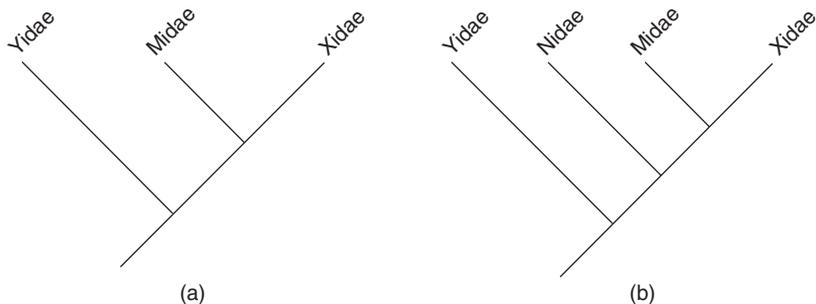

Figure 3.1. Relationship and relative relationships. (a) Xidae and Midae are more closely related to each other than either is to Yidae. (b) Exactly the same relative relationships holds for Midae and Xidae *relative to* Yidae in spite of the addition of Nidae.

effect of adding newly discovered taxa. Consider Fig. 3.1. In the first case, we state that Midae is more closely related to Xidae than to Yidae (Fig. 3.1a). Now consider Fig. 3.1b. We have discovered that another group, Nidae, is actually the sister group to Midae+Xidae. It is still the case that Midae and Xidae are more closely related to each other than either is to Yidae, but it is also true that both are more closely related to each other than to Nidae. Formerly, (Fig. 3.1a) it was hypothesized that Yidae was the sister group. Now, it appears that Nidae is the sister group (Fig. 3.1b); the status of the monophyly of Midae+Xidae is the same, but the reference point for our claim of monophyly has shifted.

Hennig (1966:73) viewed his definitions of monophyly as critical restatements (clarifications) of less restrictive definitions that had come into vogue since Haeckel (1866) coined the term. Although many authors (including Wiley, 1981a) also thought that Hennig (1966) defined monophyletic groups in terms of characters, this is not strictly true. Hennig saw characters as an epistemological means toward the end of discovering monophyletic groups, but these monophyletic groups had real ontological status. Because we cannot typically observe stem species (common ancestors), it is obvious that we need some empirical means to infer stem species. Hennig does not describe characters until 17 pages after his statements about mono-phyly (Hennig, 1966:90: "That a common stem form is shared by a group of species [a condition for a 'monophyletic group'] can be proven only by means of synapo-morphous characters, not by symplesiomorphous characters").

Some 56 pages after this, Hennig discusses the distinctions between monophy-letic, paraphyletic, and polyphyletic groups. The concept of paraphyly represents a new concept to the mix. Prior to Hennig (6, 1966), systematists generally recognized two kinds of groups relative to phylogeny, monophyletic groups and polyphyletic groups, and only a few workers distinguished between monophyly and what we now recognize as paraphyly (see next section). Certainly there was considerable debate then as to whether members of a monophyletic group were descended from higher taxa, species, or even an original pair of organisms à la Adam and Eve, but just about everyone interested in evolution agreed that monophyletic taxa were desired and polyphyletic taxa were to be avoided. However, the nature of monophyly (and whether it also encompassed what Hennig distinguished as paraphyly) was not settled. For example, Mayr (1942) held that monophyletic groups were descended from species while Simpson (1944) held that they were descended from species or from taxa of equal or lower rank ("minimum monophyly"). Hennig (1966) recog-nized that some groups that were considered monophyletic were of a different quality than what he considered true monophyletic groups and that the distinction had gone unrecognized because concepts of similarity were not being parsed cor-rectly. Hennig (1966:146) argued that when one uses similarity to group, three pos-sible outcomes obtain.

1. Monophyly obtains if the similarity used to group is synapomorphic similarity.
2. Paraphyly obtains if the similarity to group is symplesiomorphic similarity.
3. Polyphyly obtains if the similarity used to group is homoplastic similarity.

Hennig then states that paraphyly and polyphyly are of a similar nature (Hennig, 1966:146):

> The paraphyletic groups (as much as the polyphyletic groups) are distinguished from the monophyletic ones essentially by the fact that they have no independent history and thus possess neither reality nor individuality.

Although we cover it in more detail in Chapter 4, also notice that Hennig (1966) has split the concept of homology into two concepts that are relative. Some homologs are apomorphies while others are plesiomorphic when referencing a particular group of organisms.

HENNIG'S CONCEPTS PLACED IN HISTORY

Two papers by German-speaking authors analyze the development of Hennig's ideas as seen through his early papers. Richter and Meier (1994) trace the refinement of the term *monophyletic* and coining of the term *paraphyletic* through a series of works. Willmann (2003) traced the fate of the term *monophyly* in the German literature, and his analysis yields some interesting insights. He concluded that Hennig's concept closely matches Naef's (1919), who also distinguished between monophyletic and paraphyletic groups. Hennig's concept of relationship also closely matched Zimmermann's (1937, 1943). Further, Willmann's (2003) analysis considered the origin (in the German community, at least) of the concept of "minimum monophyly" that later appears in the formulations of Simpson (1944, 1961) and Mayr (1969, 1974) and which was pervasive earlier in the twentieth century. According to Willmann's analysis, Handlisch was led to propose a broader definition of monophyly that would include the origin of groups from higher taxa because the origin of species was not restricted to a single individual organism or pair of organisms. This reasoning is interesting. If Willmann's analysis is correct, then it points to the danger of treating species as ontologically identical to monophyletic groups. In other words, apparently Handlisch perceived no difference between species as taxa and groups-of-species-as-taxa. If species could originate from pairs, populations, or species, then why can they not originate from genera or orders? Willmann (2003:460) concluded that Handlisch's "vague definition of monophyly was later shared by almost all classificationists and assisted in blocking the development towards a true phylogenetic systematics."

Mayr and Ashlock's (1991) statement that from "Haeckel to 1950 a taxon was called monophyletic if it was derived from a single ancestral taxon" seems not to be indicative of the rich tradition of, at least, the German language literature. It is not even particularly indicative of Mayr's (1942) own statement that monophyletic groups were descended from single species (not "taxa" in general). Hennig (1966) was not satisfied with this restriction; he considered it to be incomplete because the concept did not specifically state that all descendants were included in the group. Further, it did not state that the ancestral species itself is a member of the group (inherent in Hennig's concept and a concept that can be traced back to earlier German systematists, see Willmann, 2003). Hennig's unique contribution to this issue was not that he originated the idea of strict monophyly, or distinguished between monophyly and paraphyly, but that he insisted that there be a correspondence between taxonomic practice and evolutionary thinking (Willmann, 2003). We

shall see in a later section that Hennig's reasoning was sound regardless of "traditional definitions."

NATURAL HIGHER TAXA AS MONOPHYLETIC GROUPS
SENSU HENNIG (1966)

Clades (monophyletic groups *sensu* Hennig) are one of the levels of biological organization predicted by our most general theories of evolution. They are one of the results of speciation. Ancestral species split for any variety of reasons, and the results are clades. Clades, like lineages, are predicted from theory. In fact, if there were no clades, we would have to completely revise our theories of macroevolution. Taxon names are different. To name a clade is a human activity. It is to hypothesize that a clade exists in nature and to acknowledge the need for communicating its existence to other scientists through the name given to the clade. There is no particular constraint on naming anything we wish. Taxa are simply named groups of organisms. Some of these names correspond to clades, but many do not and the qualities of the majority of taxa now recognized are unknown. But in phylogenetics, of the billions of possible taxonomic groupings and thus perhaps potential names, only those taxon names hypothesized to name lineages or clades are candidates for names of hypothesized natural taxa. Lineages, evolutionary species, receive binominals while clades receive uninominals. Both Patterson (1978) and Ghiselin (1980) were correct in terming natural higher taxa "individuals," as discussed by Coleman and Wiley (2001). However, the five points Wiley (1981a,b)originally raised are still pertinent, and we shall add a sixth.

1. There is no ongoing process that gives a natural higher taxon cohesion other than a common history of speciation.
2. Thus, natural supraspecific taxa must be the products of a history of speciation. That is, there is no origin of natural higher taxa except through the origin of species and species are the largest biological entities that undergo distinctive evolutionary processes.
3. We may conclude that genealogical lineage splitting and other speciation processes (even speciation via hybridization) are both necessary and sufficient to explain the origins of candidate natural higher taxa. Those higher taxa that do not accurately document these necessary and sufficient conditions cannot be natural taxa.
4. A natural higher taxon cannot overlap another at the same level in the hierarchy. That is, they cannot both contain the same species unless one includes the other in a part–whole relationship. It follows that the diagnosis of two supraspecific taxa cannot overlap. A homologous character state used to diagnose one taxon cannot be used to diagnose another. For example, we cannot diagnose a group composed of tunas and pufferfishes based on the homologous character state of having pectoral fins because pectoral fins have already been used to diagnose a larger group, the jawed vertebrates (Fig. 3.2). Pectoral fins are a property of the ancestor of all jawed vertebrates, not just the ancestor
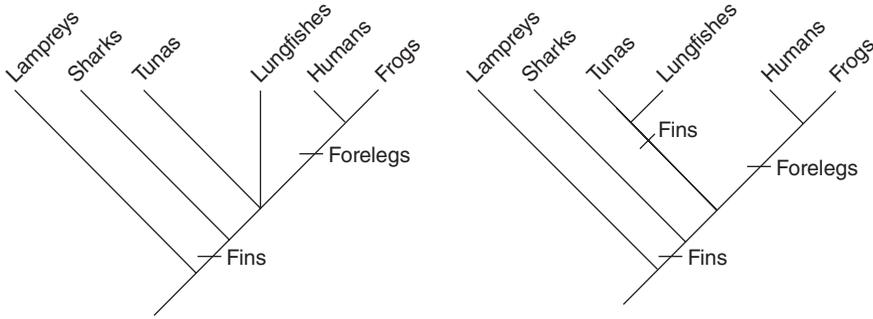
Figure 3.2. Homologous characters on trees. (a) Pectoral fins are a synapomorphy of jawed vertebrates while forelegs (modified pectoral fins) are a synapomorphy of tetrapods. (b) Using the sharing of pectoral fins shown by lungfishes and tunas to group them has the effect of using a character state twice in an analysis when the state only evolved once.

of tunas and puffers. As a property of the ancestor of jawed vertebrates, it is taken as an evolutionary innovation that evolved sometime after the origin of the lineage leading to lampreys and the origin of the common ancestor of sharks and other jawed vertebrates (exactly where cannot be discerned from this particular tree). Of course, if a character state turns out to be homoplasious (the same property evolved two or more times), then they are not homologous and can be used to diagnose different groups because they are, in the phylogenetic sense, different homologies. The properties of natural higher taxa include the evolutionary innovations of their first member (the ancestral species). This property is an evolutionary innovation, a homology, and a synapomorphy that once characterized the ancestral species as an autapomorphy. We shall discuss this quality of higher natural taxa in Chapter 5.

5. Although individuals cannot be defined by necessary and sufficient characters (Ghiselin, 1974; Wiley, 1981a), hypotheses that a particular taxon is a natural higher taxon should be justified because it is an empirical assertion of relationship. Justification comes in the form of one or more character states relevant to their hypothesized origins (synapomorphies), requiring the investigator to give evidence as to why the group should be recognized compared to numerous alternative groupings that are thought to be unnatural. However, it should be recognized that while such diagnoses might be sufficient in many cases, they are never necessary.

6. Taxa proposed to be natural higher taxa must be logically consistent with the proposed phylogeny of the group classified (Simpson, 1961; Hull, 1964) because it is a general principle of science that summaries be consistent with what they purport to summarize.

LOGICAL CONSISTENCY: THE HALLMARK OF PROPOSED NATURAL CLASSIFICATIONS

Empirically, it is the quality of logical consistency that separates hypotheses that higher taxa are natural from a large number of alternatives. Consistency is a basic

and rather powerful criterion that is so ingrained in the scientific enterprise that we are apt to forget its demands. Hull (1964) analyzed Simpson's (1961) claim that classifications should be consistent with phylogeny without mirroring phylogeny. The claim, as we shall see below, was valid. However, Hull noted that many of the groups Simpson advocated did not fulfill Simpson's own consistency criterion. As it turns out, those groups that fail the criterion are the paraphyletic groups that are still common in textbook classifications, paraphyletic groups such as Reptilia and Pongidae. Only the monophyletic group *sensu* Hennig (1966) has the quality of logical consistency relative to phylogeny. Hull's (1964) conclusions lay dormant until Wiley (1981b) called attention to them when discussing claims made about paraphyletic groups. Hull (1964:10) summarized the relationship between logical consistency, phylogeny, and classification in the following manner:

1. Of consistency, phylogeny, and classification, only phylogeny is of an empirical nature (i.e., based on data), and it is a hypothesis, not a fact.
2. Classification as now practiced portrays phylogeny by inclusion or exclusion and is a constantly diverging system because all taxa ranked at a particular level (by whatever convention) must be mutually exclusive and because two taxa once separated can never be classified together again at a lower level.
3. "[N]o implications validly drawn from a classification can contradict the classifier's views concerning phylogeny" if consistency is to be maintained between a classification and a phylogeny.
4. "Because the relationship from phylogeny to classification and that from classification back to phylogeny are each one-many relationships, few inferences specific enough to contradict phylogenetic views can be validly drawn from classification."

Hull (1964:10–11) then stated two criteria:

1. "Within very broad limits, a classification is consistent (with a phylogeny) if at least one of the possible phylogenies implied by it is the original phylogeny from which it was constructed."
2. "A classification is inconsistent if and only if all implied phylogenies conflict with the original phylogeny."

Or, as Wiley (1981b:347) stated: "Consistency obtains between a classification and a phylogeny when deductions validly drawn from the classification do not contradict any deductions validly drawn from the phylogeny." While Hull (1964) was mainly concerned with grouping, Wiley (1981b:348) was also concerned with character evolution.

1. A classification of organisms is consistent with character evolution if at least one character state diagnoses each grouping and no false deductions concerning the distribution of other character states can be drawn from the classification.

2. A classification of organisms is inconsistent with character evolution if and only if one or more diagnostic characters lead to a false deduction concerning character evolution.

Wiley (1981b) suggested that phylogenies and classifications could be directly compared if the grouping claims of the classification were converted into a classification tree and then compared to the original phylogenetic hypothesis. If the classification tree was identical in topology to the phylogeny, then it was logically consistent with, and fully informative of, the phylogeny (Fig. 3.3a, b). If the classification tree was different from the phylogeny, then two outcomes were possible. The classification tree might not fully reflect the phylogeny, but it might be consistent nevertheless, because it can be decomposed into a number of derivative classification trees, at least one of which has the same topology as the phylogeny (Hull's first criterion). An example is shown in Fig, 3.4. The classification only partly reflects the phylogeny (compare Fig. 3.3a and 3.4a). The classification tree (Fig. 3.4b) contains a polytomy because there are three clades classified at the same level within Miidae (here shown as three subfamilies). However, the classification is logically consistent with three more resolved classification trees (Fig. 3.4c, d, e). One of these, Figure 3.4d, has the same topology as the phylogeny (Fig. 3.3a). Thus, the classification is consistent with, but not fully informative about, the phylogeny.

The second outcome is shown in Figure 3.5. We have drawn the phylogeny with branch lengths (Fig. 3.5a) that imply that genera P and Q are much different from the other four genera, who share many plesiomorphic similarities. In this case, the classification (Fig. 3.5b) contains only one possible classification tree because it is dichotomous and the topology of the classification tree (Fig. 3.5c) *is not* identical with the phylogeny (Fig. 3.5a). Therefore, this classification is logically inconsistent with the phylogeny and thus *misinformative*.

Wiley (1981b) asserted that a classification containing only monophyletic groups were always logically consistent with the phylogeny and that classifications contain-
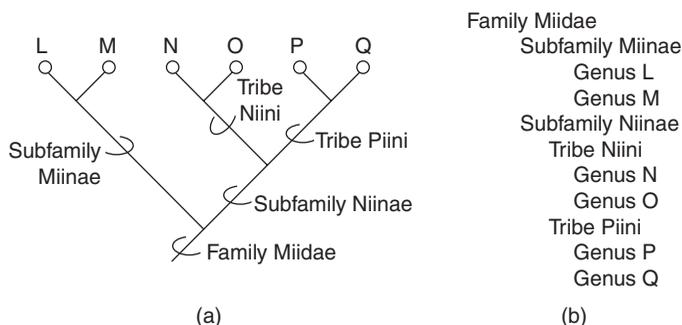
Figure 3.3. Classification and logical consistency I. (a) A tree. (b) A classification. Because the classification exactly reflects the tree, the classification is both logically consistent with and fully informative of the tree. For purposes of discussion, the tree is taken as "true" (that is, agreed to be the best hypothesis by anyone discussing logical consistency).
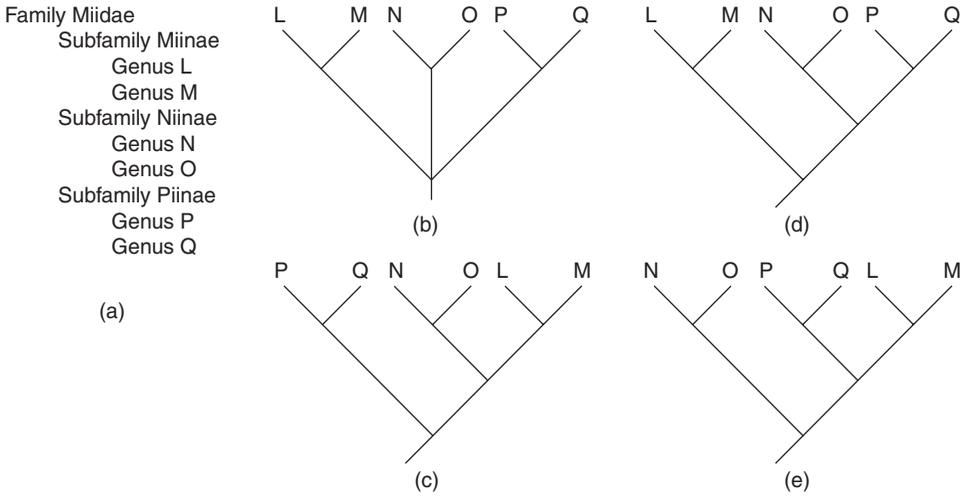
Figure 3.4. Classification and logical consistency II. (a) A classification. (b) The classification in tree form (reflecting relative subordination of the classification). (c–e) Three dichotomous resolutions of the trichotomy shown in the classification. (d) Has the same topology as the tree shown in Fig. 3.3a. The classification is logically consistent with the tree, but it is not fully informative of the tree.
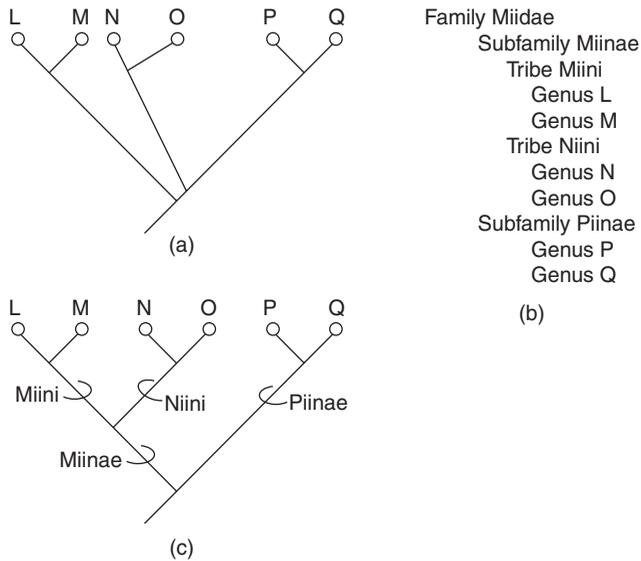
Figure 3.5. Classification and logical consistency III. (a) The tree as a phylogram, implying that LMN and O are more similar to each other than N and O are to P and Q. (b) A classification grouping LMN and O into a paraphyletic group, Miidae, based on plesiomorphic similarity. (c) The classification in tree form. Note that the classification is logically inconsistent with the original tree, and this misinforms the community.

ing even a single paraphyletic group were always logically inconsistent with the phylogeny, even when they contained tens to thousands of alternative topologies. The classic example is the current classification of tetrapod vertebrates shown in most introductory textbooks in the United States compared with two phylogenetic classifications containing only monophyletic groups (Fig. 3.6).

When evaluating a classification relative to a phylogeny, the first step is to convert the classification into a classification tree, as shown to the right of the classifications in Fig. 3.6. The second step is to determine if the classification tree contains any implicit alternatives that are internally consistent. Such alternatives are always present if the classification tree contains polytomies. However, if the classification tree is strictly dichotomous, there are no alternatives. The traditional classification, which contains Reptilia, has a polytomy between the four classes of vertebrates (Fig. 3.6a, b) and another one among the three groups of "Reptiles." As originally published by Felsenstein (1978a), every four-tomy is internally consistent with 26 possible derivative trees (including the original) in the absence of naming ancestors; and each trichotomy is internally consistent with three dichotomies plus the polytomy, yielding four possible classification trees. The total is multiplicative. Thus, there are 103 total classification trees that are logically consistent with the original classification.

Because the phylogeny of tetrapod vertebrates is strictly dichotomous (Fig. 3.6d), we can compare it to the 45 strictly dichotomous trees that emerge from the total of 103. If only one of these 45 classification trees has the same topology as the phylogeny, then we can claim that the traditional classification is logically consistent with the phylogeny. Unfortunately (for traditional taxonomy), none of the 45 classification trees have the same topology as the phylogeny. Thus, the traditional classification is logically inconsistent with the phylogeny. Why is it inconsistent? Because it contains the paraphyletic group Reptilia.

Consider the fully ranked phylogenetic classification (Fig. 3.6c). If we convert it into a classification tree (Fig. 3.6d), we observe that it is strictly dichotomous and that its topology is identical to the phylogeny (which is also Fig. 3.6d). It is both consistent with, and fully informative of, the phylogeny. Now consider the minimally ranked classification (Fig. 3.6e). It presents the reader with a polytomy of six branches, reflected in the classification tree (Fig. 3.6f). Its only knowledge claim is that all tetrapods are related. There are 945 possible dichotomous classification trees that are logically consistent with this classification. Only one of these has the same topology as the original phylogeny. However, it only takes one (Hull, 1964), so this classification is also logically consistent with the phylogeny.

We can form two conclusions from this demonstration given that logical consistency is a basic criterion of all science. No summary should be permitted that is logically inconsistent with what it attempts to summarize. Logical consistency is, as Hull (1964) concluded, a relatively weak criterion. Classifications that contain almost zero information about the phylogeny of a group may nevertheless be logically consistent with the phylogeny (e.g., Fig. 3.6e). Second, all classifications that contain paraphyletic groups are logically inconsistent with the phylogeny they are supposedly based upon. Thus, whatever their perceived benefits, they cannot be candidates for being natural classifications because they are at odds with the very empirical conclusions they claim to reflect.

Current Classification:
Class Amphibia or Lissamphibia
Class Reptilia
    Subclass Anapsida
    Subclass Diapsida
        Order Lepidosauria
        Order Crocodilia
Class Aves
Class Mammalia

(a)

(b)
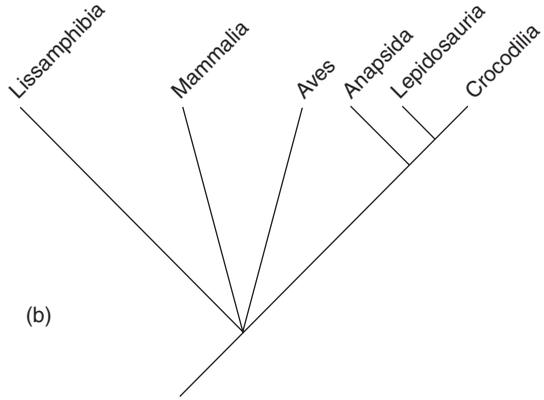
Phylogenetic Classification (fully ranked)

Class Tetrapoda
    Subclass Lissamphibia
    Subclass Amniota
        Infraclass Mammalia
        Infraclass Reptilomorpha
          Superorder Anapsida
          Superorder Diapsida
            Order Lepidosauria
            Order Archosauria
              Suborder Crocodilia
              Suborder Aves

(c)

(d)
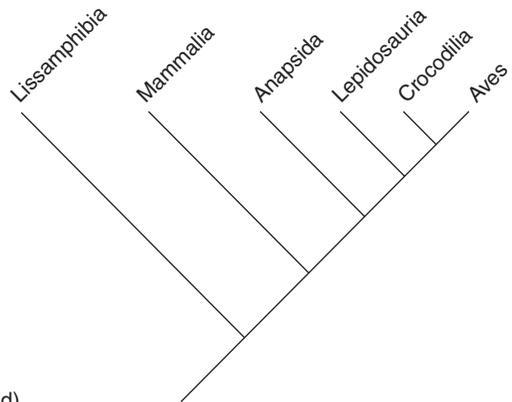
Phylogenetic Classification (minimally ranked)
Class Tetrapoda
    Subclass Lissamphibia
    Subclass Mammalia
    Subclass Anapsida
    Subclass Lepidosauria
    Subclass Crocodilia
    Subclass Aves

(e)

(f)
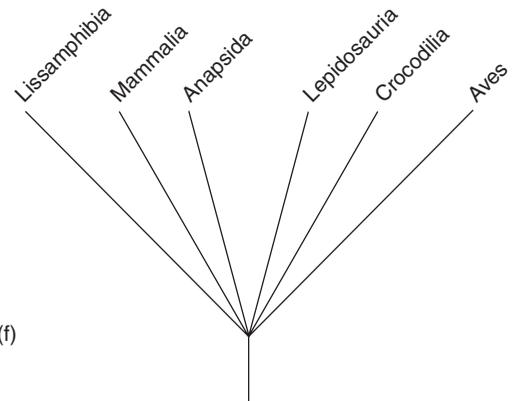
Figure 3.6. Classification and logical consistency IV. (a–b) The classification and classification tree used in many textbooks to classify tetrapod vertebrates. (c–d) A phylogenetic classification and classification tree as shown by current evidence (synapomorphies). (e–f) Totally unresolved classification and classification tree. Note that (c–d) and (e–f) are both logically consistent with the currently accepted tetrapod phylogeny but that (a–b) are not.

PARAPHYLETIC GROUPS MISREPRESENT CHARACTER EVOLUTION

Paraphyletic groups also lead to spurious representations of hypothesized homologous characters. Consider the homologies' "presence of mandibular fenestrae and the antorbital fenestrae." The presence of these fenestrae are synapomorphic for all archosaurs (Gauthier et al., 1988). Among living organisms, they are found in both crocodiles and birds. If we associate the characters with the traditional classification tree, however, we see that each homology appears twice, once as a homology of crocodiles and once as a homology of birds (Fig. 3.7). The inference is clear: the homologies are interpreted as convergent homoplasies. Thus, classifications that contain paraphyletic groups are misleading about both characters and phylogeny. Unresolved classifications containing only clades also have the potential to misrepresent character evolution, so special provisions in terms of diagnoses must be in place to prevent this misrepresentation, as discussed in Chapter 8.

Classifications can be consistent with, or inconsistent with, phylogeny only if they are meant to reflect, in some manner, phylogeny. There are many kinds of biological classifications that are not meant to reflect phylogeny and thus are not relevant to the consistency argument. Lions and pitcher plants are secondary consumers, and placing them together in a classification of ecological trophic groups may be useful and needed. Such a classification is neutral to the question of phylogenetic relationships and evolutionary descent because its knowledge claims lie in another direction.

Most natural higher taxa are composed of two or more species, and all are the result of past speciation. Some may be composed of a single known species (a monotypic family or order, for instance), but these are a byproduct of the needs of Linnean taxonomy and are really just species lineages that are "forced" to have higher taxon names in order to place them within the context of a hierarchy. Such would be the case in any classification that uses subordination and ranks to express
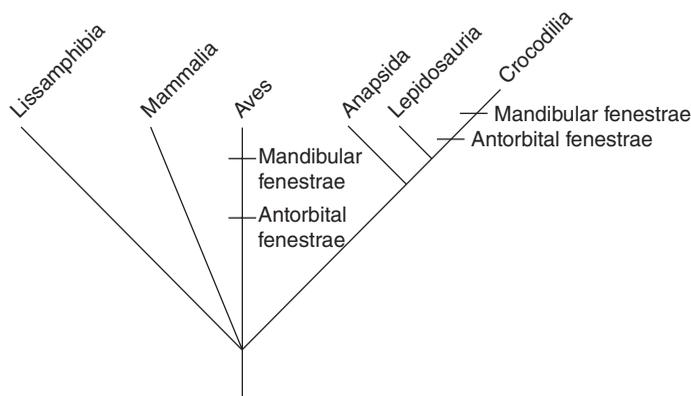

Figure 3.7. Paraphyletic groups misrepresent homologies as homoplasies. Two synapomorphies of the clade Archosauria mapped on the classification tree that maintains "Reptilia" as a group while excluding birds (Aves). Note that the mapping implies that two archosaur synapomorphies have evolved independently.

relationships. The problem is not that a monotypic family contains only a single species, but the fact that we believe that Linnean categories have some function other than to serve as relative subordination devices. Of course, one might choose not to use the Linnean System, and this topic is taken up in Chapter 8.

PARAPHYLY AND POLYPHYLY: TWO FORMS OF NONMONOPHYLY

Hull's Criterion of Logical Consistency relegates paraphyletic groups to the same status as polyphyletic groups; both kinds of groups are illogical relative to a given phylogeny. However, different authors have defined paraphyly in different ways. Hennig defined a paraphyletic group in both genealogical and methodological terms:

1. A group of species that has no ancestor in common only to themselves and thus no point of origin in time only to themselves in the true course of phylogeny.
2. A group based on symplesiomorphous characteristics.

Hennig's genealogical definitions of paraphyly and polyphyly are similar to later definitions (e.g., Ashlock, 1971; Farris, 1974; but not Nelson, 1971b). Hennig's genealogical and methodological definitions can be best shown by considering our tree of jawed vertebrates (Fig. 3.2) relative to a classification and set of diagnoses. If tunas and lungfishes are classified together, then the two groups do not have an ancestor in common only with each other because that ancestor is also ancestral to frogs and humans. If tunas and lungfishes are grouped together on the plesiomorphic character "pectoral fins," then the taxon "uberPisces" is paraphyletic. For example, the following classification and characters indicates the groupings and their group diagnoses:

Jawed Vertebrates (diagnosis: jaws, pectoral fins)
Sharks (diagnosis: males with pelvic fins modified as claspers)
uberPisces (diagnosis: pectoral fins)
Tetrapods (diagnosis: forelegs)

Hennig (1975) considered paraphyletic groups to be one of two kinds of nonmonophyletic groups, and he considered justifying such groups on the basis of symplesiomorphies a category mistake. As we have stated earlier, grouping by plesiomorphy is using single evolutionary innovations twice because each has been previously used to diagnose one or more taxa that include jawed vertebrates. This is the case with the characters jaws and pectoral fins: they have been used to diagnose a group that includes the clade composed of gnathostomes. Use of pectoral fins to diagnose the "uberPisces" (lungfishes and tunas) is the use of a single evolutionary innovation inappropriately in two places.

All authors agree that polyphyly is undesirable, but they differ in details. Hennig (1966) defined polyphyly in two ways:

1. A polyphyletic group is one in which the most recent common ancestor is not included in the group.

2. A polyphyletic group is a group based on convergent (nonhomologous) similarities.

Hennig's genealogical definition is similar to those of Mayr (1969), Ashlock (1971), and Farris (1974), but again differs from Nelson's (1971b). Nelson defined paraphyletic groups as those lacking one species or monophyletic group and defined polyphyletic groups as those lacking two species or monophyletic groups. Nelson's definitions are not in general use today because vertebrate zoologists recognize groups such as Reptilia as paraphyletic (lacking Aves and Mammalia) rather than polyphyletic.

Hennig (1975) considered the distinction between paraphyletic and polyphyletic groups to be useful descriptive terms to characterize the groupings of other authors. If an investigator justified a taxon with plesiomorphies, then the taxon is suspected to be paraphyletic, if with homoplasies (nonhomologies, convergences, or parallelisms), then the taxon is suspected to be polyphyletic. Such distinctions may result in identically circumscribed taxa being either polyphyletic or paraphyletic. Thus, Hennig's tree diagram illustrating paraphyly and polyphyly (Hennig, 1966, Fig. 45, p. 148) will appear to circumscribe the same kind of group (Fig. 3.8) unless we understand that he was referencing similarity, not tree topology, as the criterion for distinguishing between the two kinds of nonmonophyly.
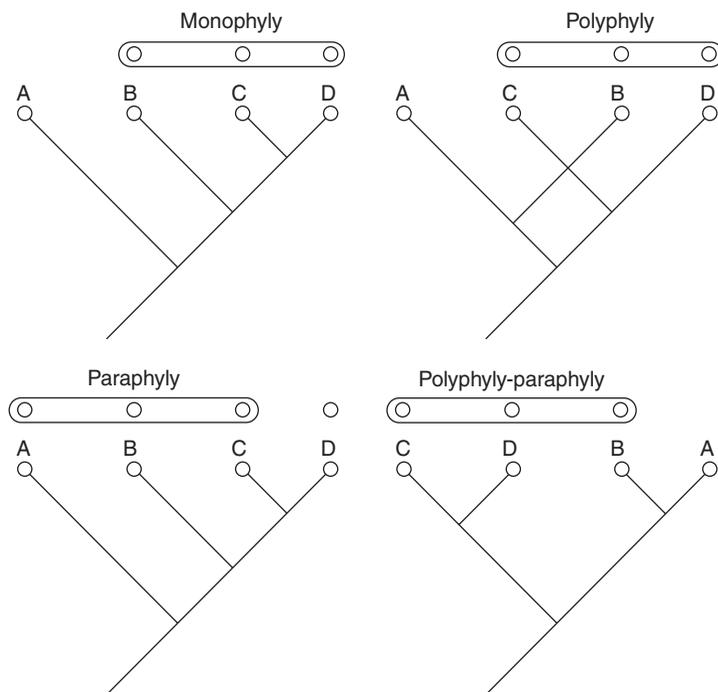

Figure 3.8. Hennig's (1966) concepts of grouping. (a–c) Trees from Hennig (1966:148). (d) Rotation of nodes on tree 3.8b results in paraphyly rather than polyphyly.

Hennig's distinctions between paraphyly and polyphyly based on the nature of similarity have the distinct advantage of basing the difference on empirical data. Distinguishing between the concepts using tree topologies has proven more difficult. For example, paraphyletic groups, by inference, include the stem species and part, but not all the descendants of that species. But if we knew that ancestor or could draw lines on trees, we could include an ancestor in just about any grouping we wished. This drawing lines about hypothetical ancestors on trees is one of the ways Ashlock (1971) attempted to rescue the concept of minimum monophyly. Farris (1974) published a more algorithmic approach, using the concept of group characters such that if a group character was unique and unreversed the group was monophyletic, one reversal, paraphyletic, and convergence in the group character itself indicated polyphyly. Oosterbroek (1987) found this approach to be ambiguous and suggested that the cause was the behavior of the group membership characters. Oosterbroek (1987) suggested that paraphyly and polyphyly be distinguished with reference to sister groups (as Nelson had done for monophyletic groups). Paraphyletic groups are groups that exclude one or more species or monophyletic group from an otherwise complete sister group system that are not, themselves, grouped together. Polyphyletic groups are groups that exclude one or more paraphyletic groups from a complete sister group system.

Two examples will suffice to illustrate these concepts. Reptilia is paraphyletic because it excludes one (Aves) or two (Aves and Mammalia) monophyletic groups. Homeothermia (Aves+Mammalia) is polyphyletic because it excludes a paraphyletic group (Reptilia). This seems to work because polyphyly always generates paraphyly in its wake.

NODE-BASED AND STEM-BASED MONOPHYLY: SAME CONCEPT DIFFERENT GRAPHS

Much has been made of the idea that monophyly can be described as being node-based or stem-based (e.g., de Queiroz and Gauthier, 1994, and the PhyloCode). This idea is an illusion based on a misinterpretation of two kinds of phylogenetic diagrams: stem-based trees and node-based trees. All monophyletic groups are composed of an ancestral species and all descendants of that species. If one uses node-based trees as the cartographic device depicting relationships, then all monophyletic groups are node-based because the ancestor is shown on the directed graph as a vertex (usually a circle at a node). If the cartographic device is a stem-based tree graph, then the ancestor is shown as a line leading to the first speciation event within the monophyletic group and all monophyletic groups are stem-based. This is discussed in more detail, with diagrammatic figures illustrating this point, in Chapter 4.

CHAPTER SUMMARY

- A taxon is a group of organisms given a name.
- In the phylogenetic system, natural taxa are monophyletic *sensu* Hennig (1966).

- Hypotheses that a particular group of organisms are monophyletic are hypotheses that the group includes an ancestral species (named or unsampled) and all of the descendants of that ancestor.
- Such hypotheses are corroborated through character analysis.
- Classifications that contain only groups hypothesized as monophyletic groups are logically consistent with the phylogenetic hypothesis upon which the classification is based. Such classifications are also logically consistent with hypothesized homology statements.
- Classifications that contain paraphyletic or polyphyletic groups are neither logically consistent with the hypothesized phylogeny nor with one or more of the hypotheses of homology.
- Genealogical concepts of monophyly are always equivalent, but the diagrams used may be different and have different interpretations, leading to the false idea that some monophyletic groups are stem-based while others are node-based.

4

TREE GRAPHS

Trees diagrams used in phylogenetics are tree graphs depicting the genealogical relationships of organisms. Haeckel (1866) is usually regarded as the first evolutionary biologist to publish tree diagrams meant to show the evolutionary relationships among actual organisms. (Darwin, 1859, contains an evolutionary tree relating hypothetical organisms.) However, tree diagrams that look for all intents and purposes like phylogenetic trees can be found in many earlier works (see Archibald, 2009) that ascribe relationships to either "transmution" or divide guidance. Many early trees, such as Agassiz (1844), are similar to more modern trees such as those appearing in Romer (1966), as pointed out by Patterson (1977). However, we can be sure that Agassiz's tree symbology, the meanings of lines, shapes, and labels, differed from Romer's. Agassiz rejected evolution; Romer embraced it. Gould (1999) has argued that Lamarck (1809) produced what may be the first truly evolutionary tree, although tree iconography is common in earlier works dating at least to Pallas in 1766 and Augier in 1801 (Archibald, 2009).

Tree diagrams have been depicted in various forms for various purposes. Haeckel's (1866) diagram was actually a tree, complete with trunk, limbs, and leaves. Romer's (1966) diagram of vertebrate phylogeny incorporates a time frame and an estimate of the numbers of species through time in each group and portrays groups emerging from other groups. The more recent and explicitly phylogenetic tree of Stiassny et al. (2004) eschews the paraphyletic groups of Romer, but captures the idea of diversity through time. Milne and Milne's (1939) tree of caddis flies is presented in three-dimensional projection and incorporates such features as habitat and casing construction. Many of Hennig's (1966) diagrams have arrows rather than lines. Some incorporate characters, but others do not. In short, tree diagrams come in many shapes and forms and may emphasize different things.

One of the problems facing phylogeneticists is the fact that different diagrams that appear in the same form have different symbologies. Indeed, Hull (1979:420) expressed the concern that "uncertainty over what it is that cladograms are supposed to depict and how they are supposed to depict it has been one of the chief sources of confusion in the controversy over cladism." The first major objective of this chapter is to sort out the various kinds of trees. We will recognize several types of trees including phylogenetic trees of two sorts (stem- and node-based), Nelson cladograms, and gene trees. Differences in how authors have used the term "tree" and what they connoted when they presented a tree have created various debates, arguments, and misunderstandings among phylogeneticists. We believe that many of these misunderstandings can be obviated by parsing out what various authors did or did not mean when they discussed or presented trees. The second major objective is to make the connection between the empirical results of character analysis and various kinds of tree diagrams.

In graph theory, a tree is any connected, acyclic graph. In general, *tree* corresponds fairly closely with the concept of hierarchy discussed by Hennig (1966:16–18). Trees consist of two basic things, vertices (singular vertex) and edges. Vertices are often termed nodes, and edges may be termed stems, lines, or internodes. A vertex with only one edge connection is termed a leaf. Internal vertices ("internodes") have two or more edges that connect them with other vertices (Fig. 4.1a). We will discuss in detail two forms of trees, one where edges are taxa (stem-based trees; Fig. 4.1b) and the other where nodes are taxa (node-based trees; Fig. 4.1c). We will then discuss Nelson cladograms where the nodes are sets and the edges are inclusion relationships. Along the way, we will mention gene trees and how they may differ from "traditional" trees and cladograms.
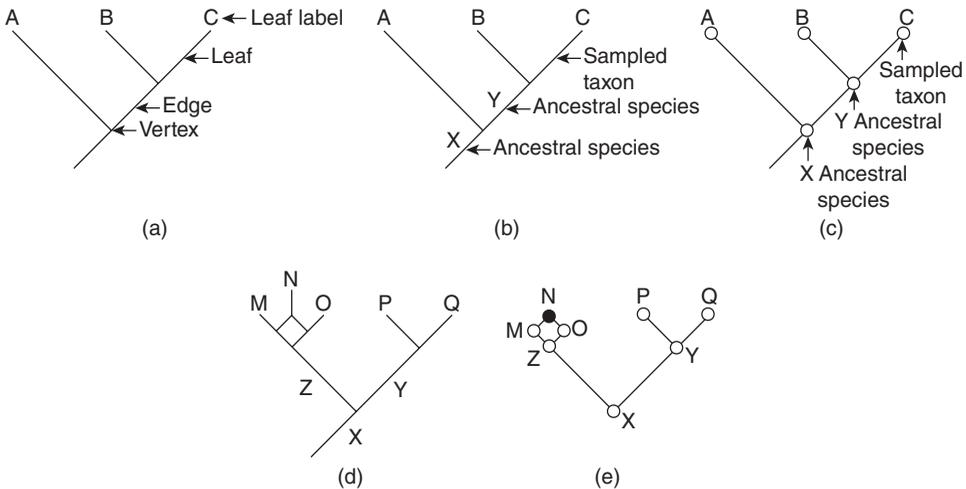
Figure 4.1. Acyclic and cyclic graphs. (a) Basic descriptive terms used to describe trees (acyclic graphs). (b) A tree with edges and leaves as taxa and nodes as speciation events. (c) A tree with taxa as vertices and edges as relationship (parent of) statements. (d) A cyclic graph of form (b) depicting the relationships of a hybrid species (N) to its parental species. (e) A cyclic graph of form (c) depicting the relationship of N to its parental species.

In contradistinction, cyclic graphs are not trees. Cyclic graphs in phylogenetics (often referred to as "phylogenetic networks") may show reticulate relationships (Fig. 4.1d, e) or may result in cases where character ambiguities preclude resolution of hierarchical relationships.

PHYLOGENETIC TREES

We recognize two sorts of phylogenetic tree graphs, stem-based and node-based, following Wiley (2010). As shown by Martin et al. (2010), these two kinds of tree graphs are simply two ways to illustrate the same kinds of relationships and are interconvertable, as logically shown by Hennig (1966) and mathematically demonstrated by Martin et al. (2010). They are simply two ways of showing the same relationships. Of course, these are not the only kinds of tree graphs, one must know what the author intended to understand the nature of the tree presented. However, we suspect that the natural way most phylogeneticists interpret trees is as a stem-based tree.

Stem-Based Phylogenetic Trees

Most examples in the literature of so-called phylogenetic trees depicting the descent of species and monophyletic groups are actually directed acyclic graphs in which the edges are taxa and the nodes are speciation events (Fig. 4.1b). These stem-based trees are attempts to capture some of the macroscopic historical processes of evolution that involve cladogenesis of species or monophyletic groups of species. The interpretation is that the internal edges are ancestral species and the terminal edges are either descendant species or descendant monophyletic groups represented by their ancestral species. We must understand that the internal edges represent the minimum number of ancestral species needed to connect descendants and not the actual number, which could only be known if we had a complete phylogeny of all descendants. One counts the minimum number of speciation events by counting the branching events. This yields the lower bound on the number of speciation events represented by the tree, not the upper bound.

Stem-based trees depicting the descent relationships of individual organism or the genes of organisms are also possible. Obviously, the meanings of vertices and edges are different for different hierarchical levels of organization. The transition between descent relationships at the level of individual organisms and at the level of species is what Hennig (1966) saw as the transition between tokogenetic and phylogenetic relationships.

It turns out that, of course, nature is too complicated to incorporate all aspects of evolutionary descent into a diagram as simple as a stem-based phylogenetic tree, and this is one reason why authors have developed different conceptual definitions for phylogenetic trees. Hennig's (1966:31) classic diagram portrays a small part of this complexity, illustrating the transition from systems of ontogeny (descent by mitosis and differentiation) and tokogeny (descent by reproduction), to systems of phylogeny and descent by speciation (Fig. 4.2). Phylogenetic descent occurs over time and space, involving one to many populations. Over time and within a lineage, relationships obtain between individuals partaking in reproduction and through reproduction establishing tokogenetic relationships. Occasionally, these tokogenetic
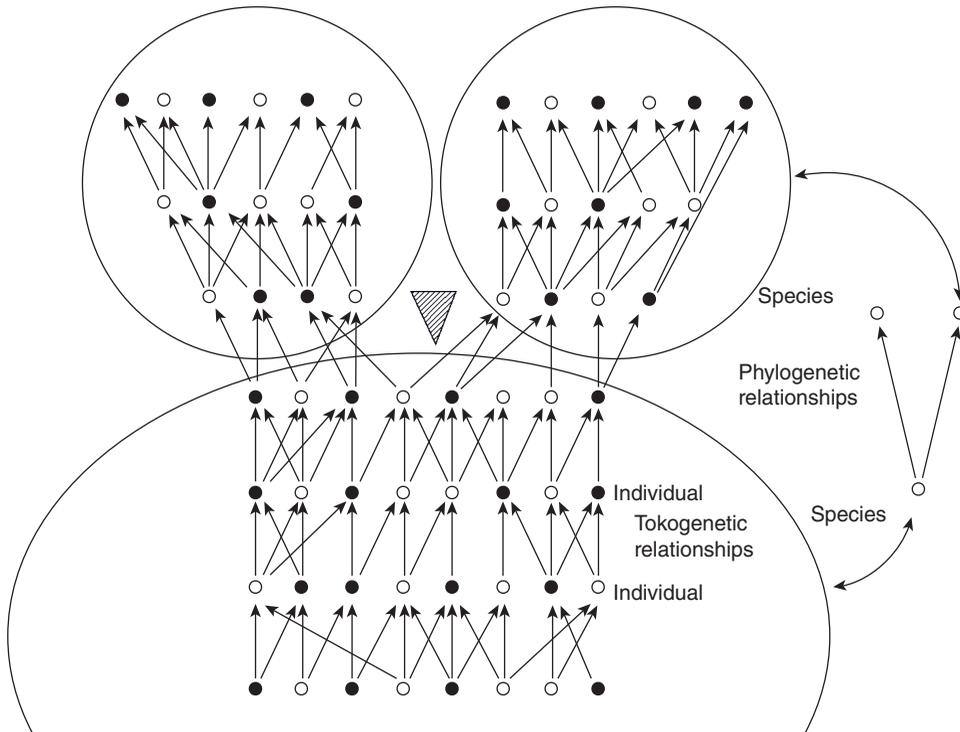
Figure 4.2. Hennig's concepts of tokogeny and phylogeny. Relationships on the left diagram portray tokogenetic relationships among individual organisms in a sexually reproducing population (individuals are open and closed circles). Species boundaries are indicated by lines around individuals. Relationships on the right are phylogenetic relationships among species on a node-based tree. From *Phylogenetic Systematics*. Copyright 1966, 1979 by the Board of Trustees of the University of Illinois. Used with permission of the author and the University of Illinois Press.

relationships are disrupted, resulting in establishment of new lineages that are new self-referential tokogenetic systems.

Analysis of this descent is restricted largely to specimens taken from these populations. Our symbolic representations can be quite simple, but they can be accurate relative to our hypotheses of relationship in the sense that if we knew the relationships we would find that the symbolic representations are logically consistent with those relationships. On the empirical level, we expect the tree to be logically consistent with the evidence at hand. It may be accurate without being complicated because the relationship between tokogeny and phylogeny is hierarchical and nontransitive: tokogenetic systems, if we can observe them, could be translated into phylogenetic systems, but the tokogenetic relationships among individual organisms cannot be recovered from a phylogenetic tree as we commonly draw them (Coleman and Wiley, 2001). An analogy is a system of highways: we can map the highways by accounting for every piece of gravel used to construct them, but we cannot account for every piece of gravel by consulting a highway map. Nevertheless, the map gets us where we wish to go; it is an accurate enough graphic representation of the macroscopic properties of the highway even though it does not account for its microscopic properties.

Consider Fig. 4.1b. It makes the following statements:

1. X speciates, giving rise to A and Y.
2. Y is the parent of B and C.
3. A, B, and C are species or clades of species.
4. C is more closely related to B than it is to A because both have the parent Y and A does not.

Phylogeneticists frequently talk about edges (Fig. 4.1b) such as X and Y as "hypothetical" ancestors. However, under the paradigm of evolution by reproduction, inheritance, and common descent, there is nothing more hypothetical about these entities than the sampled terminal taxa labeled A, B, C. Taxa A–C are hypothesized to be taxa based on specimens that have been examined. They are treated as objects of analysis, but A–C are also hypotheses brought to the analysis, with all of the background assumptions inherent in the hypotheses. The ancestral edges labeled X and Y are no more hypothetical than A–C. They represent the unsampled ancestral lineages required to assert the relationships shown among known descendant taxa.

Node-Based Phylogenetic Trees

A tree can also be drawn in another manner. Node-based trees are trees in which vertices/nodes are taxa and edges are statements of relationships or other properties shared by the taxa. These trees are also acyclic (Fig. 4.1c). The relationship between any one phylogenetic tree and the corresponding node-based tree is reciprocal: one is what graph theorists term *the line* of the other (Martin et al., 2010). Although he did not express himself in graph-theoretic terms, Hennig (1966) understood this relationship perfectly.

Hennig's (1966:59) Fig. 14, redrawn here as Fig. 4.3, illustrates the difference between a stem-based tree and a node-based tree. On the left is a stem-based tree with edges as taxa (Fig. 4.3a). On the right is a node-based tree with vertices as taxa (Fig. 4.3b). Hennig's symbolism comes straight from graph theory. The edges of the node-based tree (directed edges) are arrows (arcs in graph theory), and the arrows are statements of shared parent–child properties. Vertices of degree two or higher (i.e., vertices joined by two or more edges) are not empty and are ancestral species (unnamed), and the leaf vertices of degree one are taxa represented by specimens and given a name. In most instances, the leaf vertices are not shown as vertices but simply as labels.

To develop this and similar diagrams, Hennig (1966) used the symbology of Greg (1950), which Greg derived from Woodger (1952). The arrows (arcs) do not represent lineages, ancestors, or any other kind of taxon. Rather, they represent a concept: relationship. In particular, they state that one vertex is the ancestor of (or parent of) another vertex. In graph theoretic terms, we would say that the tail of the arc is the ancestor of the head of the arc. Arrows that point from only one entity to only one entity and in only one direction are characteristic of Hennig's (1966) concept of hierarchical relationships and typical of directed acyclic graphs. (Undirected acyclic graphs are termed *unrooted trees*, and we will deal with these later in the chapter.)
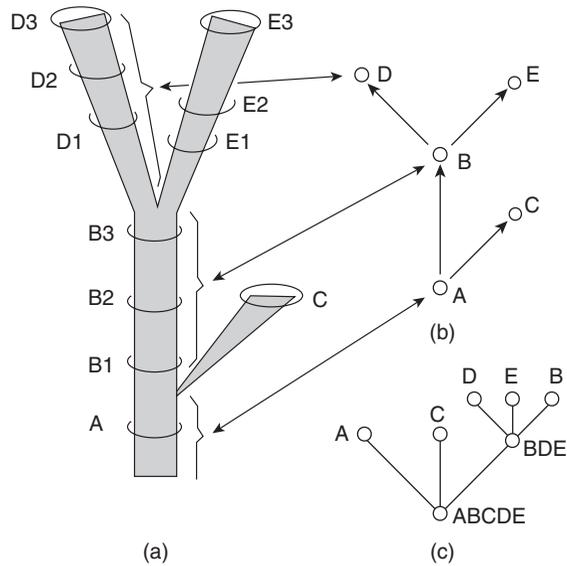
Figure 4.3. "Species category in the time dimension" (Hennig, 1966:59, Fig. 14). (a) A stem-based tree. Letters are symbols for species, and the number applied to each letter are labels of samples of each species at a particular time. (b) A node-based tree (species are nodes) with single-headed arrows symbolizing parent of relationships and nodes representing labeled species. Note the correspondence between the extended lineages in (a) and the nodes in (b), as shown by double-headed arrows and brackets. (c) A Nelson tree with ancestors displaced as terminal nodes and internal nodes interpreted as sets of taxa. Figures in (a–b) redrawn from *Phylogenetic Systematics*. Copyright 1966, 1979 by the Board of Trustees of the University of Illinois. Used with permission of the author and the University of Illinois Press.

    In the phylogenetic system, hierarchies represented as directed acyclic graphs are the markers of phylogenetic relationships. Hennig (1966) frequently, but not consistently, symbolized ancestral species as open circles at the nodes and known, terminal, taxa with solid circles at the tips. So, vertex B in Fig. 4.3b is the ancestor of vertices D and E. Vertex A is the ancestor of B and C and of the entire clade. This is opposed to the nonhierarchical tokogenetic (and acyclic) relationships represented by graphs where two edges lead to a single entity in sexually reproducing species (Fig. 4.2) or two parents of a taxon of hybrid origin.
    Hennig (1966; Figs. 4, 6, 14, 15) made clear his concept of the relationship between stem-based trees and node-based trees. The vertices at nodes of node-based trees symbolize the ancestral lineages of stem-based trees and the nodes of stem-based trees represent lineage splitting. Thus, the vertices of node-based trees do not represent lineage splitting, a speciation event, or any other process event; they simply represent the objects of study, either sampled (leaf) or unsampled or unrecognized (vertex). Conversely, vertices of acyclic stem-based trees exactly represent lineage splitting in the same manner as a fork on a road map represents a fork in the road.
    The edges of node-based trees represent some statement of ancestry relationship that exists among vertices while the edges of stem-based trees represent lineages that are hypothesized to exist in nature. Edges of node-based trees are symbols for "parent of" relationships among vertices. An edge of a stem-based tree graphically

represents a lineage, just as a line on a map represents a road. We will see in Chapter 8 that other interpretations of phylogenetic trees are incorrect and have created problems in logic and consistency with the proposed PhyloCode.

CYCLIC GRAPHS

Cyclic graphs can also be drawn in two configurations, somewhat analogous to stem- and node-based trees. Fig. 4.1d shows a cyclic graph with two speciation events and one hybridization event that leads to the origin of species N. Most of the edges are lineages, but the graph also shows a tokogenetic event, with "mating" lines leading from the two parental species to the reticulate vertex at the base of the N edge. The node-based tree is more straightforward (Fig. 4.1e), stating that N is equally related to its two ancestors M and O; it is the child of both. An example of the cyclic graph is show in Fig. 4.4. It is a map of haplotypes taken from two species of North American sand darters (genus *Ammocrypta*). The haplotype map demonstrates that coalescence has not occurred among the haploptypes on this short segment of the mitochondrial gene cytochrome-*b*.
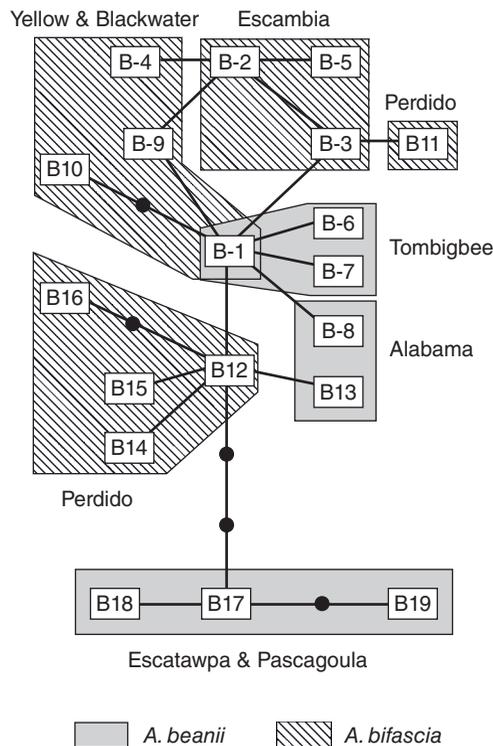
Figure 4.4. A network (cyclic graph) of haplotypes observed in a study of cytochrome-b for samples of two fishes, *Ammocrypta beanii* and *A. bifascia*. Numbered letters represent observed haplotypes. Closed circles represent unobserved haplotypes of one mutation step. Haplotypes of *A. bifascia* are in striped polygons, and those of *A. beanii* are in shaded polygons. Note that species boundaries do not correspond to haplotype relationships. From Wiley and Hagen (1997); copyright Academic Press, used with permission.

CLADOGRAMS

The term *cladogram* is commonly used to describe phylogenetic tree graphs. However, what and how many different kinds of cladograms there might be has been somewhat of a mystery (Hull, 1979). To many, a cladogram is simply a phylogenetic tree with unsampled ancestral species. To others, it is a common ancestry tree. As such, it might be stem or node based. To Gary Nelson (1979ms) and many of the so-called pattern or transformed cladists, it is any kind of acyclic graph where entities were clustered according to some property relationship. This is the general concept of a tree graph, but it begs the question of exactly what the vertices, edges and leaves mean. Disagreements about what connotes a cladogram have spawned significant scientific debate. Yet, much of these debates can be resolved by recognizing the implicit assumptions that various authors were using when they invoked the term *cladogram*. There appears to be no particular justification for accepting Nelson's (1979) concept of cladogram as the most preferred or valid.

The signal characteristic of what we refer to as Nelson cladograms is that all sampled taxa are leaves. Internal edges or nodes specify an undefined relationship property. Thus, even an acyclic graph generated by phenetic clustering, a graph usually termed a phenogram, is one kind of cladogram according to Nelson (1979ms). Nelson (1979ms) was never published, but the idea that cladograms were fundamentally different from phylogenetic trees (and scenarios) appears in Eldredge and Tattersall (1975) and Tattersall and Eldredge (1977), is cited as early as 1977 by Platnick (1977), and formed the conceptual framework for cladogram as defined in Eldredge and Cracraft (1980), Nelson and Platnick (1981), and many other subsequent works. Using such a perspective, cladograms can be trees of common ancestry (Platnick, 1977), x-trees (Nelson, 1979ms, "x" being unspecified), or synapomorphy schemes (Nelson and Platnick, 1981). Assertions that cladograms are not trees (Nelson and Platnick, 1981:171) tried to draw the distinction between trees with specified ancestors and those without, but the assertion is incorrect from a graph theoretical perspective (Hendy and Penny, 1984). In fact, any acyclic graph is a tree. As we will see later, this does not mean that Nelson cladograms are simply stem- or node-based trees. Thus, considerable confusion (expressed by Hull, 1979) surrounded the exact meaning of Nelson's cladograms.

Nelson Trees in Phylogenetics

Any acyclic graph is a cladogram in the broadest sense of Nelson (1979ms). This would include phenograms, acyclic graphs that portray one-to-many relationships based on similarity properties. For purposes of discussion, we will restrict ourselves to phylogenetic cladograms, acyclic graphs generated by grouping by synapomorphy and with all taxa residing as leaves. We will term such trees *Nelson trees*, following Matrin et al. (2010). Nodes represent taxa, and the edges are inclusion relationships. Because all named taxa are displaced to the leaf position, Nelson trees will differ from either stem-based trees or node-based trees if an actual ancestral species happens to be included in the analysis. (Englemann and Wiley [1977] note that such a circumstance is unlikely but not impossible.) To see this, consider Fig. 4.3 again. A Nelson tree of the relationships among the taxa ABCDE would look like Fig. 4.3c.

It is apparent that the topology of the Nelson tree in Fig. 4.3c is different from Figs. 4.3a and b because the ancestral species A and B are displaced to a leaf position. One interpretation that might be made is that Nelson trees depict the relationships among sets of taxa. This is shown by the labeled nodes in Fig. 4.3c. Internal nodes are inclusive sets, and edges link more inclusive sets of taxa, with the character properties of each set being synapomorphies. This interpretation leads to an important similarity when we consider the empirical evidence: Exactly the same monophyletic groups are circumscribed in Nelson trees and either node-based or stem-based trees. Synapomorphies diagnose monophyletic groups in the phylogenetic system. Further, ancestral species are the founders of monophyletic groups and belong to the group they founded. Thus, synapomorphies that circumscribe a monophyletic group must, by definition, include the ancestral species because it is in the ancestral species that the synapomorphies were fixed and passed on to its descendants. Nelson might reply that his cladograms do not imply any descent relationship. However, his concept of synapomorphy does ultimately rely on reproduction and inheritance.

From Nelson Trees to Phylogenetic Trees

When it was first appreciated that what we term Nelson trees might be different from phylogenetic trees (e.g., Eldredge and Cracraft, 1980), it was thought that there were many phylogenetic trees for any one cladogram. (In this particular discussion, for purposes of brevity, we will use the term *cladogram* to refer to *Nelson tree* and *tree* to refer to either form of a phylogenetic tree.) For any three taxa, the number of cladograms is four, but the number of trees was thought to range from 13 to 22 (Cracraft, 1974; Harper, 1976; Platnick, 1977). For example, 13 are shown in Fig. 4.5. To some, this suggested cladograms would have limited utility for evolutionary biology, although some argued that the situation was not all that bad (e.g., Wiley,
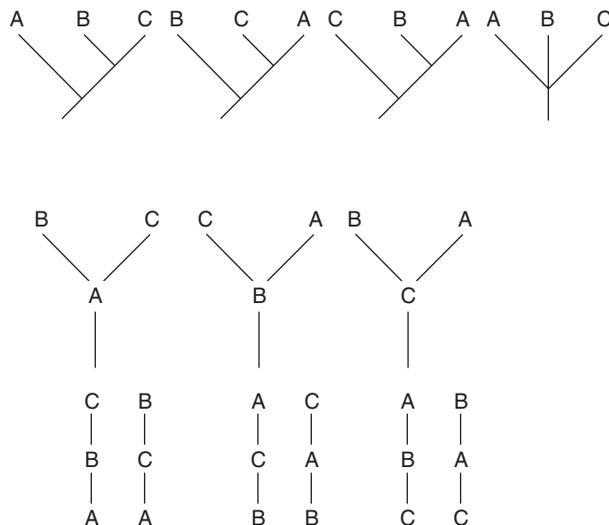

Figure 4.5. Various interpretations of the relationship of three entities: A, B, and C.

1979a, b, 1981a, 1987; Eldredge and Cracraft, 1980). The major issue is whether cladograms of taxa at different levels of organization (different kinds of objects) might be different from phylogenetic trees at the same levels of organization. We suspect that they will not be, and we consider two basic levels of phylogenetic analysis of entities to show this. The first level is where a decision has not been made as to whether the samples analyzed form tokogenetic or phylogenetic systems. The second level is where some a priori decision has been made that the entities analyzed are parts of a phylogenetic system; that is, they are species or monophyletic groups of species.

Gene trees of individuals or haplotypes are examples of analyses in which no commitment has been made about the species-level entities that might be present in the analysis. Consider the tree shown in Fig. 4.6, the parsimony analysis of individual haplotypes among several presumed species of sand darters (based on the same data that the haplotype network in Fig. 4.4 was constructed). Obviously, the vertices and edges of this tree do not portray speciation events and species lineages. Rather, they display a hypothesis of the relationships among haplotypes, a gene tree. In some cases, the hierarchy of the haplotypes seems to correspond to a hierarchy associated with species (for example, the haplotypes near the root of the tree), but in other cases, haplotype relationships are either unresolved or shared among the presumed species used in the analysis. We might suspect that this particular gene tree demonstrates incomplete coalescence among the species labeled V, M, and B. Attempts to translate this graph into anything but a gene tree would be futile. Gene trees cannot be automatically translated into phylogenetic trees. This presents a special problem as we enter the era of genomics where systematicists will employ an increasing number of genes and gene products in phylogenetic research. A review of the challenge of discriminating between gene trees and species trees has been published by Degnan and Rosenberg (2009). They suggest that there are a number of outstanding questions that need clarification to achieve the integration of micro- and macroevolutionary processes in order to achieve an understanding of gene tree and species tree discordance.

Now, consider the cladogram in Fig. 4.7 relating lanternfishes. The taxa analyzed are considered (for purposes of the analysis) to be monophyletic groups. There are no species in the analysis, so there is no possibility that some species occupies a vertex. Further, no taxon can occupy a vertex because all taxa are hypothesized to be monophyletic (and monophyletic taxa cannot be ancestors). So again in this case, the Nelson tree would have the same topology as a phylogenetic tree.

For fully dichotomous cladograms involving species or monophyletic groups, the topology of the cladogram will exactly match the topology of the phylogenetic tree hypothesis given the adoption of species as individuals and the rejection of speciation via phyletic gradualism (Wiley, 1979a, b, 1981a, 1987) except when dealing with taxa of hybrid origin. This is true even for dichotomous cladograms involving species that lack autapomorphies (Eldredge and Cracraft, 1980; Wiley 1981a). One might object and respond that what we took as a single species might be two species or that an ancestral species might have been sampled before the evolution of the autapomorphy that it would pass to its descendants. Either case might be true, but neither helps. These assertions are simply assertions that our empirical analysis may be disconfirmed by the discovery of new characters or new specimens. But as empirical scientists, we are tied to the evidence we have before us, not to an infinite number
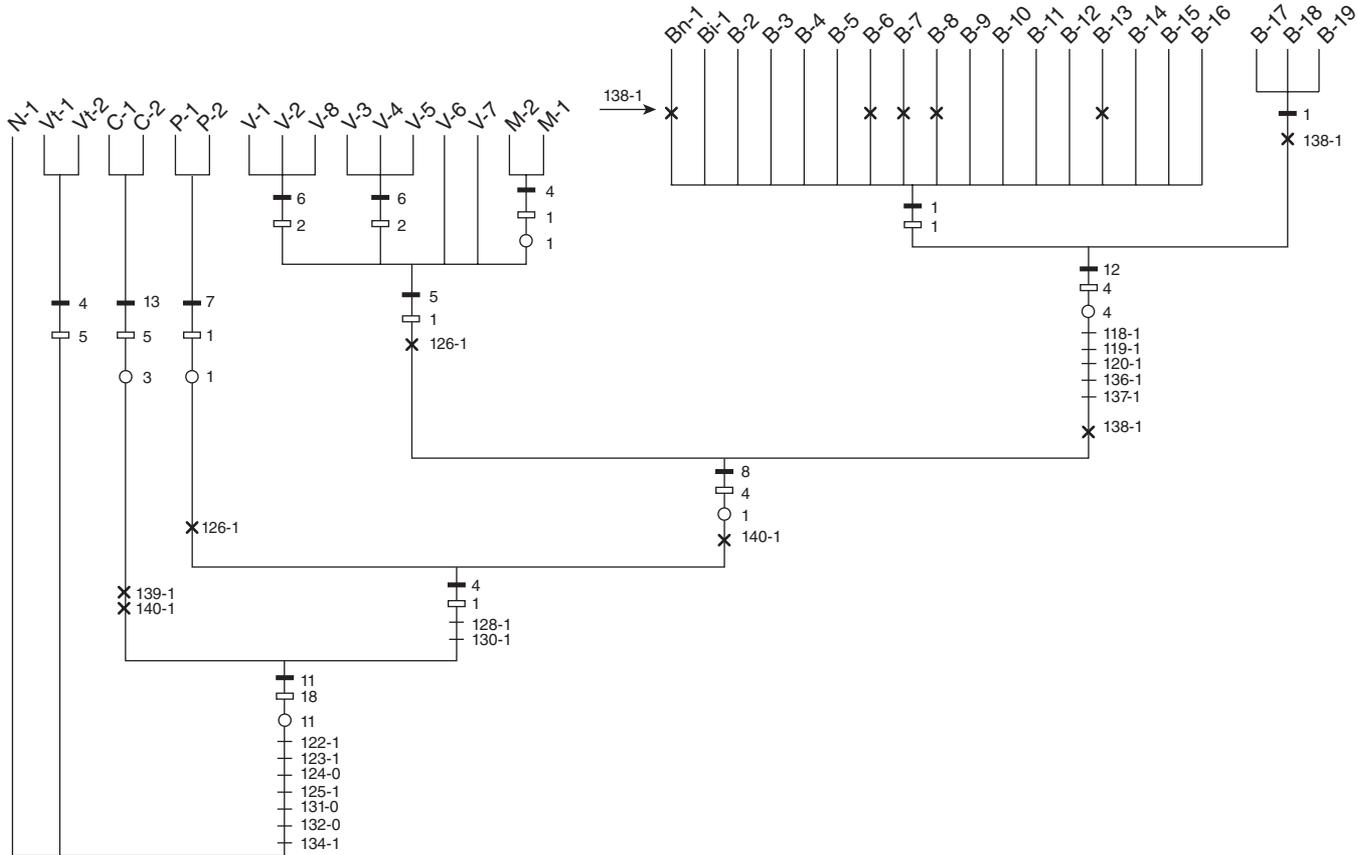
Figure 4.6. A total evidence parsimony analysis using equally weighted DNA bases and morphological characters. Dashes are characters that are unique and unreversed; "x" denotes characters that show homoplasy. From Wiley and Hagen (1997); copyright Academic Press, used with permission.
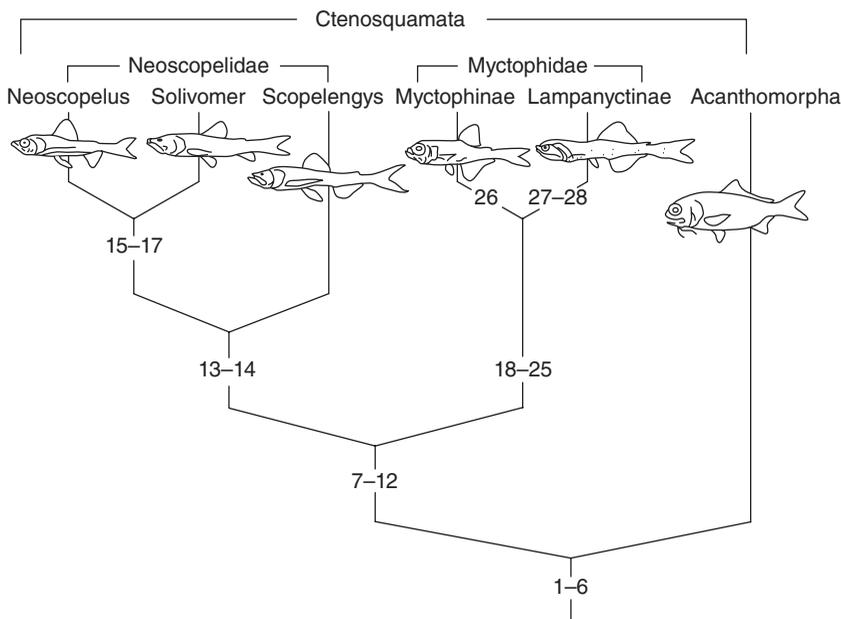
Figure 4.7. Scheme of relationships among ctenosquamate fishes (lanternfishes and kin plus higher teleosts). From Stiassny (1996); copyright Academic Press, used with permission.

of possible future discoveries. Empirical analysis does not guarantee that we arrive at the correct result, only that the results we arrive at can be justified by the data we have.

That Nelson trees including taxa of hybrid origin do not directly translate into evolutionary trees is well known to botanists (Bremer and Wanntorp, 1979; Humphries, 1980, 1983; Funk, 1981, 1985; see also Platnick, 1985). Various kinds of phylogenetic networks emerge from the analysis of a set of taxa that contain species or clades of hybrid origin (Funk, 1985). One simple pattern is for the hybrid species to be intermediate in sampled morphology and derived from two sister species, resulting in a trichotomy (Fig. 4.8b). Another simple pattern is for the hybrid to show more derived characters of one parent species than the other, either due to character sampling or to dominant expression, resulting in a dichotomy (Fig. 4.8c). The effect of including hybrids in a phylogenetic analysis was investigated in some detail by McDade (1990, 1992, 1997) using known hybrids produced in the laboratory. McDade (1992) summarized the predictions about the behavior of hybrids in phylogenetic analysis made by various authors (e.g., Bremer and Wanntorp, 1979; Nelson and Platnick, 1980; Hill and Crane, 1982; Humphries, 1983; Nelson, 1983; Wanntorp, 1983; Funk, 1985): (1) increased levels of homoplasy expressed as poorer fit of the character set to the tree, (2) increased number of equally parsimonious trees that will collapse into poorly resolved consensus trees, and (3) distortion of the patterns of relationships among taxa that are not of hybrid origin.

McDade's earlier work on patterns of character variation among the hybrids (McDade, 1990) led her to conclude that many of the theoretical predictions might
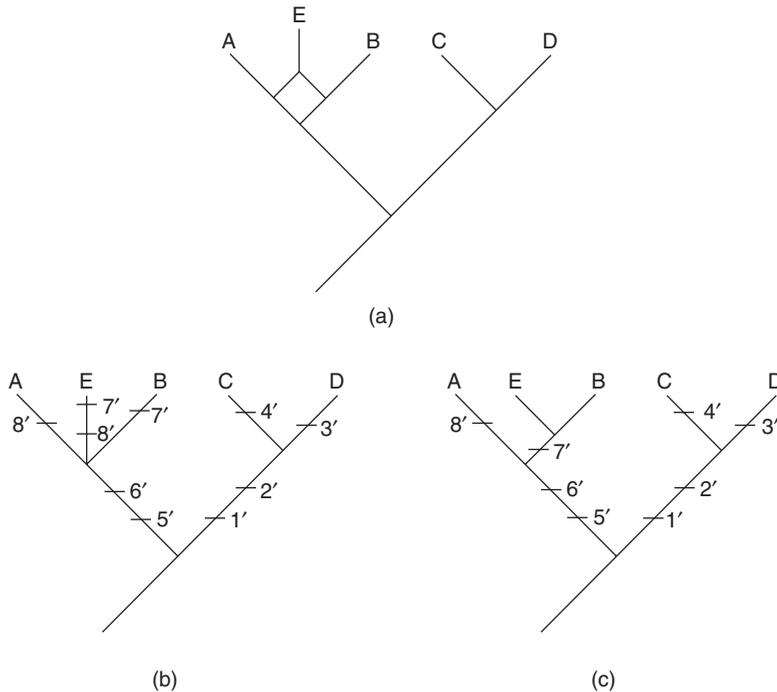
Figure 4.8. Cyclic phylogeny and trees. (a) A cyclic phylogeny showing the hybrid origin of species E. (b–c) Two simple results that might obtain from phylogenetic analysis of the taxa. Note that actual analysis discussed in the text may lead to different cladograms than those shown here.

be false. She tested the predictions using combinations of one to five hybrids inserted into a matrix composed of parental species of nonhybrid origin that varied between 16 and 11 species to maintain a constant sized data matrix.

In general, McDade found that hybrids mostly had intermediate characters between the parental species (McDade, 1990) and that hybrids are most frequently placed basally in the clade that includes the most apical parent. This was accompanied by only modest increases in homoplasy and no drastic increases in the number of most parsimonious trees. These findings contradict the predictions that hybrids will express the derived conditions of each parent and lead to poorly resolved results. Thus, their existence does not create intractable problems for phylogenetic analysis. Indeed, because of this sometimes it may be difficult to ascertain the nature of hybrid species through phylogenetic analysis alone. Here is a case when several trees are compatible with a single cladogram. To determine which one will often require noncladistic criteria that can be applied to identify the hybrids and their parents (as suggested by Wagner, 1983; see also cited literature in McDade, 1992, for molecular approaches to hybrid identification).

In the absence of taxa of hybrid origin (that is, most of the time for the vast majority of groups analyzed), the only time a cladogram will have a topology different from a tree is when an ancestral species is actually present in the analysis and
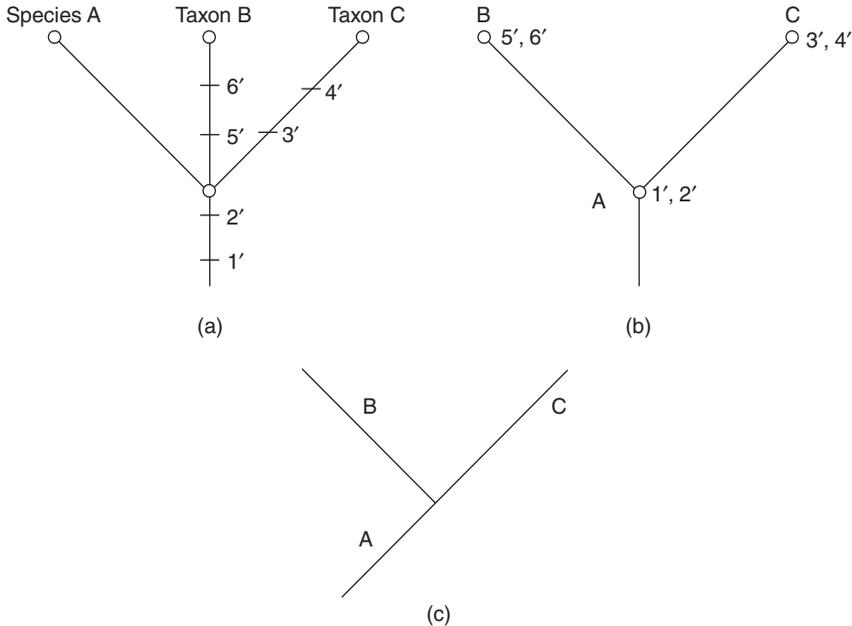
Figure 4.9. A cladogram and two trees. (a) The cladogram with the ancestral species A displaced to leaf position. (b) A node-based tree. (c) A stem-based tree. Note that character homologies are not distored by the cladogram (a) but that the implications of relationship are different.

is displaced to a leaf position in a polytomy with its descendants (Fig. 4.9a; that is, the case that Eldredge and Cracraft, 1980, noted, Fig. 4.13). This requires a species to be in a polytomy with either other species or monophyletic groups of species and that the ancestor candidate cannot have any autapomorphies. Of course, this topology has an alternative explanation, an unresolved dichotomy. The rejection of the hypothesis that the species is an ancestor requires only that one find an autapomorphy for it or that one resolve the polytomy (for example, finding a synapomorphy in B and C to resolve Fig. 4.9a). Further, there are other types of information that can be used to adduce this, for if we are dealing with fossils and the species is a contemporary of the other taxa, then the case for an unresolved polytomy increases. (Having said this, even this would not definitively reject the fact that that species is the ancestor, but the evidence at hand would suggest the hypothesis should be rejected.) Thus evidence outside the realm of character analysis, such as stratigraphic and biogeographic data, is needed to corroborate the hypothesis that one is actually dealing with an ancestral species (e.g., Prothero and Lazarus, 1980).

Please note that we do not claim that because Nelson trees and phylogenetic trees usually have the same topologies (ancestral species and hybrid taxa excepted) they have the *correct* topologies. For example, a maximum likelihood tree of molecular characters may only imply one evolutionary tree, but the tree it implied might be quite incorrect due to a variety of factors such as incomplete lineage sorting that make gene trees different in topology than the phylogeny of the organisms

(Maddison, 1997). Or, a morphological analysis might have concentrated on too narrow a range of morphology that turns out to be correctly analyzed but not congruent with a larger set of morphology. A nice example of reaching an incorrect conclusion based on a narrow range of morphology is the analysis of Wiley (1979d) where Wiley concluded that coelacanths were the sister of sarcopteryigians+actinopterygians on the basis of ventral gill arch morphology. The overwhelming weight of evidence, viewed more broadly, is that coelacanths are sarcopterygians.

GENE TREES

Except for the sand darters discussed earlier, we have worked under the assumption that character evidence was evidence of the descent of species. However, this may not always be the case; genes may have their own descent patterns (e.g., Fitch, 1970; Goodman et al., 1979) that results in a gene tree that is different from a species tree. This could be due to a number of factors (Maddison, 1997) including gene coalescence (Hudson, 1990). The human–chimp–gorilla controversy illustrates the problem. Ruvolo (1997) found that among 14 DNA data sets, 11 support the hypothesis that chimps are the closest living sister species of humans but two supported the chimp–gorilla relationship and one supported the gorilla–human relationship.

   Another common reason for mismatch between gene trees and species trees is the availability of only short gene sequences (easily fixed by gathering more data; Saitou and Nei, 1987) and incomplete lineage sorting (not easily fixed although the probability can be addressed under an assumption of neutrality given that the gene tree is correct; Hudson, 1992; Nei, 1986; Pamilo and Nei, 1988). Less common in eukaryotes and more common in prokaryotes is horizontal gene transfer (Maddison, 1997).

   The most common way to guard against interpreting a gene tree that does not reflect a species tree as a species tree is to employ multiple independent (unlinked) genes in the analysis. Returning to our example, Ruvolo (1997) used the multiple-locus test of Wu (1991) to accept the chimp–human sister group relationship while rejecting the alternatives at $P = 0.002$ within a likelihood framework. The test evaluates gene–tree/species–tree mismatches using the likelihood ratio test. The use of multiple unlinked genes to detect the mismatch between gene trees and taxon trees is becoming standard in phylogenetic analyses of DNA sequence data, regardless of the method of analysis.

INDIVIDUALS VERSUS SETS OF INDIVIDUALS USED IN AN ANALYSIS

As empirical constructs, trees are the end result of character analysis. The fact that species and monophyletic groups are ontological individuals and that, for example, Nelson trees deal with sets does not create any logical difficulties. After all, it would be intractable to work with entire species or monophyletic clades when we seek to reconstruct trees. Instead, we must be content to work with exemplars or sets of specimens picked for a particular analysis to represent, for example, a species. We never work with all specimens of a species and we never are sure that we work

with all species of a clade (some might be extinct and others undiscovered); thus we work with a subsample of diversity, a set of selected taxa. Further, we do not work with all characters, or even all apomorphic characters (though, of course, it would be nice if we could find them all); instead, we work with a selected set of characters. However, we should not confuse the way we study a thing, an epistemological issue, with the ontology of the thing itself. Just because we work with sets of specimens does not mean that species are sets. Just because our set of specimens has a set of studied characters, does not mean that characters are sets. Ontological issues, such as the issue of whether taxa are kinds or individuals (we endorse the latter view) is the background knowledge that forms part of the basis for why we select particular sets of taxa to compare and what we think of the nature of the taxa we select to analyze. Our understanding of the nature of characters also guides us in our selection of characters and our interpretations of the results.

REPRESENTING CHARACTER EVOLUTION ON TREES

Phylogeneticists have traditionally shown the support for trees by mapping characters unto them. There are various ways to map characters, which we discuss in Chapter 5. In this section, we are concerned with the potential confusion that might be caused by working with different sorts of trees.

Traditionally, there are three common ways to map characters on trees. Hennig (1966:91) showed support by placing the synapmorphies across leaves (Fig. 4.10a). Modified versions of such diagrams move the character bars down the tree to make the groupings more explicit and easier to read. The empirical claim is that edges bisected by character bars connect known taxa that have the character, and thus bisection of two edges denotes sharing of the same character. This method, however, is almost never used today.

Another place to map characters on a phylogenetic tree is along the ancestral lineage edges (Fig 4.10b). This now common method apparently was first used by Peter Ashlock. Characters (or their modifications) symbolized by lines, bars, circles, etc., that bisect an ancestral edge are shared by descendants of that ancestor unless replaced by a modification (or loss) of that character. The order of the characters and the relative time of appearance along the individual edge are arbitrary and indeterminate because there is no ancestor to analyze. The only claim that can be made is that it is hypothesized that the characters in question were fixed as autapomorphies somewhere along this ancestral edge, between the two known speciation events. In fact, we cannot even claim that multiple innovations were fixed in the same ancestor, only that between the two hypothesized speciation events two innovations were fixed. It is quite possible (and perhaps probable) that one day one or more taxa will be discovered that bisect that edge, thereby modifying the original hypothesis from one inferred ancestor to two or more inferred ancestral edges.

The natural place to show characters on node-based trees or Nelson trees is at the nodes (Fig. 4.10c) or in a list that references the nodes (the usual practice due to publishing constraints). The claims about characters placed at nodes are exactly the same as claims about characters placed on edges, but one must bear in mind that characters move up when converting stem-based trees to node-based trees, not down. The edges of node-based trees are relationship terms that point toward the apex of the tree. In terms of graph theory, this is because stem-based
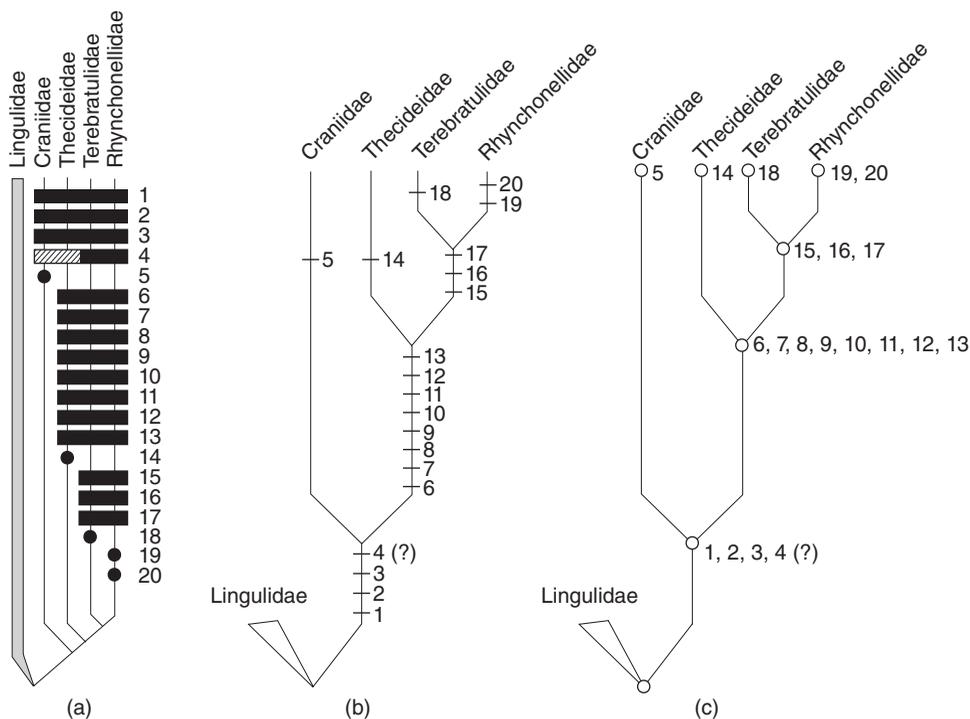
Figure 4.10. Ways of mapping characters on trees. (a) Mapping across leafs. (b) Mapping on edges and leafs. (c) Mapping on nodes. Group example (a) from *Phylogenetic Systematics*. Copyright 1966, 1979 by the Board of Trustees of the University of Illinois. Used with permission of the author and the University of Illinois Press.

trees are "planted" by an edge while node-based trees are "rooted" by a node (Martin et al., 2010).

UNROOTED TREES AND THEIR RELATIONSHIP
TO PHYLOGENETIC TREES

Modern phylogenetic methods use computers, and most software packages use unrooted tree graphs for their computations. Thus, it is necessary to understand the relationships between unrooted trees and rooted trees. Unrooted trees are acyclic graphs. They specify a particular number of possible routes one might follow to get from one taxon to another, but they restrict the number of routes taken. In other words, they are logically consistent with a limited number of possible rooted trees. For example, with four terminal taxa there are 15 possible dichotomous trees (Fig. 4.11a). If we take a particular unrooted tree (Fig. 4.11b), only 5 of the 15 are possible dichotomous solutions (enclosed in boxes in Fig. 4.11a). This is because unrooted trees specify particular paths through the tree that preclude other possible trees. To get to B from C, you encounter a node that branches to A, and you cannot proceed from B to C by bypassing this node. Thus, A and B will be adjacent on any rooted solution to the unrooted tree. Now, being adjacent on the unrooted tree does not mean that A and B will form a monophyletic group on the rooted phylogenetic tree:
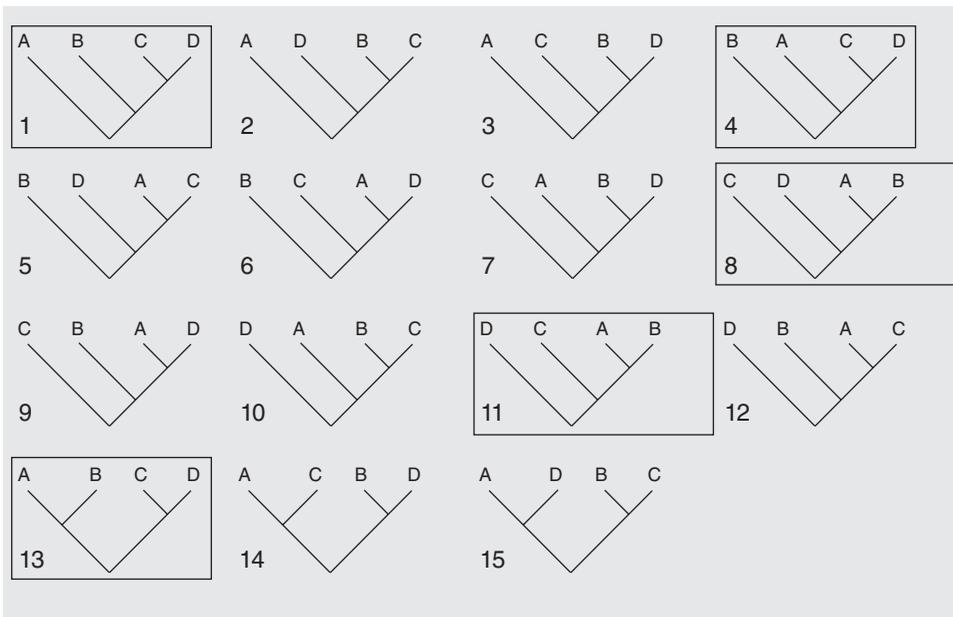
(a)
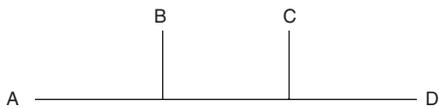
(b)

Figure 4.11. Rooted and unrooted trees. (a) The 15 possible dichotomous trees for 4 taxa (A, B, C, D). (b) An unrooted tree of the same taxa. Note that only 5 of the 15 trees in (a) are logically consistent with the unrooted tree. These are boxed in (a).

of the five possible phylogenetic trees only three hypothesize that A+B form a monophyletic group. However, the unrooted tree specifically prohibits a monophyletic group A+C to the exclusion of D. The inclusion of outgroups allows one to convert an unrooted tree into a rooted one.

NODE ROTATION

Leaves and edges can be rotated freely without changing the hypothesis of relationships. This free rotation can cause students and others unfamiliar with trees to misinterpret relationships. In Fig. 4.12, we show the same genealogical relationships among some major groups of tetrapods under four rotations. Perhaps because of the lingering effects of the *Scala Naturae*, many people expect to see humans positioned on the upper right, and they become disoriented when presented with other rotations (our experience in teaching undergraduates). However, the relationships of humans to the other tetrapods does not change under any of these rotations.
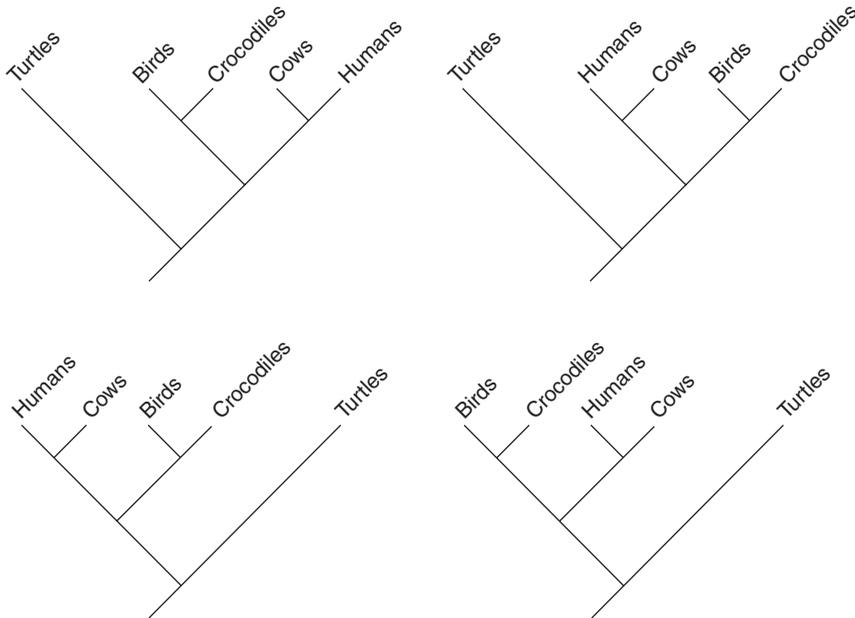
Figure 4.12. The same phylogenetic tree under different node rotations.

OTHER KINDS OF TREE TERMINOLOGY

The various kinds of trees dealt with by phylogeneticists belong to the class of additive trees. Edges connect vertices and leafs, but edges can also have weight. For example, the weight of an edge might be a measure of how different a child is from its parent or a descendant species is from its ancestor. In phylogenetics, this weight can take many forms, but the usual form is expressed as a patristic distance, a measure of how much evolution has occurred between parent and child or between two children. Additive trees are those trees where the differences in similarity between taxa are expressed as patristic distances along edges. For example, in Fig. 4.10b, the difference between thecideids and terebratulids is four steps and between thecideids and rhynchonellids is six steps because we travel from these taxa along the leaves and edges. If expressed as a distance, it is termed a patristic distance. What we count in the measure depends on the metric used. Some methods count steps (i.e., maximum parsimony), others use estimates of branch length (i.e., likelihood). Phenetic trees usually belong to another class of trees, nonadditive trees, because they measure difference across taxa and not along the internal edges of the tree. The only time phenetic trees are additive is when evolution is proceeding at a strictly constant rate. Note that the term *additive* is also used to describe methods of analysis using distances (Swofford et al., 1996).

We have gone to some lengths to distinguish between three types of trees in order to understand their relationships. In the literature, such fine distinctions are not made. With this in mind, below are some of the more common terms for trees.

Cladogram. As used in many analysis packages, a cladogram is a parsimony tree where the weight of the edges is not relevant. This is the original concept of Camin and Sokal (1965) and is usually what is meant when encountered in the literature.

Phylogram. Usually, this refers to a phylogenetic tree in which the length of the edges is proportional to the amount of relative change along the edge. The length of the edge, or branch length, may be an expression of the number of discrete steps or a measure of patristic distance.

Consensus Tree. Consensus trees are graphs that summarize the common knowledge claims of different phylogenetic trees. Most consensus trees are computed for different trees containing the same taxa, as when there is more than one most parsimonious tree for a particular set of data. *Supertrees* are a special class of consensus trees where knowledge claims of two or more smaller trees are combined into a larger tree. Supertree consensus techniques require that some number of taxa be common to both smaller trees (see Bininda-Edmonds, 2004, for a recent complication of papers on supertrees).

Network. Some, such as Huson and Bryant (2005), use this term to refer to any and all kinds of phylogenetic graphs. Others use the term for unrooted trees. Herein we use "network" to refer to cyclic graphs, that is, graphs where there are two distinct pathways to one or more taxa. Phylogenetic descent containing reticulate speciation or horizontal gene transfer events can be expressed as networks, and the use of networks to address questions of gene trees versus phylogenetic trees will surely increase as the properties of networks are better understood (e.g., Nakhleh and Wang, 2005). On the face of it, networks might seem anathema to phylogeneticists. However, networks are likely to be of increasing use to phylogeneticists as they (1) grapple with the acyclic nature of some phylogenetic trees caused by species of hybrid origin and (2) use networks to study character incongruence. The use of networks in phylogenetics is still in its infancy, but a review has been provided by Morrison (2005).

Classification Tree. This is a tree diagram of a classification made for the purpose of checking the logical consistency between a proposed phylogeny and a formal classification. The use of classification trees was covered in Chapter 3.

It is also common to characterize trees using the method of analysis. A *Wagner tree* is a parsimony tree that results from a Wagner analysis. A *maximum parsimony tree* is a tree that results from using the maximum parsimony optimality criterion and may be referred to as a Wagner tree. A *maximum likelihood tree* results from an analysis using the maximum likelihood optimality criterion, and a *Bayesian tree* is one that results from a Bayesian analysis. Both maximum likelihood and Bayesian trees consider branch length in their calculations and are thus forms of phylogenetic trees. The interpretation of maximum parsimony trees depends on the intention of the author and their interpretation of the nature of edges and nodes.

CONCEPTS OF MONOPHYLY AND TREES

We discussed the nature of monophyletic groups in some detail in Chapter 3, but it is worthwhile to revisit that concept here. The reason being that some phylogeneticists have claimed that there are three ways of characterizing monophyly (e.g., de Queiroz and Gauthier, 1992, 1994), and they have attempted to distinguish these
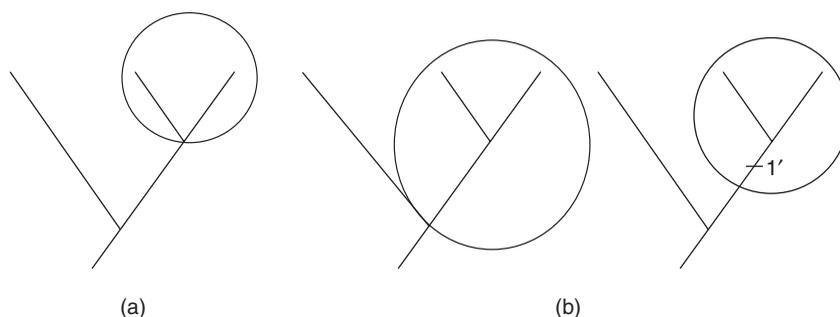
Figure 4.13. Different purported ways of circumscribing monophyly: (a) node-based, (b) stem-based, and (c) apomorphy-based.

three types through reference to a tree (Fig. 4.13a–c). Node-based monophyletic groups supposedly originate at vertices. Stem/branch-based monophyletic groups include both a vertex and the edge below the vertex. Apomorphy-based monophyletic groups are defined by synapomorphies. If the intent of such distinctions is to circumscribe groups in a particular way, they need not always be problematic. For example, a stem/branch-based definition of the teleosts would signal that investigators desired to include all taxa between the branch leading from living *Amia* and relatives to the first living teleosts in Teleostei (Wiley and Johnson, 2010). By contrast, a node-based definition would restrict all taxa below the living teleosts to some group outside Teleostei. Finally, an apomorphy-based definition would restrict the parts of Teleostei to those taxa that have, say, an interoperculum. Taken literally as definitions, however, such assertions imply that there really are three kinds of monophyly, and there really is only one kind (Martin et al., 2010). Monophyletic groups are part–whole systems that include an ancestral species and all descendants of that ancestral species. So, why propose three kinds of definitions? We prefer to have a single concept of monophyly, while instead recognizing that there are different cartographic ways of representing the same genealogical hypothesis. We endorse this strategy because it more fully recognizes the complex aspects of what is a tree, while endorsing the unitary nature of monophyly. Further, it is a more accurate description of nature and also the history of the science of phylogenetics as it dates back to Hennig (1966).

It is worth describing how de Queiroz and Gauthier's (1992, 1994) concepts of monophyly in fact correspond to the tree types elucidated herein. For example, in node-based trees, monophyletic groups are always node-based because vertices/nodes are symbols for ancestral lineages. Likewise, in stem-based trees, monophyletic groups are always stem-based because edges are ancestral species. Adding taxa that subtend an edge simply creates more edges in the graph, and some edge always will be the unsampled ancestor, never a node. A node on a stem-based tree is an event, not a thing, and events are not ancestors. Likewise, in Nelson trees monophyletic groups are always apomorphy-based because apomorphies are used to cluster terminals into what we interpret as a monophyletic group. Further, the known parts of a monophyletic group in a Nelson tree always correspond to the known parts of the same monophyletic group in a stem-based tree and node-based tree. Thus, recognition and diagnosis of a particular monophyletic

group in all three kinds of graphs will be based on apomorphy and they are all apomorphy-based.

The problem with conflating the meaning of trees when speaking of monophyletic groups can be illustrated simply. The edge below a vertex in a node-based tree describes a property relationship that exists with the parent of the node in question. Thus, a stem/branch-based concept applied to a node-based tree will define a group with at least one more known member than that applied to a stem-based definition. In contrast, the edge below a vertex in a stem-based tree points up, to a relationship property shared with children or descendants of the edge in question. Again, this is because stem-based trees are planted by an edge while node-based trees are rooted by a node.

CHAPTER SUMMARY

- Diagrams depicting relationships can take many forms.
- Stem-based trees are diagrams of the actual course of phylogenetic descent as hypothesized by the investigator; branch points are speciation events, and internodes are unsampled (or sampled) ancestral species.
- Node-based trees are directed acyclic graphs with vertices as ancestors and edges as statements of common ancestry.
- Nelson trees are directed acyclic graphs with edges as statements of shared synapomorphies diagnosing sets of taxa.
- There are many terms applied to tree, most of which relate to the empirical method used to make the tree.
- Leaves and edges of a tree can be freely rotated without distortion of the genealogical relationships implied by the tree.
- Node-based, stem-based, and apomorphy-based concepts of monophyly are synonyms, referring to monophyly as it relates to different kinds of tree diagrams.

5

CHARACTERS AND HOMOLOGY

Characters, like gold, are where you find them.

—G. S. Myers

The hypothesis that a character in one organism is the same or different than a character in another organism is at the very base of systematic research: all systematic research depends on our ability to make this choice intelligently. In this chapter, we will first formulate a concept of character as it applies to a single organism. Such characters constitute parts of organisms and properties of organisms. We then develop a concept of a shared character and suggest that they are properties of groups of organisms. These foundations lay the groundwork for a discussion of homology. We begin with some historical considerations. We then move to theoretical matters and review some of the progress that has been made in solving the "problem of homology." We will distinguish between several concepts of homology and recognize two of particular importance to systematists. Taxic homologies are those that are properties of clades while transformational homologies are properties of nested clades. We then turn to empirical matters including empirical tests of homology, strategies of coding characters, and various forms of parsimony. Several parts of this chapter are based on Wiley (2008).

A CONCEPT OF CHARACTER

In the first edition of *Phylogenetics*, Wiley (1981a) expressed the opinion that *character* was a primitive term and he reviewed attempts by Cain and Harrison (1958),

Davis and Heywood (1965), Mayr (1969), and others to define terms such as *character* and *taxonomic character*. We think the problem is much simpler. At the most basic level, when a systematist states that some particular feature of an organism is a character he or she is making a hypothesis that this feature is a property of that organism. For purposes of analysis and comparison with other organisms, the particular form of the character is usually termed the *character state*. This is very similar to the concept of "attribute state" outlined by Jardine (1969) and discussed by Kitching et al. (1998) in their useful compilation of various definitions of characters. The concept of character is most closely related to the concepts of the "anatomical singular" of Riedl (1978) and "quasi-independent parts" of Wagner (1999).

> *A character is a property (feature, expression, part) of an organism that is quasi-independent from other properties of the organism* (modified from Wagner, 1999).

Use of the term *quasi-independent* rather than *autonomous* acknowledges that the organism is a whole and individual entity and that all characters are ultimately linked by genetic and ontogenetic processes. A quasi-independent property is a property that may evolve independently from other parts such that a modification of the character may appear as an evolutionary novelty during descent without the necessary change of another quasi-independent character. Empirically, evolutionary novelties that appear on different branches of a phylogeny are usually considered quasi-independent so long as the tree hypothesis is not rejected by new data. There is always a question of fully demonstrating autonomy when states of different characters appear on the same branch/edge/node unless genetic or ontogenetic evidence is presented that indicates the freedom of characters to vary independently. However, until such time that nonautonomy is demonstrated, the characters and their states should be treated as autonomous, thus exposing them to further analysis.

Character analysis begins with "factorization" (Wagner and Stadler, 2003). A "correct" factorization results when the systematist "breaks down" the organism to the exact number of quasi-independent parts as those parts exist in nature. This goal is never fully obtainable, so on the empirical level we must assert hypotheses of characters. The expectation is that our hypotheses of characters correspond to characters in the real world, and we seek confirmation of our hypotheses by performing additional tests whenever possible. Subsequent research will determine whether our character hypotheses continue to stand up to new knowledge as we and other investigators revisit the organisms studied.

Systematists are usually interested in heritable characters. Thus, the information flow of interest is the flow from parent to offspring rather than issues related to somatic mutations or ecophenotypic variation. Heritable characters arise through the interaction of information flow, matter, various processes such as ontogeny (e.g., morphological characters), replication (e.g., DNA sequences), post-transcriptional processing (e.g., proteins), or the interactions of other characters that are, themselves, inherited (e.g., behavioral characters). In addition, systematists are usually interested in characters whose states show relatively little variation within species but some level of variation between species. (Minimally, we are most interested when two species share a character state not shared with a third species.) So, we are most interested in changes in the information that occur over relatively long time scales (scales usually measured over millennia and in terms of speciation events).

It is convenient to think that any single organism has characters. If two specimens have the same character, we indicate that similarity by assigning a name or code to that particular character that indicates that both specimens have the same character. This name or code is usually termed the *state* of that character. Specimens with a different form of the same character are assigned different states. This concept of character and character state corresponds to Hennig's (1966) concept of the transformation series (=character) and character (=character state). Empirically, this is manifested in a data matrix where (usually) rows are specimens/taxa and columns are characters/transformation series and cells are character states/characters. States may be alternate forms of a character or even the absence of a character. Thought of in this manner, character is a property actually instantiated by an organism and state is the interpretation of that property by a systematist, rendering the difference between character and character state as the difference between a property and how we interpret that property. Data matrices are filled with the properties we observe in specimens. Columns are filled with related properties (ones we have reason to think are homologous). Given this, we offer a formal definition of character state.

> *A character state is an interpretation of the character of an organism that is used to compare the character of that organism with another organism.*

The concept of character as a property of a single organism and its states as the manifestation of that property as interpreted by an observer differs from some other treatments. For example, many systematists define *character* as a feature of one organism that differentiates it from another organism (e.g., Sokal and Michener, 1958:1410; Mayr and Ashlock, 1991:160). Ashlock (1985) attempted to "restore" the term *character* to its original meaning by substituting the term *signifier* for *character state*. Wiley et al. (1991) attempted to preserve Hennig's meaning of character as that attribute that fills a data cell, thus using the terms *transformation series* for columns and *characters* for cells.

CHARACTER STATES AS PROPERTIES

Character states are usually defined as attributes or features of entities and thus seem intuitively to be properties of the entities. Indeed, this is exactly how most systematists have always viewed characters. However, the implication of character states being properties has not been fully explored. Bealer (1999) provides an accessible discussion. In set theory, to say that a rose bush has the character state "xylem present" is to say that the rose bush is an instance of xylem and that xylem is predictable given a rose bush. But xylem is an intentional property of the rose bush. Xylem has the quality of intentionality because distinct properties (the quasi-independent ones) can be truly predicated of the same entity. The property of having xylem does not equal the property of having phloem, although our rose bush may have both xylem and phloem. This makes properties different from sets. Sets satisfy the principle of extensionality. Two sets are identical if they have the same extensions. Xylem and phloem cannot be considered sets simply because, if they were, they would be the same character because they have the same extensions (plants

with both xylem and phloem). Although this may seem a philosophical nicety, it is relevant given that there is no true set of character states of an organism as character states are properties, defined by intention, not sets defined by extension. Thus, character states are parts of organisms and organisms are not the sum of some set of properties. They are individuals that can partake in relationships. The relationships of two characters of the same organism are based on DNA replication, growth, ontogeny, mitosis, common history, or other processes. The relationships of two qualitatively identical character states in two entities may have other kinds of relationships, as discussed below. Does this mean that there is no such thing as a "set of characters"? No, in the empirical world whatever characters an investigator is examining is a set of characters, defined by extension by the very acts of selecting and examining said characters. But this does not mean that the properties of an organism comprise a set of characters, only that investigators will study a set of characters.

SHARED CHARACTER STATES

Two or more organisms share the same character state when this character property is truly and simultaneously instantiated in both organisms at the same time. For example, the property of having a neurocranium is supposed to be simultaneously instantiated in a lamprey and a shark. This amounts to asserting an identity relationship between the parts of two organisms. Identity is a murky philosophical concept, but philosophers generally recognize two kinds of identity relationships (S. J. Wagner, 1999), numerical identity and qualitative identity.

Numerical identity statements assert that two things are the same thing. For example, Clark Kent is Superman, or the child in a family picture is the adult of today. Clark Kent and Superman always have the same space coordinates at any one time. But, Lois Lane sees only one manifestation of the superhero at any one time.

Numerical identity statements are made when an investigator claims that a particular morphological feature has changed over time, as in the case of ontogenetic transformation. Statements such as "the limb bud is identical to the leg" would be typical of such numerical identity statements. Serial homologs, discussed below, would not: the nephridia of the 10th segment of an annelid worm does not share a numerical identity with the nephridia of the 12th segment as each has a different space coordinate in spite of the fact that they might share time coordinates.

Qualitative identity statements (hypotheses of shared properties) assert that two characters/objects/entities share the same intrinsic properties even though they occupy different time or space coordinates. Qualitative identity can take the form of exact similarity where two character properties are exact copies of each other in terms of their own intrinsic properties. Qualitative identity is rarely asserted, except in molecular data where bases and certain amino acids may be said to be qualitatively identical. In morphology, individual variation and complexity preclude assertions of qualitative identity.

The common concept, and the most useful one in morphology, is the concept of relative qualitative identity; two characters (or objects, or entities) are identical if they agree on some list of intrinsic properties, fixed for a given context. Their agreement in having these intrinsic properties can cause us to (1) give them the same

name or code and (2) seek explanation for why they both have this set of intrinsic properties. There are at least three potential reasons for such an agreement.

1. A direct historical relationship. One neurocranium is the ancestor of the other, or both are descended from an ancestral neurocranium. This kind of relationship can be immediately rejected because neurocrania are not replicators (*sensu* Hull, 1981) and, thus, cannot have either ancestor or descendant relationships. Neurocrania, like all character properties, are made anew each generation.
2. An indirect historical relationship. Individual organisms with neurocrania replicate, giving rise to new organisms that develop neurocrania using information passed from one generation to the next. Thus, one organism is the ancestor of the other, or both are descended from an ancestor that had a neurocranium and the information specifying the development of the neurocranium has not changed significantly enough to cause the systematist to call the structure anything except a neurocranium.
3. An indirect ahistorical relationship. Individuals having a neurocranium are members of a kind defined by the property of having a neurocranium. This might be explicable by the independent acquisition of what was mistakenly identified by the systematist as the same structure; it is also conceivable that certain evolutionarily independent lineages might share a common developmental system that allowed the same type of structure, a neurocranium, to form.

The choice between two and three is the choice between character properties being bound by history and being free of history, the same individuals versus kinds controversy we met in a different form in the species discussion. It is not a dichotomy between meaningful properties and nonsense, and we shall distinguish between choices two and three by referring to historical character properties and kind character properties, respectively.

HISTORICAL CHARACTER STATES AS PROPERTIES

Historical character state properties are contingent, spatiotemporally restricted properties that have a relationship among themselves or part–whole relationships with larger individuals. The relationship is relative to a reference point, the entity with which the character property shares the relationship. The neurocranium of a particular shark has a part–whole relationship with the shark. Two cells within the cartilaginous matrix of the neurocranium have a relationship of being cells in the matrix and a part–whole relationship with the neurocranium. But the two cells also have a part–whole relationship with the shark. This, establishing the frame of reference, is essential to describing the relationship. The frame of reference can be described as the reference point; to what whole do the parts belong in the part–whole relationship. The reference point can be any real (or hypothesized real) individual. For example, it is true (but not very useful for systematic purposes) that the two cells in the neurocranium are also "parts of the universe," as well as parts of the neurocranium of the shark.

Individuals can partake of part–whole relationships. *Rana pipiens* and *Homo sapiens* are two parts of Tetrapoda. Individuals of each species have right femurs. Right femur, as we shall argue in the homology sections below, is a property of Tetrapoda instantiated in constituent parts (individual adults) of Tetrapoda, specifically adult individuals of *Rana pipiens* and *Homo sapiens* (and most other tetrapods) as character states with the same code. These species instantiate the property of having a right femur because they are spatiotemporally connected by reproduction and development, and the information that flows through these processes has not changed within the limits we place on the qualitative identity conjecture that defines the intrinsic properties of right femur. Thus we score the character "right femur" with the state "1."

The historical character properties of having femurs, neurocrania, or xylem are examples of qualitatively identical and spatiotemporally restricted states in which the part–whole relationship is expressed through the entities that have the properties and not through the properties themselves. The properties are indicators of deeper relationships or deeper processes that demand explanation; they do not provide explanations (Hennig, 1966; Roth, 1994). Or in a paraphrase of Linnaeus, the genus gives the characters, the characters do not give the genus.

AHISTORICAL KIND PROPERTIES

Historical character properties are quite common; every entity from a particular quark to a particular galaxy has them. However, most talk in science and normal conversation is not talk about entities and their properties; it is instead talk about kinds, properties of kinds, and members of kinds and how these instantiate the properties enumerated. Kinds, as reference points, are not entities; they are concepts defined by kind properties. The relationship between an individual and a kind is not a part–whole relationship, but a class–member relationship. The properties of kinds are not spatiotemporally limited. Every individual, regardless of its history, that instantiates the property of a particular kind is a member of that kind. Further, an entity can be a member of many kinds at the same time, even when these kinds overlap.

Just as kinds come in two varieties (Chapter 2), so do kind properties. The properties of nominal kinds allow us to distinguish between a motorcycle and a bicycle. However, these properties are not predicted by scientific process theories although process theories may explain how they operate and even lead to new inventions to improve the member of that nominal kind. In contrast, the properties of natural kinds are supposed to be predicted by process theories. A theory is proposed that is meant to explain the behavior of individuals. Part of the theory asserts that individuals have certain properties and that these individuals will behave in some predictable manner under certain conditions. For example, individual atoms that are members of the natural kind helium have a certain number of protons and are predicted to be relatively unreactive at Earth-like temperatures and pressures, because their electron orbits are full.

Natural kinds are fundamental to science. It is through studying the patterns of actions, reactions, origins, and histories of groups of individuals that natural kinds are discovered. The behavior of individuals as members of natural kinds, relative to

the properties predicted by the covering theory, is the way process theories are tested. If the behavior of the individuals is that predicted by the theory, then the theory is confirmed; if not, the theory is disconfirmed. Although each individual (for example, atom) has a history and historical properties (origin, spatiotemporal location, etc.), this history is not of particular scientific interest when scientists are investigating the properties of particular kinds. Members of the kind "uranium" are expected to behave in certain ways given what we think we know of the properties of uranium and the processes of atomic physics and chemistry. Given that even small quantities of the substance uranium consist of astronomical numbers of atoms of uranium and the behavior of each of these individual atoms is simple and not affected by historical origin or spatiotemporal location, it is typically not necessary to study the behavior of individual atoms in a sample of uranium. Instead, they can be statistically summed and lumped together. Because the properties are both necessary and sufficient, the theories are open to rigorous tests, given specified conditions. If individuals do not behave according to the properties of the kind to which they belong as members then the theory might be rejected or modified to accommodate the "anomalous" behavior (anomalous relative to the theory, not to nature). For example, if the decay of a lump of uranium does not result in the origin of helium atoms as predicted, the aspects of the theory of atomic chemistry are suspect or something went wrong with the detection equipment.

HISTORICAL GROUPS AND NATURAL KINDS

Historical groups are similar to natural kinds in several important ways, although they do differ. They are similar because both are composed of constituents (parts or members, as appropriate) that share properties. Both are important to science because they are not arbitrary like nominal kinds. Historical groups function significantly in science because they are the result of the operation of natural processes on their parts. Natural kinds function significantly in science when they can predict or explain how individuals will behave while undergoing natural processes (if, of course, the theory behind the kind is sound). They differ because the character properties of natural kinds are ahistorical, necessary, and sufficient while the character properties of historical groups are historical, not necessary, and not sufficient. Consider the kind "hydrogen" and the group "Angiospermae." Membership in the kind "hydrogen" is granted to any atom that has one and only one proton. No hydrogen atom has an ancestor that is, itself, a hydrogen atom, and hydrogen atoms form all the time, in different manners, and have independent origins. The necessary and sufficient property is achieved by convergence. (Thus, systematists would call a group of hydrogen atoms a polyphyletic group.) Membership in Angiospermae is granted to any and all species that have descended from the original angiosperm ancestral species. Any character property shared by this ancestor and its descendants is one historical property of Angiospermae by descent. But these characters are neither necessary nor sufficient. Part of Angiospermae may lose some of these characters and modify others. The observation that this particular plant has flowers during the spring may be considered sufficient to place the plant in Angiospermae (given no convergence), but it is not necessary (Wiley, 1981a) because it is still an angiosperm when it is not flowering. Those plants that instantiate flowers "achieve"

this property in the opposite manner to the hydrogen atoms; they achieve it by common descent from an ancestor who also instantiated the property. (Systematists would call the group monophyletic.) Sciences where historical groups matter, such as phylogenetics, usually have as their subject complex individuals that behave differently depending on their historical origins and spatiotemporal position.

The final difference between historical groups and natural kinds is how they function in theory. Successful scientific theories correctly predict the kinds and properties of those kinds within their domains. Failure to find individual members of the predicted kind under the specified circumstances leads to the rejection or modification of the theory. Failure of individual members to respond in ways that are predicted by the theory, given their properties, can lead to rejection or modification of the process theory. Individuals are important because they are needed to test the theory, but particular individuals are not important. Any individual that belongs to the kind will suffice for the test. Macroevolutionary theory predicts that we should find monophyletic groups when we perform character analysis. But macroevolutionary theory does not predict that we have to find any particular monophyletic group, like Angiospermae. Discovery of the existence of monophyletic groups, in general, acts as confirmation of macroevolutionary theory, but the rejection of any particular group's monophyly does not reject the theory.

HOMOLOGY

Owen (1843) is generally credited with coining the word *homolog*. But the idea that parts of organisms are comparable in some fundamental sense can be traced back at least to Aristotle (Panchen, 1994) and the general notion of comparability of parts must be a basic part of human language (Patterson, 1988). Owen used *homolog* to denote the comparative similarity in structure between parts of two organisms "under every variety of form and function." For example, the right forelimb of a bird would be considered homologous with the right forelimb of a human in spite of differences in function and considerable differences of form. Thus, homologs met the criteria expected of structural or positional characters. This certainly did not preclude them from also meeting the criteria of functional characters. Homology can be contrasted with the term *analogy*. Analogous characters denoted similar function without necessary underlying similarity (wings of birds and butterflies). Analogous characters met the criteria of functional characters but failed in various ways to meet criteria associated with structural characters. Although many consider these words as having opposite meaning, this was not the original intent (Panchen, 1994). Homologous parts can have analogous functions (the radius of the wings of birds and the radius of the wings of bats), just as nonhomologous parts can have analogous functions (wings of bumble bees and wings of birds). Today, most biologists use the term *homolog* to denote comparable (similar or identical) characters shared through common descent and *analog* to denote characters that perform similar functions but have very different morphologies.

Laubichler (2000) provides an interesting account of some aspects of the development of the homology concept. He suggested that twentieth-century concepts fall into three categories; idealistic, historical, and causal-analytical. The idealistic concept is traced directly to Owen (1843) and is based on comparison relative to

an idealized archetype. General homologs are linked by the archetype. The Darwinian revolution replaced the archetype with the common ancestor, but this left the question of the criteria for identifying homologs open and evolutionary morphologists relied upon development to link the homologs of different species. As Striedter and Northcutt (1989), Hall (1992), Laubichler (2000), and others have pointed out, the flexibility of developmental programs effectively barred the criterion of development from the same embryonic anlage as providing a *general* criterion for identifying homologous characters (although it certainly works well in circumstances such as Haeckelian recapitulation). This led Spemann (1915) to consider Lankester's (1870) recognition of the difference between homology and homoplasy. Lankester asserted that homologs could be traced directly to a common ancestor while homoplasies could not. Thus, Lankester argued for the historical concept of homology. Spemann (1915) sought a causal-analytical criterion for homology within an experimentalist framework. Remane (1952) provided one component of the integration of the historical and causal-analytical approaches. He outlined criteria that would lead to hypotheses of homology that could be defended on probabilistic grounds. We shall review these in a later section of this chapter. About the same time, Hennig (1950) linked the historical concept of homology with a specific research program designed to reconstruct phylogenetic relationships. The combination of Remane's (1952) criteria and Hennig's (1950, 1966) methods has led to the homology concepts used in this book. It has not, however, totally solved the problem of accounting for homologs in a mechanistic manner (see G. P. Wagner, 1999). We begin our general discussion of homology with an overview of Haszprunar (1992). This will provide a background for subsequent discussion that will be carried to more detailed consideration of homology in systematics.

Haszprunar's Homology Synthesis

Haszprunar (1992) sought to tie together various concepts of homology in order to understand their relationships. He recognized four types of homology, each associated with a different level of biological organization.

    Iterative Homology. Iterative homology encompasses the concepts of serial homology and homonomy. Iterative homologs are comparable parts of the same individual organism at the same time of life. Hypotheses of iterative homology amount to qualitative identity statements of two or more characters simultaneously instantiated in the same organism. In metameric organisms, iterative homology relates directly to the comparability of parts of different segments, as in the iterative homology that obtains between parapods of different segments of polychaete worms (Fig. 5.1a) or insect antennae and insect legs. In modular organisms, it relates directly to a comparable part of different modules, as in the leaves of vascular plants (Fig. 5.1b) or flower sepals and leaves. It may also be related to other types of repeated parts such as the tentacles of octopi and corals or the hairs of mammals. The relationship among such structures is termed *homonomous*. Haszprunar suggests that the importance of this level of homology to phylogenetics lies in the ability of the investigator to individuate, and we would add factorize, the organism in order to compare its parts among organisms. Iterative homology is relative qualitative identity within the same organism.
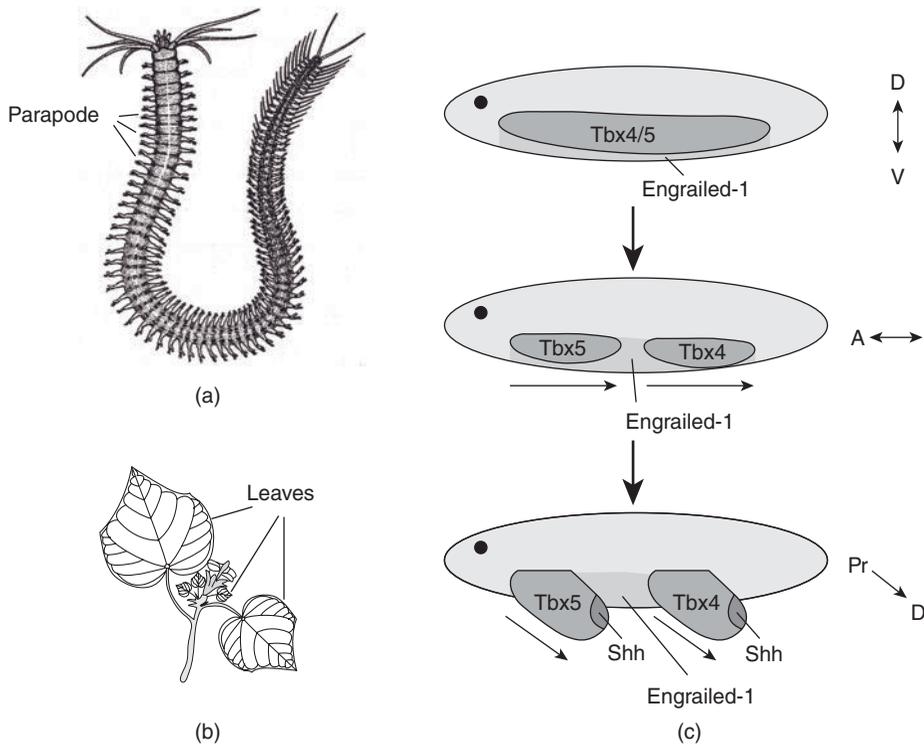
Figure 5.1. Iterative and ontogenetic homology. Parapods of different segment (a) and leaves on a vascular plant (b) illustrate iterative homology (homonomy). (c) Ontogenetic homology: Tanaka et al. (2002) present a hypothesis of ontogenetic homology between the undifferentiated body wall cells and the paired appendages of cartilaginous vertebrates (and by inference gnathostomes). Differentiation during development is marked by expression of various signaling proteins. See color insert.

Ontogenetic Homology. Ontogenetic homologs are comparable parts of the same individual organism at different times of development and growth. Hypotheses of ontogenetic homology are assertions of numerical identity. The most direct example would be parts that are coupled in an ontogenetic pathway, such as fin-fold to fin (Fig. 5.1c). Haszprunar uses the example of antennae and the mandibles of a fiddler crab; the antennae of the nauplius larva become the mandible of the adult crab. The use of the concept of ontogenetic homology on the systematic level represents an attempt to study the differentiation and growth of the organism and to provide a basis for comparisons between organisms. Studies that mix larval, juvenile, and adult characters together in phylogenetic analyses or studies that compare relative growth patterns between taxa (studies of neoteny, etc.) use this concept of homology. Some of the most fascinating examples that conjoin elements of ontogenetic and serial homology come from the polychaete worms (Dick, 1998; Bely and Wray, 2001). In some taxa that undergo asexual budding, a tiny protruberance grows off of a single segment. This protruberance ultimately differentiates and develops into an entire individual organism, which later detaches. What it entails when a single segment or part of the organism gives rise to an entire, complete organism

shows how the concepts of ontogenetic and serial homology, and even individuality, can sometimes blend and blur together.

Di- and Polyphyletic Homology. Haszprunar restricts di- and polyphyletic homology to the relationship of characters within a single species, but there is no reason to do so. Given that apomorphies do not arise all at once, we can expect that within a single species apomorphies will appear as polymorphisms and their dynamics will be explicable by standard population genetic forces until fixed. The relationship is that between apomorphy and plesiomorphy as they vary within a single lineage.

Supraspecific Homology. Supraspecific homologs are properties of species and clades and are frequently termed phylogenetic homologies (e.g., Ax, 1987). Although Haszprunar (1992) did not specifically discuss different kinds of supraspecific homologs, we recognize two: taxic and transformational homologs.

Taxic homologs are relative qualitative identities of properties where individual organisms share a unique history of descent without modification. Taxic homologs are apomorphies at one particular place in the history of descent and shared taxic homologs are synapomorphies that diagnose clades.

Transformational homologs are different apomorphies (different character states) that diagnose two clades, one of which is nested within the other (Fig. 5.2a; see also Patterson, 1982). Transformational homologs are arrayed in transformation series as different character states of a character.

Concepts of Homology in Systematics

Wiley (1975) and Patterson (1982) recognized four classes of homology concepts that have been used in systematics: classical, evolutionary, phenetic, and phylogenetic. Classical concepts trace to Owen (1849). Stripped of archetypes and essentialism, classic concepts are concerned with identity and comparability. Homologous characters, at least as a first guess, are the "same" character in different organisms (Owen, 1849; Nelson, 1989), or anatomical singulars (Riedl, 1978). Evolutionary homology concepts are usually traced to Darwin (1859). Homologous characters are those properties whose identities can be ascribed to common ancestry (e.g., Simpson, 1961; Hennig, 1966; Mayr, 1969; Bock, 1969; and Wiley, 1981a, all give slightly different versions). However, subsequent to Darwin (1859), an important emphasis in evolutionary homology statements was placed on the role that evolutionary processes, in particular, natural selection, might play in causing such "homologous" structures to appear. Here, the case became clear whereby homoplasious or convergent characters might be called *homologs*, which is a major problem in phylogenetics (Patterson, 1982; see for examples). Phenetic homology, or "operational" homology, concepts trace to the phenetic movement of the mid-twentieth century (Sokal and Sneath, 1963; Sneath and Sokal, 1973). Characterizations ranged from very algorithmic (Key, 1967) to very classical (Sneath and Sokal, 1973). Phenetic homologs fulfill the criteria of structural and compositional correspondence. The motivation for phenetic concepts was to escape circular reasoning: in particular, the circularity of viewing homologs as based on common ancestry but concluding that evidence for common ancestry is derived from the sharing of homologs. Hennig (1953) and Ghiselin (1966) effectively dismissed the problem of circularity, while Hull (1967) pointed out that no hypothesis is subject to direct proof, which means to us that operational approaches are no less subject to assumptions than

evolutionary approaches to homology and there is no special reason to dignify a phenetic homology concept.

Phylogenetic concepts (cladistic concepts of Patterson, 1982) of homology are based on discovering monophyletic groups. Hennig expressed this concept in terms of transformation series (Hennig, 1966:93) and Wiley (1975) in terms of ancestry. Workers such as Bock (1969) and Hecht and Edwards (1977) also tied homology directly to monophyly, but their concept of monophyly was not the phylogenetic concept *sensu* Hennig (1966) and also included para- and even potentially polyphyletic groups. Phylogeneticists, however, treated homology as synapomorphy (Wiley, 1975, 1976; Bonde, 1977; Patterson, 1982; de Pinna, 1991; and others). Patterson (1982:60–61) concluded that homology is the relationship characterizing monophyletic groups and that synapomorphies were the *only* properties of monophyletic groups. We will take a slightly different view, but agree with Patterson that synapomorphies are *one* of the properties of monophyletic groups, unique ancestry being the other principle property. We embrace the original concept of Hennig (e.g., 1966) that the principle property of monophyletic groups is unique ancestry relationships among members of the group and that the evidence for such genealogical relationships resides in character properties. With our revised view, we hope to reach a reconciliation between the concept that homology is *only* synapomorphy and the criticisms of that assertion (e.g., McKitrick, 1994; Ghiselin, 2005).

Phylogenetic Characters and Phylogenetic Homology: An Overview

At the very base of the phylogenetic system is what Hennig (1966) termed his "auxiliary principle." The principle can be stated simply: characters meeting various similarity criteria (relative qualitative identity) are to be considered homologous unless evidence is presented that demonstrates that they are homoplasious. This principle is important in two respects. First, it offers the hope that additional tests might be performed to confirm or disconfirm the assertion. Second, the auxiliary principle is actually needed to proceed in systematics because without it we could dismiss any similarities as evidence of relationship without evidence and arbitrarily; then systematics becomes a matter of opinion and no longer evidence-based and scientific. The auxiliary principle is just a restatement of a basic principle applied to all science, the parsimony principle. If the characters of two specimens can be explained by a single origin, then invoking two origins requires adoption of an ad hoc assumption or some additional knowledge about the evolution of the characters in question that is not inherent in their similarity.

In subsequent sections, we will make the point, following Hennig (1966:95), Wiley (1975), and Patterson (1982), that all true taxic homologies (homologs that share an identity) are synapomorphies at some level in the tree of life and that plesiomorphic homologies are simply apomorphic homologies at another (more inclusive) level in the tree. We shall also see that taxic homologs are not the only kind of homolog we meet in restricted parts of the tree of life. The homologous relationship between preexisting characters and their modifications (i.e., between plesiomorphies and their homologous apomorphies) can be described as transformational, and thus plesiomorphies and apomorphies form transformational homologs. If we always considered the entire tree, we could reformulate that statement: basal apomorphies and apical apomorphies form transformational homologs at different levels of the hierarchy of descent.

Taxic Homologies as Properties of Monophyletic Groups

All true taxic homologies are character properties of one or more true monophyletic groups. The problem is, in phylogenetics we have no entrée to absolute truth. So empirically, when we assert that a particular similarity observed in two or more organisms is homologous, we are asserting a hypothesis that the organisms belong to a monophyletic group whose properties include the character state in question.

Because we cannot discover true homologies, each of our statements, from the characters of individual specimens to assertions as to the homologies of clades, are hypotheses subject to testing by one or more methods (McKitrick, 1994, and others). Following Hennig (1966:94), De Pinna (1991), McKitrick (1994), and others, we believe it is useful to distinguish between "true homology" and hypotheses of homology. In this section, we will suggest that taxic homologies are synapomorphies and transformational homologies are nested hypotheses of plesiomorphy and apomorphy, and these are the two kinds of homology of primary interest to systematists. We begin with definitions of phylogenetic homology, taxic homology, transformational homology, and phylogenetic homoplasy.

Phylogenetic Homology. The character states of two organisms are phylogenetically homologous if they are either taxic homologs or transformational homologs, as defined below.

Taxic Homology. A character state shared by two or more species or clades is homologous if the shared character state is a diagnostic property of a monophyletic group to which both taxa belong. This corresponds to Haszprunar's (1992) taxic homology. The inference is that the ancestor of the clade had the character state when it speciated to give rise to its descendants.

Transformational Homology. Different character states are homologous if one is the direct historical antecedent of the other and if the species or clades having the properties are nested as monophyletic groups. The inference is that the information specifying the antecedent character state has been modified to specify the derived character state.

Phylogenetic Homoplasy. Theoretically, homoplasies are character states that share a relative qualitative identity but diagnose polyphyletic groups. The clearest examples are character states such as base residues that occupy the same position in a gene, but appear twice or more on a tree. Such character states are members of a kind (thymine, for example), but homoplasious. Other character states initially thought to share an identity (morphological features coded as the same state) may turn out to not share an identity when we look at them in detail (see sections on Remane's criteria later in this chapter). Such character states are coding "mistakes," and the character states in question are not members of the same kind. (Note that true homoplasies are members of the same kind.) Homoplasies may separately diagnose two or more monophyletic groups. Thus, some homoplasies, taken together, are homoplasies; but taken separately, each may be independent taxic homologies of the monophyletic groups with which they are associated as a diagnostic property.

The empirical relationships between concepts and data are shown in Fig. 5.2. Consider the tree (Fig. 5.2a) to be the empirical result of a phylogenetic analysis with the synapomorphies from the matrix (Fig. 5.2b) plotted on it. The monophyly of the clade A–F is supported by the character states 1 and 2. They form a transformational homolog with their plesiomorphic character states -1 and -2, respectively. They also form a transformational homolog with character states 1+ and
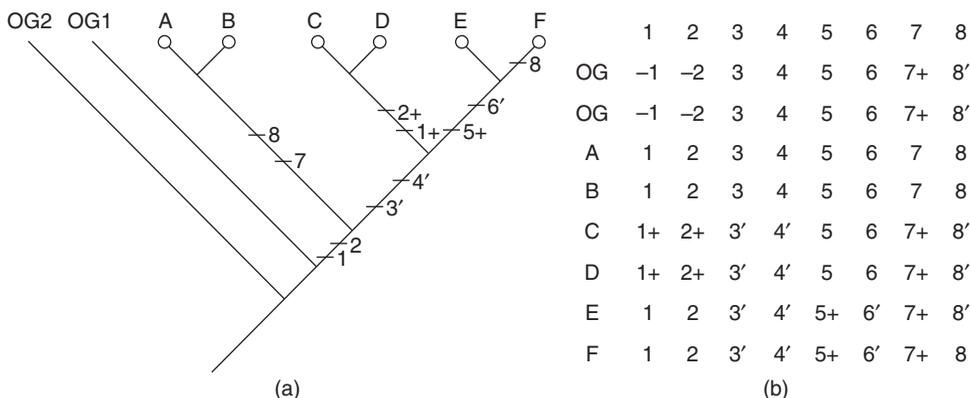
	1	2	3	4	5	6	7	8
OG	−1	−2	3	4	5	6	7+	8′
OG	−1	−2	3	4	5	6	7+	8′
A	1	2	3	4	5	6	7	8
B	1	2	3	4	5	6	7	8
C	1+	2+	3′	4′	5	6	7+	8′
D	1+	2+	3′	4′	5	6	7+	8′
E	1	2	3′	4′	5+	6′	7+	8′
F	1	2	3′	4′	5+	6′	7+	8

Figure 5.2. Concepts and data. (a) A tree based on characters presented in (b) the matrix. Relationships of the ourgroups (OG1, OG2) are assumed true based on prior empirical evidence.

2+, which form taxic homologs for the clade CD. The entire transformation series for the first data column consists of the transformational homology hypothesis -1→1→1+. Character state 8' forms a transformational homology with 8. Character state 8 is a taxic homology of the clade AB and a separate taxic homology of taxon F. Thus, 8 shows levels of both homology (shared with A and B) and homoplasy (the group ABF would be polyphyletic if grouped by 8).

Most authors describe the relationships of homologs in terms of descent. One of a homologous pair of characters is said to be "ancestral" to the other. This is the manner in which Wiley (1981a) described the relationships among transformational homologs. However, we must realize that "ancestral" in this context is a metaphor because characters are not replicators. As we noted above, characters do not give rise to other characters. (Sattler, 1984, 1994, also notes this, but draws different consequences than those drawn here. See the end section of Chapter 6.) Rather, organisms give rise to other organisms and characters are reconstituted each generation anew from the information passed on from the previous generation. This information may be genetic, cytogenetic, epigenetic, or behavioral; the only requirement is that it is heritable. This is why workers such as Osche (1973, 1982), Van Valen (1982), Brooks and Wiley (1986), Haszprunar (1992:15), and Roth (1994) embraced a concept of homology as a manifestation of the flow of information across generations. And this is why we refer to the relationship of a plesiomorphic character as antecedent rather than ancestral relative to a derived homolog when discussing transformational homologs.

> Homologies are similarities of complex structures or patterns, which are caused by a continuity of biological information (in the sense of Riedl's and Hazsprunar's "instruction").

Of course, there is no need to restrict the concept to complex structures or patterns of morphology, but in every manifestation, homology relates to continuity of information, genetic and epigenetic. No continuum of information flow means that no

homology relationship exists, as is the case with all homoplasies. Haszprunar's definition is largely synonymous with VanValen's (1982). As Roth (1994:306) put it: "Information is that which is conveyed, directly or indirectly, by replication between generations. …"

Although highly theoretical, the concept provides a link to empirical systematics by replacing metaphorical talk about homology with talk that can be related directly to phylogeny and evolution. Right hands do not give rise to other right hands, but the reemergence of right hands generation after generation does depend on transmission of instructional information from generation to generation that specifies the ontogeny of a right hand in individual organisms that are parts of their clade. Second, exceptions and evolutionary modifications are explainable; if the instructions are not passed on and processed correctly, the character does not appear. Third, as the material manifestation of information transmission, the phenomenon of homology is not restricted to emergent morphology (including the morphology of gene sequences). For example, there are behavioral synapomorphies and these may be just as valuable in systematics as morphological homologies. For reviews and literature relating to the use of behavioral characters in phylogenetics, see Wenzel, 1992; Brooks and McLennan, 1991, 2002; Rendall and Di Fiore, 2007; note that homoplasy does not necessarily occur at higher frequency in behavioral characters than in morphological characters (McLennan et al., 1988). Similarly, in this view, morphological homologies are no less suited for phylogenetic analysis than DNA sequence homologies. The idea that one kind of data is inherently better than other kinds of data is not viable under this concept, and hypotheses of homology from whatever source can and should be allowed to compete on an even playing field as potential evolutionary innovations (e.g., discussion in Hillis, 1987).

Transformational Homology: Linking Different Hypotheses of Qualitative Identity in a Transformation Series

We cannot escape dealing with transformational homologs for they are inherent in every data matrix we prepare. To claim that any shared character state is a synapomorphy for a group (a taxic homology) requires that it be associated with a preexisting character state in order to polarize the character states, and this creates a transformational hypothesis. Given that there is no spontaneous generation, most taxic homologs have some preexisting progenitor, even if it is undifferentiated from the normal somatic cells of the organism or changed such that the information is no longer expressed ("absence"). Some have only a metaphorical progenitor, but these seem mostly confined to gene duplications and other additions on the genetic level, where the duplicated gene, for example, is truly absent before the duplication event. If true character states are properties of real organisms, true taxic homologs can be considered properties of truly monophyletic groups of organisms. This realist assertion provides the basis for forming empirical hypotheses of taxic homology. But homologs come in two forms; homologs that are similar enough to be called the same character state (they share an identity), and homologs that are different enough to be called by a different name (they do not share an identity but they do share a historical relationship). Consider a relatively unproblematic example. Most tetrapod vertebrates have right legs while most teleost fishes have right pectoral fins. These character states are certainly different in many respects, in fact, in
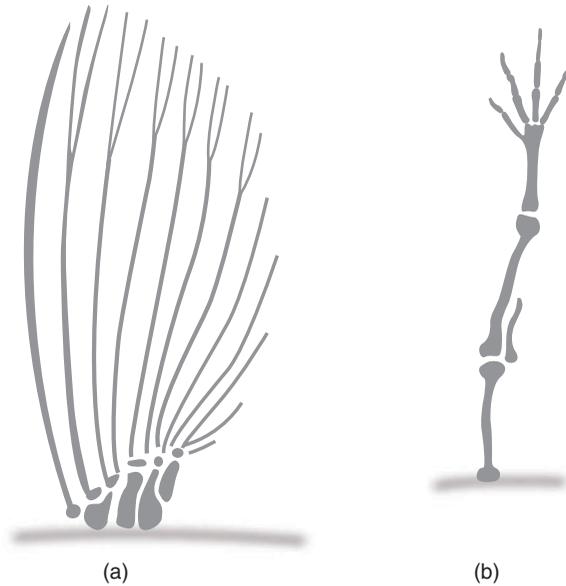
Figure 5.3. The right pectoral fin of a fish (a) and the right leg of a tetrapod (b).

most respects from a strictly morphological viewpoint (Fig. 5.3). Yet, traditionally we treat these as homologous. Why?

Most of us willingly accept the hypothesis that the right legs of all tetrapods are homologous. We observe that right forelegs share certain shared informational properties such as a specific topographic relationship with other body parts and some number of similarly positioned bones, and this leads us to believe that the regular (if not universal) appearance of legs in each generation of tetrapods reflects a common underlying informational system passed on to each generation through reproduction (Riedl's and Hazsprunar's instruction). Further, we interpret the disappearance of legs in such organisms as snakes as modifications of this system and in no way comparable to the lack of legs in jellyfishes (Patterson, 1982). The fins of fishes seem to have certain shared informational properties. Although diverse, we hypothesize that all fishes that have right pectoral fins share the same basic informational system that leads to the appearance of this fin during the growth of the organism. Right pectoral fins are qualitatively identical among fishes. Right legs are qualitatively identical among tetrapods. Why do we think that right fins and right legs can be placed in the same transformation series, that is, why do we think that together they form a hypothesis of transformational homology? These are crucial empirical issues to the rigorous formulation of homology and synapomorphy that we address below.

DISCOVERING AND TESTING HOMOLOGY

A basic principle of phylogenetic realism is that homologies and phylogenies require discovery. A basic principle of phylogenetic empiricism is that discovered homolo-
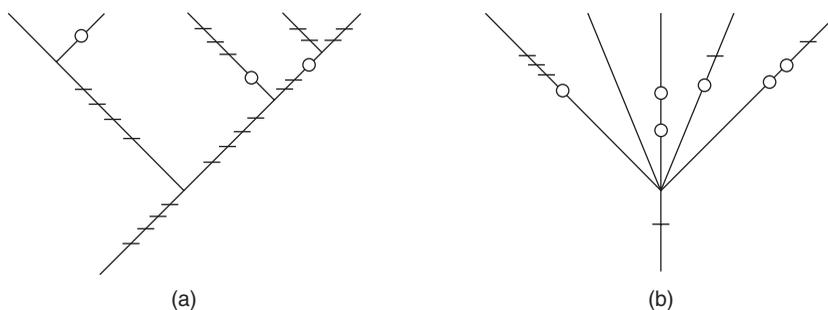
Figure 5.4. Strong and weak phylogenetic signal. The tree in (a) has few homoplasies (circles) and strong signal (dashes) while the tree in (b) has many homoplasies and no signal.

gies are observational hypotheses (Hennig, 1966), not facts, because we have no perfect method of observing real homologies as they exist in nature. Given this, the assertion that two or more organisms share a homology or the assertion that a particular synapomorphy is a character property of a particular monophyletic group are both probabilistic conjectures (Patterson, 1982; Haszprunar, 1998; Sober, 2000) whose veracities are always open to further testing as opposed to deductive conclusions (e.g., Rieppel, 1980). *This can be viewed as a manifestation of the general principle of parsimony and the auxiliary principle.* The degree to which we accept a hypothesis of homology is directly related to the severity of tests applied to the conjecture. Thus, we can separate conjectures of homology into two basic categories. The first is what Sober (1988) calls a "match," de Pinna (1991) calls "primary homology," and Wiley (1975, 1981a) calls "initial hypotheses of homology." The second category comprises those homology hypotheses that have survived phylogenetic analysis and appear, in the end, as homologies. Homology discovery and testing is locked into a dance of confirmation and disconfirmation. Hypotheses are asserted, tested, and evaluated through the process of reciprocal clarification or illumination (Hennig, 1966). This is one of the reasons, in addition to paleontological and ontogenetic information, why we can treat the lack of limbs in snakes as different from the lack of limbs in jellyfish.

The process, from the modern perspective, can be broken into two processes: (1) matrix building and (2) matrix analysis. The end result is a consilience of the quality of the matrix and the quality of the phylogeny. If we have carefully prepared the matrix and the result is a robust phylogeny, the probability that we have made good homology conjectures is increased because our homology conjectures are mutually reinforcing; they are highly covariant, reinforcing each other by displaying a strong historical signal (Fig. 5.4a). In contrast, a weakly supported phylogeny may result from a matrix that contains conflicting or no historical signal (Fig. 5.4b). The matrix may be carefully prepared, but the data do not speak to the problem. Or the matrix may be a sloppy mix of ill-conceived homology conjectures reflecting the quality of the science performed. So matrix preparation is a critical phase in phylogenetic analysis and how we arrive at matches is critical to the outcome. There is also the worst-case scenario; we might encounter data that are positively misleading in that they contain a stronger ahistorical signal than historical signal due to convergence. We will discuss these kinds of data in another section.

Patterson's Tests

Colin Patterson (1982) outlined a theory of how homologies are tested in phyloge-netic analysis. It works regardless of the method of analysis, although it is most transparent in parsimony analysis. There are three basic tests.

1. The test of similarity. Two character states are likely to be homologous if they are similar in detail, develop in the same manner, or occupy the same place in the organisms relative to other character states. These are restatements of Remane's (1952, 1956) criteria, discussed in some detail below.
2. The test of conjunction. If the supposed homologs occur in the same organism, then phylogenetic homology relative to different organisms is refuted.
3. The test of congruence. Character states that pass the tests of similarity and conjunction are likely to be homologous if they are congruent with the histo-ries of other character states that pass the same tests. This is Hennig's (1966) criterion. Homologies are those state properties that pass all three tests while various forms of homoplasy are those that fail one or all tests.

Tests of similarity and conjunction are tests applied during matrix building, and they also can be used to evaluate character state matches after matrix analysis. If we consider each data column in a matrix to be a homology hypothesis, then the job at hand is to fill the columns with robust matches *sensu* Sober (1988): robust primary homology statements *sensu* de Pinna (1991). Only after this phase can we move to the third test, the analysis of the matrix. Below, we discuss each of the tests.

Similarity and Remane's Criteria

Similarity is the primary means of asserting that the same character state is simul-taneously instantiated in two or more organisms as an initial match. It is the major criterion for asserting that a qualitative identity obtains between states of two organ-isms. However, similarity is a complex property and can have several meanings. Asserting qualitative identity usually means that two states agree in their intrinsic properties, an important criterion for assessing qualitative identity. Similarity might also mean that there is some similarity relationship relative to other parts of the organism. Do the characters occupy the same position in the organism relative to other parts, whether they agree in their intrinsic properties or not? Many aspects of the similarity criteria were captured by Remane (1952 and subsequent editions, the second edition, 1956, is most often cited), who outlined three primary qualities of resemblance: (1) similarity in position, (2) similarity in detail, and (3) similarity traced through intermediate forms in development or phylogeny. If the latter, the test partly depends on phylogenetic congruence. For those interested in Remane and the development of his ideas, see the review of Zachos and Hoßfeld (2006).

Similarity in Position: Morphology   The most common use of similarity in both morphological and molecular data is application of the criterion of position. This criterion takes into account the topographic position of a character relative to other characters and to the body as a whole. Two examples illustrate this criterion in morphology.
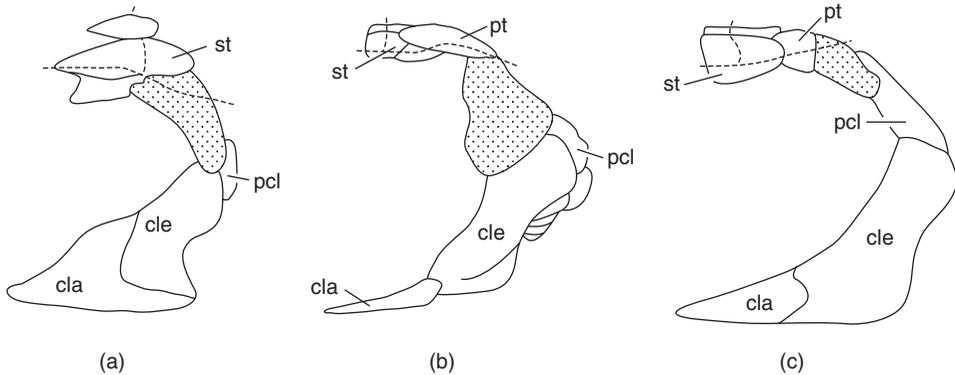
Figure 5.5. Remane's criterion of topographic position I. The dermal shoulder girdles of (a) a sturgeon, (b) a bichir, and (c) an osteolepiform. The supracleithrum (stippled) has a similar topographic position to the cleithrum (cle) and the postcleithrum (pcl) in (a) and (b) and a similar topographic position in relation to the posttemporal (pt) in (b) and (c). The sturgeon differs dorsally in having lost the posttemporal and having an enlarged supratemporal (st) while the osteolepiform differs in having an enlarged postcleithrum that displaces the supracleithrum from its association with the cleithrum. From Wiley (1981a); redrawn and modified from Jollie, 1973.

The supracleithrum of basal osteichthyan fishes is the dermal bone (bone formed in the membrane, not in cartilage) above the cleithrum that carries part of the lateral line canal and articulates with bones of the skull (Fig. 5.5). This relationship is relatively simple in sturgeons and bichirs (basal actinopterygians), but not in *Eusthenopteron* (basal sarcopterygian) where an enlarged postcleithrum displaces the supracleithrum dorsally. Sturgeons and bichirs have three points of topographic similarity and share two of these points with *Eusthenopteron*. The match is fairly straightforward, because all three have a postcleithrum and the displacement of the supracleithrum from its contact with the cleithrum is relatively easy to identify (given that contact is the plesiomorphic state).

The mouth parts of sponging and biting flies illustrates use of the criterion of topographic position in characters that have been modified to perform different functions. Each appendage is accounted for in spite of the fact that the form of each appendage has changed (Fig. 5.6). Thus, although each may not be qualitatively identical, their states can be gathered into a single transformation series.

In our example of fins and tetrapod limbs, the criterion of topographic position is confirmed by showing that although the articulation differs, both fins and legs articulate with an internal support that is ultimately connected to the body axis. That support structure is the shoulder girdle.

Similarity in Position: Molecular Characters   Hillis (1994) provides an extensive discussion of homology at the molecular level. Here we are concerned primarily with homology at the level of DNA and protein sequences. Similarity in position has a special place in such molecular studies because the character states belong to discrete classes that reoccur. For example, there are only 4 kinds of base pairs (ATCG) and 20 common amino acids. There is little possibility of ferreting out homoplasy in the matches based on special similarity (discussed below) because
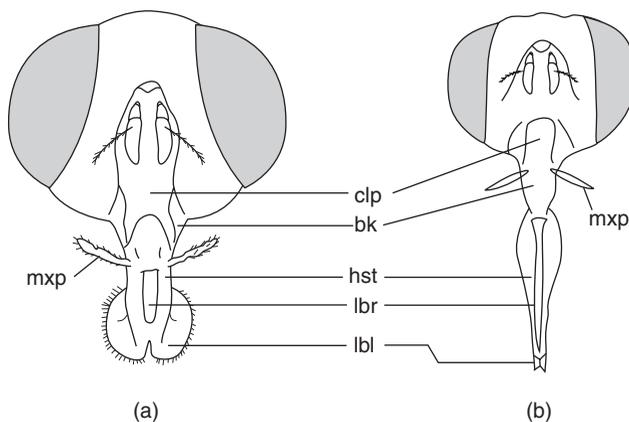
Figure 5.6. Remane's criterion of topographic position II. Frontal views of (a) a sponging fly and (b) a piercing fly. The labeled mouth parts have the same relative topographic positions although they differ in shape and sometimes in function. Abbreviations: bk, rostrum; clp, clypeus; hst, haustellum; lbl, labellum; lbr, labrum; mxp, maxillary palp. From Wiley (1981a); modified from Borror et al., 1976.

there is no difference between, say, thymine residues nor any intermediates between thymine and guanine residues; this is because thymine is a relatively simple object, at least when compared to complex cellular and morphological structures. So for DNA and amino acid sequences, similarity in position within the sequence as shown by the alignment is the primary criterion (e.g., alignments in Fig. 5.7). Hillis et al. (1996:549) define alignment in the following manner:

> Alignment. The juxtaposition of amino acids or nucleotides in homologous molecules to maximize similarity or minimize the number of inferred changes among the sequences. Alignment is used to infer positional homology (qv) prior to or concurrent with phylogenetic analysis. …

Alignment in systematic research is usually carried out on more than two sequences, but the basic principles apply if only two sequences are aligned. The first principle is that any two (or more) sequences can be aligned if a sufficient number of gaps are introduced. The usual strategy (first proposed by Needleman and Wunsch, 1970; see variations and reviews in Doolittle, 1990; Miyamoto and Cracraft, 1991; Hillis et al., 1996; and Phillips et al., 2000; others are discussed below) is to seek an alignment that minimizes gaps and sequence variation. This is a basic parsimony argument.

Gaps. Because any two sequences can be aligned without mismatches if a sufficient number of gaps are introduced, it follows that gaps should be kept to a minimum necessary to align the sequence. The investigator assigns penalties that are based on the number of gaps or the lengths of the gaps. There are a variety of ways to assign gap penalties depending on the nature of the sequence and the position of the gap within the sequence. For example, gaps within protein-coding genes that cause frame shifts might be assessed a greater penalty than gaps within a variable loop region of a ribosomal gene.
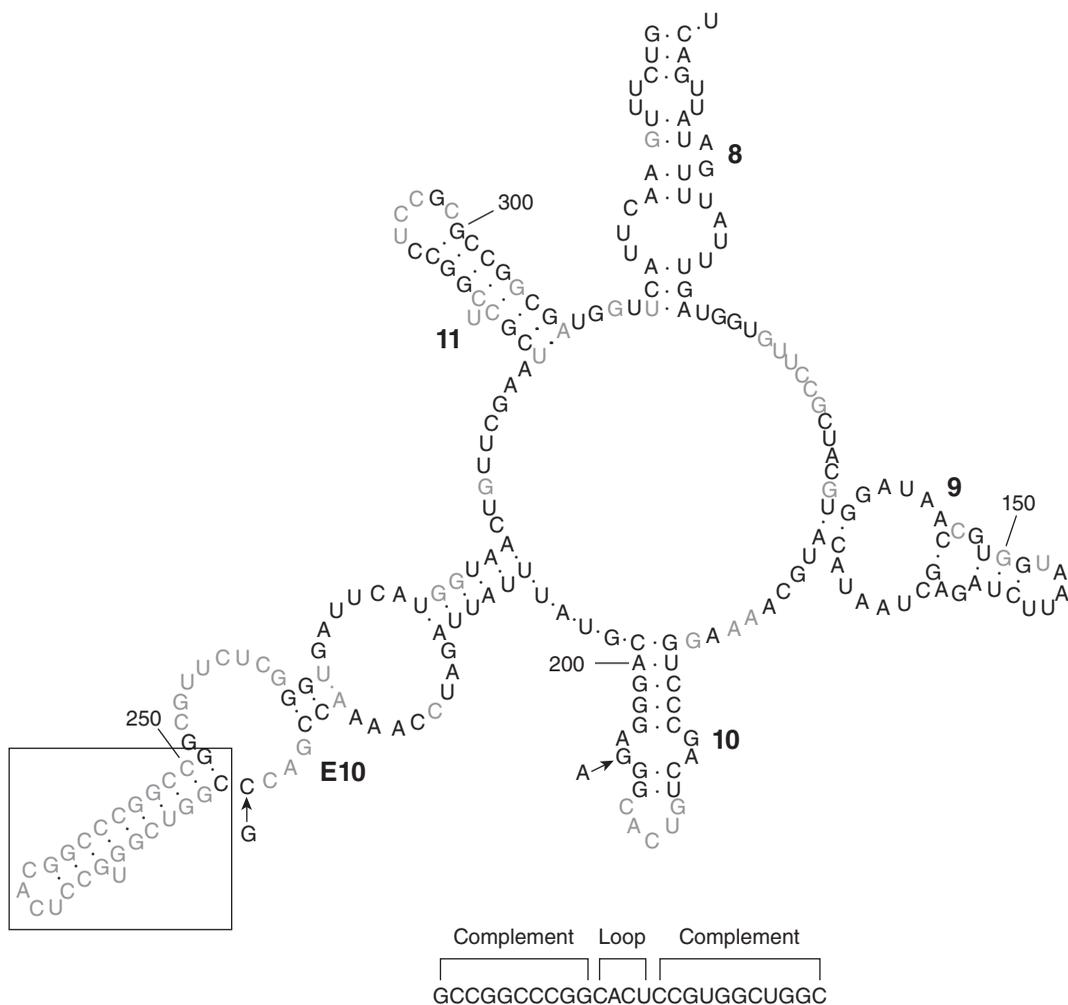
A) Nucleotide sequences

```
1 TGTACCAAGAATGTTATAGTGCTCCCACTAACTTTCTTGTGAACACATGGCTCATCTTCA 60
2 -GTACCAAGAATGTTATAGTGCTCCCACTAACTTTCTTGTAAACACATGGCTCATCTTCA 59
3 -GTACCTAGAATGTTATAGTGCTCCCACTAACTTTCTTGTAAACACATGGCTCATTTTCA 59
4 AGTACCAAGAATGTTATAGTGCTCCCACTAACTTTCTTGTAAACACATGGCTCATCTTCA 60
5 -GAACCAAAAATATGTACTTACTCCCATTGAATTTTTGATACACACAATCATCAACAATA 59
  * *** * *** * * ****** * * *** * * ***** *** *
1 TTTTTGATAAAACCAAACTCTTTGATTGCATCATTAAAACGAAGATTCCAACTTCGAGAA 120
2 TTTTTGATAAATCCAAACTCTTTGATTGCATCATTAAAACGAATATTCCAACTTCGAGAA 119
3 TTTTTGATAAAACCAAACTCTTTGATTGCATCATTAAAACGAAGATTCCAACTTCGAGAA 119
4 TTTTTGATAAAACCAAACTCTTTGATTGCATCATTAAAACGAAGATTCCAACTTCGAGAA 120
5 TTCATCTCAAAACCAAACAAAACAATTATTTAATGAAACTTGTGGTACCATTGACGGGAA 119
  ** * *** ****** *** * ** *** * *** ** ***
1 GCTTGCTTTAATCCATAATTGGATCTTTGTAGCTTACATATCTT--TCCAGCATCCTTTG 178
2 GCTTGCTTTAATCCATAAATGGATCTTTGTAGCTTACATACCTT--TCTAGCATCCTTTG 177
3 GCTTGCTTTAATCCATAAATGGATCTTTGTAGCTTACATATCTT--TCCAGCATCCTTTG 177
4 GCTTGCTTTAATCCATAAATGGATCTTTGTAGCTTACATATCTT--TCCAGCATCCTTT- 177
5 GCTTGTTTGAGCCCATAGATGGATTTTTGTCAATTGCAAACCATATTCTTTGAGTCTTTC 179
  ***** ** * ***** ***** ** * ** ** * * ** * ****
1 G-TTGACAAAACCTTCAAGTTGTGTCATGTACACATCCTCTTCAAGTTTCCCATTAAGGA 237
2 GATTGACAAAACCTTCAAGTTGTGTCATGTACACATCCTCTTCAAGTTTCCCATTAAGGA 237
3 GATTGACAAAACCTTCAGGTTGTGTCATGTACACATCCTCTTCAAGTTTCCCATTAAGGA 237
4 ---TGACAAAACCTTCAGGTTATGTCATGTACACATCCTCTTCAAGTTCCCCATTAAGGA 234
5 GACT--CAAAATTTTCTGGTTGCACCATATAAATTGTTTCTTCAATGTCGTCATTAAGAA 237
  * ***** *** *** *** ** * ******* * ******* *
1 AAGTTGT--- 244
2 AAGCTGT--- 244
3 AAGCTGT--- 244
4 AAGCTGTTTT 244
5 ACACAGT--- 244
  * **
```

B) Amino acid sequences

```
1 YQECYSAPTNFLVNTWLIFIFDKSKLFDCIIKTNIPTSRSLLSINGSLLTYLS-SILWID 59
2 YLECYSAPTNFLVNTWLIFIFDKTKLFDCIIKTKIPTSRSLLSINGSLLTYLS-SILWID 59
3 ---VPRMLCSHLSCEHMAHLHFNQTLLHHNEDSNF--EKLALIHNWIFVAYISFQHPLVD 55
4 ---VPRMLCSHLSCKHMAHLHFNQTLLHHNEDSNF--EKLALIHKWIFVAYISFQHP-FD 54
5 NQKYVLTPIEFLIHTIINNIHLKTKQNNYLMKLVVPLTGSLFEPIDGFCQFANHILVFRL 60
  ..* : :. : . . . . : : : .
1 KTFKLCHVHILFKFPIKESC- 79
2 KTFRLCHVHILFKFPIKESC- 79
3 KTFKLCHVHILFKFPIKESC- 75
4 KTFRLCHVHILFKFPIKESCF 75
5 KIFWLHHINCFFNVVIKKHS- 80
  * * * *:: :*:. **: .
```

Figure 5.7. (a) An aligned DNA sequence of copia-like retrotransposons in clones of Egyptian cotton and relatives. Labels: 1–3, *Gossypium barbadense*; 4, *G. arboretum*; 5, *G. darwinii*. (b) An aligned protein sequence of the same DNA region. From Abdel Ghany and Zaki (2003).

Substitutions. Substitution cost can be specified by the investigator through the use of a cost matrix (Sankoff and Cedergren, 1983). For example, the investigator can specify that a greater cost be assigned for transversions than transitions.

Refinement based on topographic position or models of change. Once an initial alignment has been achieved, the investigator can examine the results in an effort to improve the hypotheses of homology. One example is alignment of ribosomal gene sequences. Ribosomal genes code for ribosomal RNA which, like transfer RNA, has a secondary structure in the mature molecule (Fig. 5.8). Kjer (1995) demonstrated that alignments could be significantly improved by comparing the

Figure 5.8. Alignment of ribosomal sequences as inferred from secondary structures. (Upper) Part of the secondary structure of the sponge *Amphimedum queenslandica* (from Voigt et al., 2008). (Lower) Alignment of part (shown in box) of the sequence in the E10 branch of the mature molecule. Note that complementary bases do not appear together in the sequence.

aligned sequence with a model of the secondary structure that identified regions containing loops (unpaired sequences) and stems (paired sequences). For example, Xia et al. (2003) demonstrated the need for such refinements in analyses of tetrapod phylogeny using the nuclear 18S rRNA gene; results from 18S had previously been used as an example of the disparity between morphological and DNA data. In particular, morphological analysis grouped birds with crocodiles (e.g., Gauthier et al., 1988; Eernisse and Kluge, 1993) while early analysis of the 18S rRNA gene grouped birds with mammals (e.g., Hedges et al., 1990). Xia et al. (2003) found that when the 18S rRNA sequences were refined by adjusting the alignment according to secondary structure (and better sequences were obtained), the results no longer

support the bird–mammal relationship and instead agree with the morphology-based analysis.

Alignments of sequences can be performed in one of two basic ways, a priori alignment and simultaneous (or direct) alignment/tree finding. As named, a priori alignment is a procedure where matches are determined before any phylogenetic analysis is performed. It is the most common form of alignment. There are a variety of programs that perform multiple alignment. A common method is to use the program Clustal (Higgins and Sharp, 1988; Higgins et al., 1992). This approach implements the method of Feng and Doolittle (1987) to order the taxa using clusters of sequence similarity. One can reiterate the process until a stable alignment is obtained (TreeAlign: Hein, 1990, 1994). Other programs include MUSCLE (Edgar, 2004) and MAFFT (see Katah and Toh, 2008, for overview). There are a number of other such alignment programs for both pair-wise and multiple alignment. We note that Sankoff et al. (1973) apparently were the first to publish a formal multiple alignment algorithm (see Sankoff, 2000).

Simultaneous alignment refers to aligning sequences and tree reconstructions simultaneously, as is performed in the program POY (Varon et al., 2010; Wheeler, 1996, 2001, 2003a, b). POY begins with a number of trees, each associated with a particular alignment. Each tree is refined according to the optimality criterion adopted (parsimony or likelihood) by branch swapping and other tree manipulations, and the new tree is evaluated by producing an alignment inferred from the tree. The alignment is termed an implied homology. The process is reinterated until both the tree structure and the associated alignment is optimized. Simultaneous analysis and alignment is controversial. See Wheeler (2006) for a full account of the method and Ogden and Rosenberg (2007b) for comparison of performance of POY and a priori alignment.

Special or Intrinsic Similarity   If the characters of two or more specimens are good matches using positional criteria, we should expect them to be similar in their details when we examine them more closely. If this expectation is not met, then we may have cause to reject the homology conjecture. Equally important, after we apply the test of congruence, we may find that parts that agree in gross topological similarity may differ in detail, allowing us to dismiss the characters as homoplasies that are properties of polyphyletic groups or kinds. Two examples illustrate the use of this criterion to dismiss matches.

Gegenbaur (1873) proposed that the endoskeletal shoulder girdle of gnathostome vertebrates was the serial homolog of gill arches. This is a reasonable assertion of serial homology given the topographic positions of these elements (Fig. 5.9) and Gegenbaur's hypothetical derivation of the girdle from a prototypical arch. (Gegenbaur was one of the great nineteenth-century comparative anatomists.) However, when we look at the details, Gegenbaur's proposal falls apart. In fact, the endoskeletal (cartilaginous) part of the shoulder girdle is derived from lateral mesoderm while the gill arches are derived from neural crest cells, indicating that they have quite different embryological origins and thus are not good candidate serial homologs (see Balinsky, 1970, for embryology and Zangrel and Case, 1976, for the Gegenbaur thesis).

Another example is the case of vertebral centra of bowfins and teleosts fishes. These have the same topological relationships relative to vertebral elements
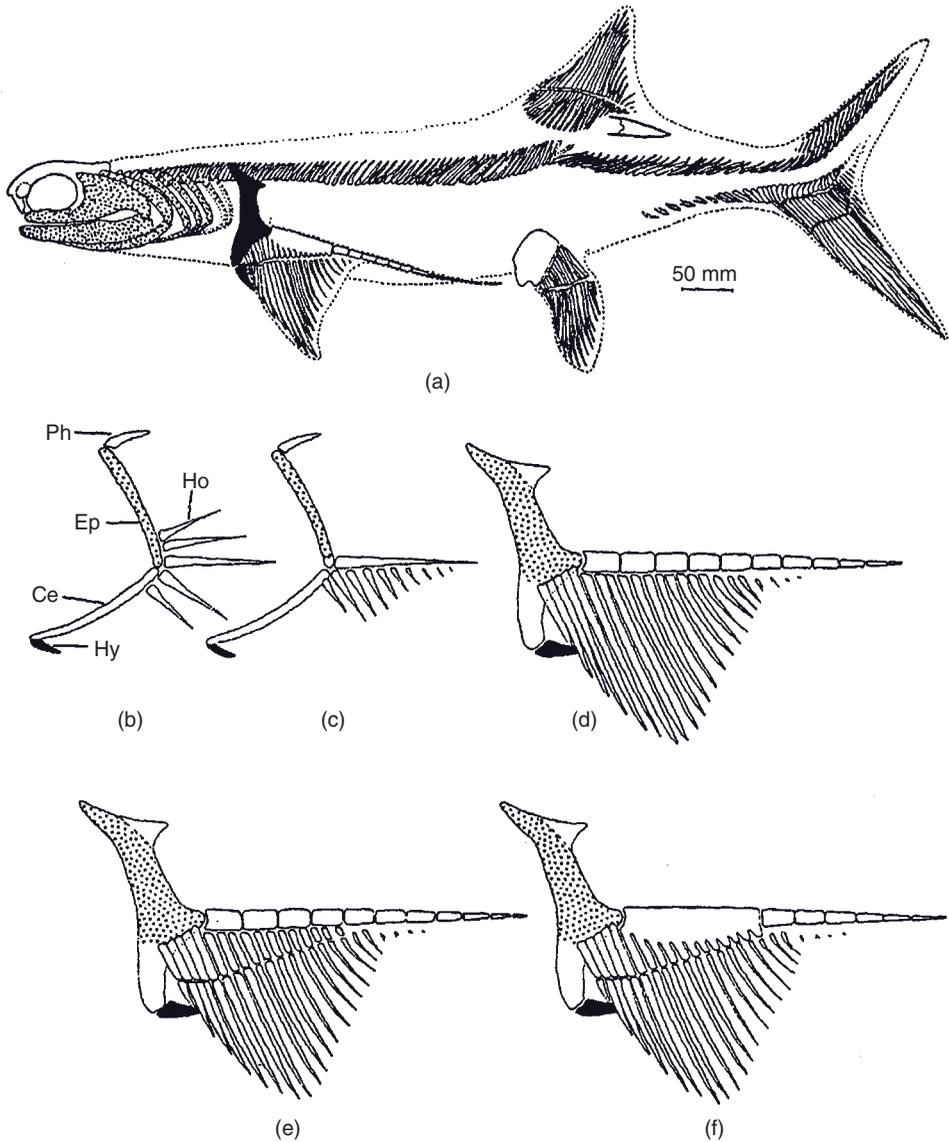
Figure 5.9. *Cobelodus aculeatus* (a) and Gegenbaur's hypothesis of the origin of the gnathostome shoulder girdle (b–f). In (a) the shoulder girdle is shaded black and the jaws and gill are stippled. In (b) the parts of the unmodified parts of the endoskeletal visceral arches are labeled, and subsequent modifications to produce the shoulder girdle are shown in (d–f). Ce, ceratobranchial; Ep, epibranchial; Ho, holobranch; Hy, hypobranchial; Ph, pharyngobranchial. From Wiley (1981a); modified from Zangrel and Case (1976).

(Fig. 5.10). However, Schaeffer (1967) points out that the centra are developed from different tissue layers and thus have different embryological origins.

Another use of the criterion of special similarity can be applied to cases where seemingly dissimilar characters might be good transformational matches because of their special similarities. For example, the transverse ventralis muscle, a muscle of the ventral gill arches in lungfishes, has a different insertion than the oblique ven-
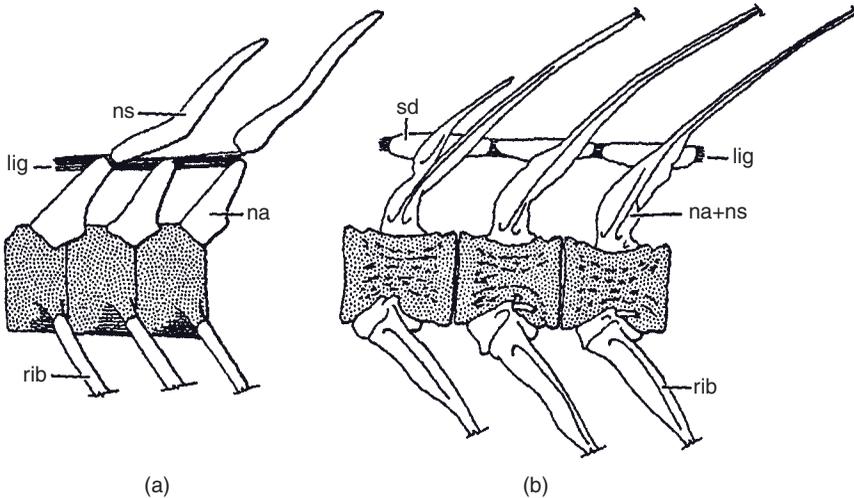
Figure 5.10. Series of vertebrae of (a) the holostean fish *Amia calva* and (b) the teleost fish *Salmo salar*. Although the centra of both species have the same topographic relationship to other ossifications, the centra develop from different tissue layers and are not homologous. na, neural arch; ns, neural spine. From Wiley (1981a); modified from Jollie, 1973.

tralis muscles of actinopterygian (bony) fishes. Further, some lungfishes have displaced transverse ventrali. Yet, during embryological development, the oblique ventrali of actinopterygians go through a transverse stage during their development such that it can be hypothesized that these are homologous with the adult transverse ventrali of dipnoans (Wiley, 1979d).

Returning to our original example of homology involving the fins of fishes and the legs of tetrapods; it fares well by some but not all of the criteria of determining homology we have described. The internal structure of an adult actinopterygian fin is totally different from the internal structure of an adult tetrapod leg. Shark pectoral girdles are made entirely of cartilage while those of bony fishes are made of a combination of dermal bone, cartilage, and bone ossified from cartilage. But if we look at development, a different conclusion can be drawn. Tetrapod legs are simply repatterned fish fins. They are repatterned as signaled by differential expression of Hox genes in response to differential cell growth and signaling by the sonic hedgehog gene (Shubin et al., 1997; Wagner and Chiu, 2001; Davis et al., 2007; Dahn et al., 2007).

Stacking Transformations: Intermediate Forms   The criterion of transformations through intermediate forms was the evolutionary counterargument to idealistic morphology. For example, we can see a reasonable transformation from an amphibian to a mammal in the evolution of the inner ear (Fig. 5.11) or the intermediate nature of the legs/fins of advanced sarcopterygian fishes. Initial matches made with this criterion are open to testing through congruence. It is common in phylogenetic research to match three or more characters because of similarities using other criteria (positional and special). If these map correctly on the resulting phylogeny, they would seem confirmed.
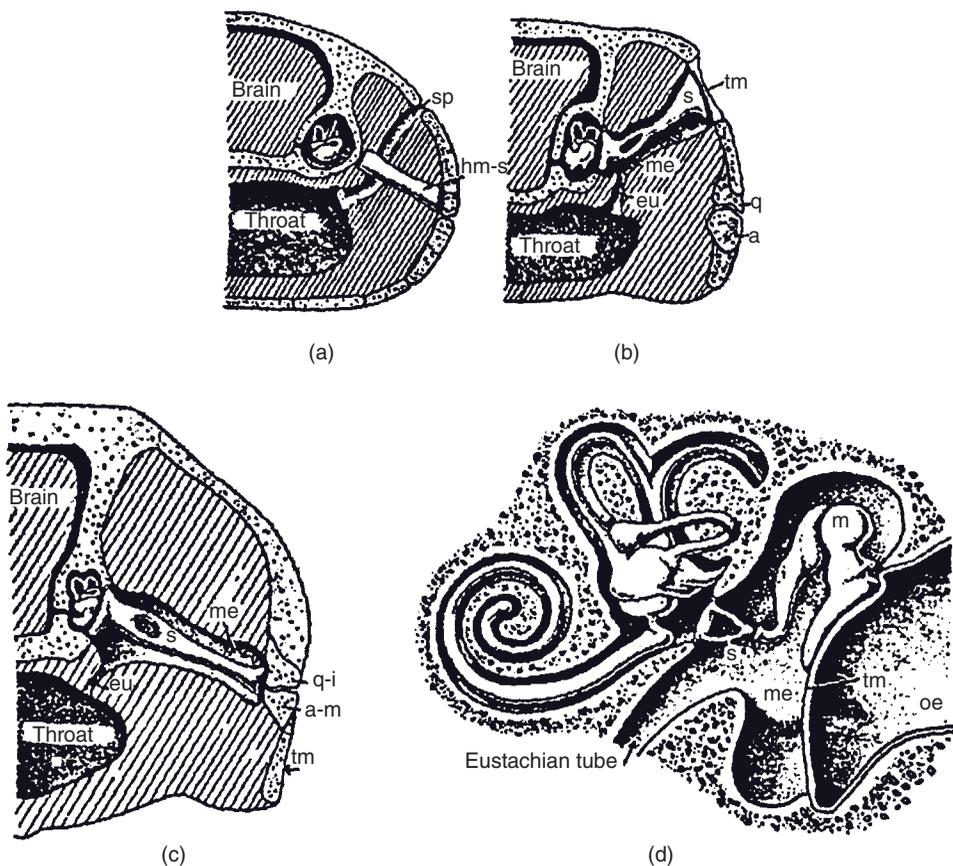
Figure 5.11. Remane's criterion of intermediate forms, the hypothetical evolution of the mammalian ear. Cross-sections through the skulls of (a) a fish, (b) an amphibian, (c) a reptile, and (d) a mammal. Abbreviations: a, articular; eu, eustachian tube (homolog of part of the spiracle); hm, hyomandibular (homolog of the stapes); i, incus (homolog of the quadrate); m, malleus (homolog of the articular); me, middle ear cavity (homolog of part of the spiracle); q, quadrate; sp, spiracle; tm, tympanic membrane. From Vertebrate Paleontology by A. S. Romer. Copyright 1966 by the University of Chicago Press. Used with permission.

Conjunction

Patterson (1982) coined the term *conjunction* for the class of tests that may lead to rejection of a primary homology conjecture by demonstrating that the supposed homologs are not actually comparable. De Pinna (1991) discussed the conjunction test in some detail. The conjunction test may be simply stated.

"If two supposed homologs are found in the same organism, they cannot be homologs" (Patterson, 1988:605). Patterson's (1988) example, discussed in some detail by de Pinna (1991), was a hypothetical one suggesting that the forelegs of humans and birds might be shown to be nonhomologous if angels were ever discovered because they have both. A more realistic example is paralogous gene sequences. Paralogs are two genes found in the same organism that are derived via gene dupli-
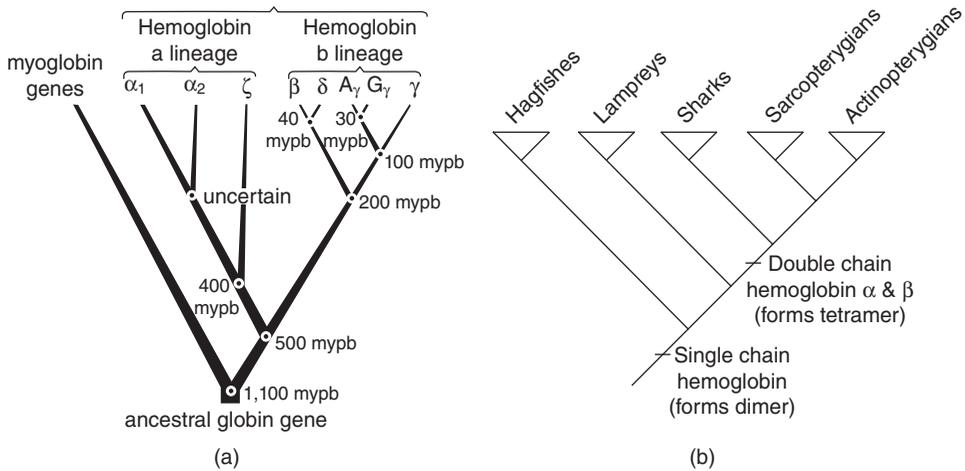
Figure 5.12. Globin evolution. (a) The relationships among certain globin genes myoglobin genes (redrawn from Ayala, "Evolution," in Encyclopedia Britannica). (b) The relationships of these genes on a phylogenetic tree.

cation from a common or orthologous ancestral gene (Fitch, 1970) (orthologous genes are the same gene found in two or more organisms). The hemoglobin gene family (Goodman et al., 1975; Dickerson and Geis, 1983; Hardison, 1996) is an ancient gene family and evolution of the hemogolin-α and hemoglobin-β paralogs in vertebrates serves as an example (Fig. 5.12). Gnathosotme vertebrates have both hemogolin-α and hemoglobin-β. These gene paralogs are interesting because they illustrate the differential levels of character properties. The presence or absence of paralogs are properties of monophyletic groups. For example, the presence of both genes forming tetramers in the functional proteins appears to be a synapomorphy of Gnathostomata. The zebrafish (*Danio rerio*, a teleost) and the frog *Xenopus laevis* have these genes on the same chromosome (Chan et al., 1997; Hosbach et al., 1983), a plesiomorphic condition similar to lampreys where the functional hemoglobin is composed of dimers rather than tetramers. Additional duplications occur in *Xenopus laevis*, and higher tetrapods (chickens and humans) which have created different genes belonging to both families, and these are found on different chromosomes, resulting in additional synapomorphies based on gene position. Sequence analysis of a random mix of paralogous gene sequences will not produce meaningful phylogenetic signal (Fitch, 1970, and others). Analysis of each taxon entered with all paralogous sequences will result in a series of duplicated relationships (see Goodman et al., 1975), each showing the gene tree of one of the sets of paralogs. The presence or absence of such genes families can help ferret out evolutionary patterns among the genes themselves, and hemoglobins have been found in one form or another in most organisms (Hardison, 1996). Thus, hemoglobin homology exists at different levels. At the level of taxa, shared DNA sequence properties of orthologous genes are relevant, but analyzing paralogous sequences would be a category mistake. At the level of taxa, presence and absence of particular members of the gene family is appropriate. At the level of the gene family, shared DNA sequence
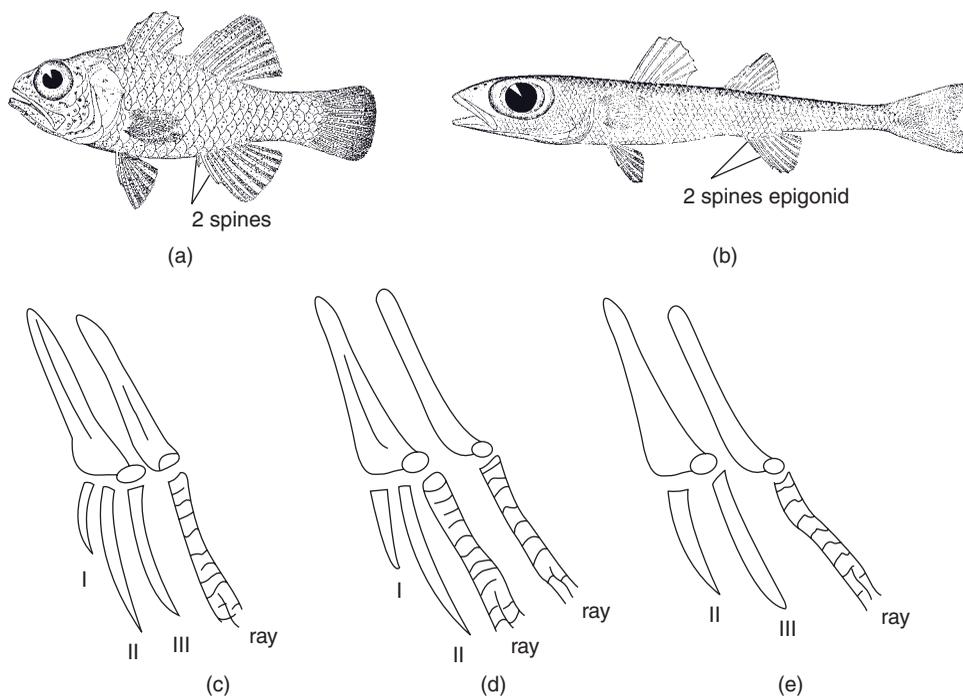
Figure 5.13. The character state "presence of two anal fin spines" is rejected as homologous between the teleost fish families (a) Apogonidae and (b) Epigonidae. The plesiomorphic condition (c) found in most percoid fishes is the presence of three spines, two supernumery spines (I, II) and one spine (III) associated with the anterior pterygiophore. In apogonids the two spine are spines I and II (d) while in epigonids the two spines are II and III (interpretation from Johnson, 1984). Fig. 5.13a from Jordan and Evermann (1900), Fig. 5.13b from Goode and Bean (1895).

properties are relevant to working out the descent of the genes themselves; as properties of the genes (contra the organisms), paralogous base positions are homologous. Thus, what is homologous at one level (sequence of paralogous genes relative to gene family evolution) is homoplasious at another level (sequence of paralogous genes relative to taxon descent). See Fig. 5.12a.

By way of a concrete aspect of this example, the base at position 146 of the hemoglobin-α gene is not homologous to the base pair at the same position in the hemoglobin-β gene because humans have both genes. But the base at position 146 in the hemoglobin-α gene might be quite comparable to that in the hemoglobin-β gene if the issue is to determine whether hemoglobin-α and hemoglobin-β are derived from lamprey myoglobin or lamprey hemoglobin.

Morphological examples are common in modular plants and metameric animals where efforts to homologize various parts may be rejected by showing that the characters are actually properties of different segments or modules. An empirical example is furnished by the work of Johnson (1984) on the relationships of fin spines and their internal supports (pterygiophores) in higher teleost fishes (Fig. 5.13). Among the 90+ families of percoid fishes (basses, snappers, etc.), the usual condition for the anal fin is to have three fin spines (usually labeled I, II, and III). The cardinal
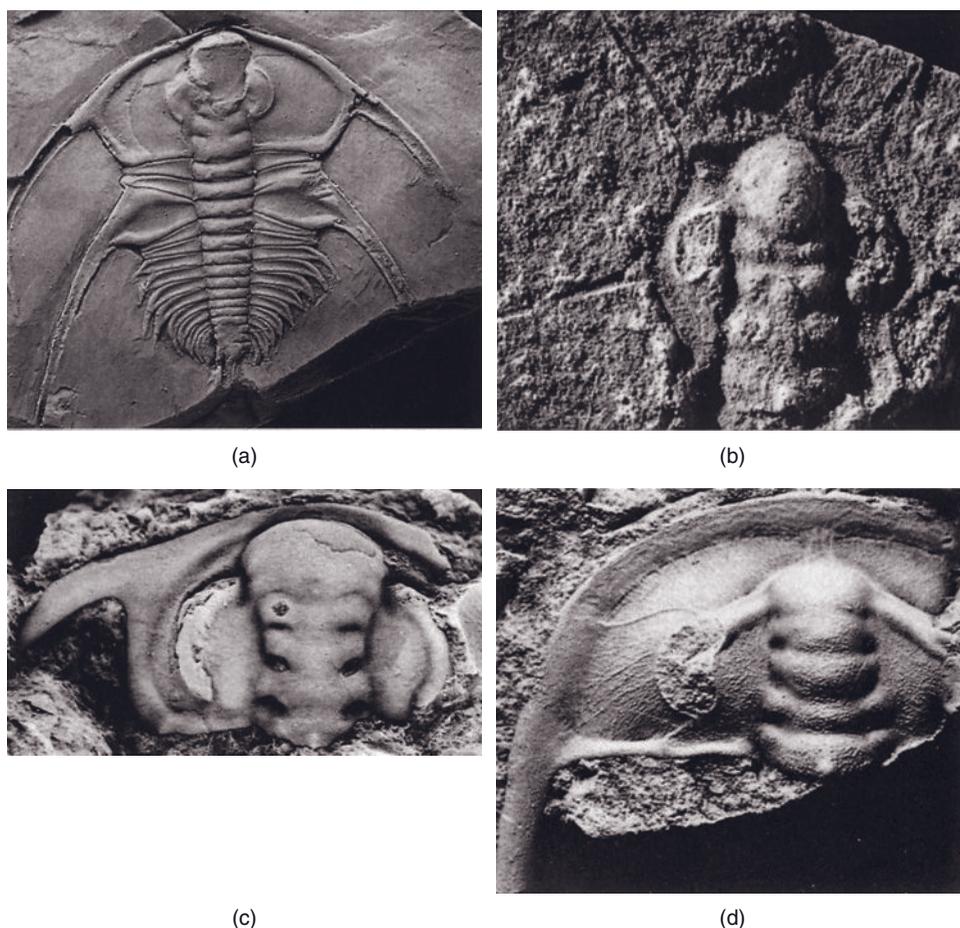
Figure 5.14. Different types of cephalic spines in trilobites illustrated using four species of Early Cambrian olenelline taxa. (a) *Bristolia harringtoni*, (b) *Olenelloides armatus*, (c) *Holmiella preancora*, (d) *Fallotaspidella musatovi*. From Palmer and Repina (1993), used with permission of the Paleontological Institute, University of Kansas. See color insert.

fishes (Apogonidae) and the deep-water cardinal fishes (Epinonidae) are unusual (but not unique) in having only two anal fin spines, and this was thought to be a uniting character. Closer inspection reveals, however, that the two spines of apogonids are spines II and III while the two spines of epinonids are spines I and II based on their positions relative to the supporting pterygiophores. Thus, the number of spines is not a property of a monophyletic group but rather a homoplasious similarity in two clades (Johnson, 1984). Trilobites, an extinct group of arachnomorph arthropods, show related phenomena. One example comes from a group of trilobites assigned to the Olenellina (Lieberman, 1998, 1999, 2001). Species can differ in the relative position and number of certain spines on the head (Palmer and Repina, 1993, see Fig. 5.14). The anteriormost pair of spines on the trilobite species shown in Fig. 5.14b are not homologous with the anteriormost spine of those species shown in Figs. 5.14a, c, or d. Instead, the second pair of spines in Fig. 5.14b is homologous with these; the homolog of the anteriormost spine pairs of the species shown in

Fig. 5.14b is not found in the adults of any of the other figured species. Moreover, the terminal pairs of spines on the head in Fig. 5.14b are homologous to the second pair of tiny spines in Fig. 5.14c, while that spine is effectively entirely lost in the species shown in Fig. 5.14a and d (see also McNamara, 1978).

Phylogenetic Homology (Forging Congruence between Hennig's and Patterson's Views)

Owen's (1843:379) general concept of "the same organ under every variety of form and function" provided the major preevolutionary criterion for hypothesizing homology. Working within the evolutionary paradigm, similarity is a means of hypothesizing reasonable statements of homology between characters (Patterson, 1982; Stevens, 1984), but it is only a beginning. Conjunction can sort out what Riedl (1978) termed *anatomical singulars* and can be decisive in sorting out homology among metameric morphological characters and paralogous genes (Patterson, 1982, 1988). However, once the evolutionary paradigm is embraced and a hierarchical view of the tree of life accepted, the criterion of congruence, or phylogenetic homology, becomes the decisive criterion (Hennig, 1966:93–95). Hennig recognized that Remane's criteria were only accessory and that "the real principal criterion—the belonging of the characters to a phylogenetic transformation series—cannot be directly determined." In determining whether transformational or taxic homology obtained, Hennig (1966:120) clearly saw the process of sorting out homologous and homoplasious characters as intimately connected to a phylogenetic hypothesis:

> The deciding whether different characters of several kinds are to be regarded as homologous, and therefore generally comparable with one another for the purposes of phylogenetic systematics, is a question of determining whether they can be regarded as transformation conditions of a character that was present in a different condition in a stem species, which did not have to be the stem species of only the compared species. But in deciding whether corresponding characters of several species are to be regarded as synapomorphies, convergences, homologies, or parallelisms we must determine whether the same character was already present in a stem species that is common only to the bearers of the identical character.

In other words, taxic homologs are synapomorphies while transformational homologies are hypotheses that require a nested phylogenetic hypothesis, as outlined above. Similar statements equating homology with synapomorphy are found in Wiley (1975, 1981a), Patterson (1982, 1988), Ax (1987), dePinna (1991), and others.

Avoiding Circularity: How Congruence Works

Congruence obtains when the groupings implied by one synapomorphy do not conflict with the groupings implied by another synapomorphy. Sometimes, congruent homologies could confirm the same monophyletic group, sometimes they could confirm nested monophyletic groups. The idea of congruence is closely allied with the idea of Hennig's auxiliary principle and parsimony because congruence is ultimately a question of testing phylogenetic hypotheses. Hennig (1966:121–122) stated his principle thusly:

I have therefore called it an "auxiliary principle" that the presence of apomorphous characters in different species is always reason for suspecting kinship [i.e., that species belong to a monophyletic group], and their origin by convergences should not be assumed a priori (Hennig, 1953). This is based on the conviction that "phylogenetic systematics would lose all the ground on which it stands" if the presence of apomorphous characters in different species was considered first of all as convergences (or parallelisms), with proof to the contrary required in each case. Rather, the burden of proof must be placed on the contention that in individual cases the possession of common apomorphous characters may be based only on convergence (or parallelism).

Hennig tied the auxiliary principle to the test of character congruence (Hennig, 1966:122). Characters are tested against one another to determine their phylogenetic utility. In the final analysis, this is again the method of "checking, correcting, and rechecking. …"

In unweighted parsimony analysis, congruence is achieved by minimizing the number of times one must hypothesize that similar characters are homoplasious and maximizing the number of statements of homology/synapomorphy (see Farris, 1983). Circularity is avoided because matches (based on Remane's criteria, for example) are tested by their congruence with other, independently analyzed matches. This is backed up by the expectation that evolution involves descent with modification. In statistical phylogenetic analysis and in weighted parsimony, congruence is a more complicated matter that involves maximizing homology given an a priori model of character change, with parsimony entering the picture in the selection of the model. That is, one picks the simplest model that can adequately explain the data. In both cases (parsimony and likelihood), the parsimony principle does not seek simplistic explanations but explanations as simple as the data allow.

An empirical question now arises. If taxic homology is equivalent to synapomorphy at the appropriate level in the tree of life, then how do we establish which homologies are synapomorphies at a particular hierarchical level? We shall explore this question in the next two chapters. For now, we turn to some practical issues concerning characters and their coding in a phylogenetic analysis.

WORKING WITH CHARACTERS

Systematists work with characters in two general ways. Descriptions of species and revision of groups are meant to provide basic data on the diversity and characteristics of taxa. Phylogenetic analyses are meant to place taxa in a historical framework. Although these two approaches frequently intersect, their purposes are different. Below, we take up the phylogenetic aspects of working with characters.

Discussions on the nature of characters and homology are issues of relevance to the practical matter of identifying characters suitable for phylogenetic analysis, but there is more involved. In effect, this is the distinction between ontology and epistemology. There is a considerable literature on the appropriateness of using certain kinds of characters in phylogenetic analysis and the actual distinctions between different kinds of characters (e.g., Pimentel and Riggins, 1987; Cranston and Humphries, 1988; Thiele and Ladiges, 1988; Chappill, 1989; Stevens, 1991;

Scotland, 1992; Thiele, 1993; Kitching et al., 1998; Wiens, 2001). Thiele's (1993) useful review of how systematists have discriminated and partitioned kinds of characters reveals the complexity. He recognized three kinds of comparative character data.

Qualitative Data in Phylogenetic Analysis. Qualitative character states share an identity. Character states that share the same identity and that are placed in the data column are those character states thought to be matches and potentially homologous. Two character states are qualitatively different if they are different in kind identity, and this is expressed by giving them different names or codes. Legs are qualitatively different than fins. Thymine is qualitatively different from guanine. Qualitatively different character states placed in a transformation series are asserted to have a homologous relationship as transformational homologs.

Quantitative Data in Phylogenetic Analysis. Quantitative data are similar in kind but different in degree, as in two leaves that have different lengths and widths. Their most basic description is in the form of numbers or adjectives that act as surrogates for numbers. What are usually expressed are properties of other properties. For example, two organisms have frontal bones. One frontal is long (character state "a"), the other short (character state "b"). The properties long and short are relative properties of the frontal bones and also properties of the organisms themselves. The fact that we ascribe different properties to the same bones assumes that we have accepted the taxic homology of the bones themselves. This acceptance will be important when we consider how morphometrics may play a role in phylogenetics, and this is discussed in a separate section below. It would be unusual to ascribe the same properties to, say transformational homologs; for example, contrasting a long pectoral fin with a short leg.

The usual practice in phylogenetic analysis is to use qualitative data. Indeed, Pimentel and Riggins (1987) advocated the use of such data to the exclusion of quantitative data. The problem is, many kinds of qualitative data are quantitative data in disguise (Stevens, 1991), with the distinction between qualitative being at times a semantic one (Wiley, 1981a; MacLeod, 2001). This is part of a broader debate in science as to whether all science must be strictly quantitative in form. Of course, it is unlikely that this debate will be resolved any time soon, and when applied to phylogenetics, we suspect that both quantitative and qualitative characters, however these may be defined, will have phylogenetic utility.

Thiele (1993) suggests that quantitative differences marked by widely discontinuous patterns of variation are useful and justifiable. The question is: what does "widely discontinuous pattern" mean? Stevens (1991) has pointed out that qualitative expressions may hide patterns of continuous variation. This brings us to the two categories of quantitative data discussed by Thiele (1993).

Continuous Data. Continuous data are mathematical properties such as length or other dimensions such that there is a continuous range of values measured. The kind of continuous data most interesting to systematists are measures of intrinsic property values taken from different specimens on characters thought to be homologous.

Discrete Data. Discrete data differ from continuous data in that the property can be expressed by only a few values. One common kind of discrete data is meristic, or count data. Data such as the number of stamens of a plant, the number of setae per segment of a worm, the number of segments of a trilobite, or the number of pectoral fin rays of a fish are examples of discrete data, and such data are expressed

as integers. Other kinds of discrete data usually considered qualitative include presence/absence data and DNA sequence data.

Overlap. Thiele (1993) suggests that one of the most important properties of character states for phylogenetic analysis is the extent to which they overlap. In traditional practice, character states may be coded as discrete if they show a disjunct distribution of values between species or populations. Broadly overlapping character states between taxa are frequently dismissed as not useful. Overlapping character states are frequently coded as discrete, and the extent to which some statistical evaluation is made before this decision may be unclear. The process of evaluating such character states has been described as "filtering" (Thiele, 1993). Inherently discrete character states that "overlap" indicate character polymorphism, and the taxa whose parts/specimens show polymorphism can be coded as such (that is, with both codes).

Thiele (1993) and Wiens (2001) ask two important questions concerning characters. First, are most character states usually described in qualitative terms actually quantitative? Second, has "traditional phylogenetics" dismissed many quantitative characters that have phylogenetic signal?

Qualitative versus Quantitative Characters: Avoiding Vague Characters

In his review of character analysis, Wiens (2001) makes two important points about character descriptions. First, many character descriptions, cast as qualitative character states (wide versus narrow; long versus short), are simply vague descriptions of quantitative characters and, thus, should be restated by defining the trait in a quantitative manner. Second, many quantitative character states that come in the form of counts with ranges in values (e.g., 1–3 scales or 4–6 scales) are not accompanied by a rationale for why the particular ranges were picked. That is, why not 1–2 scales, 3–4 scales, as opposed to 1–3 scales, etc.? Different treatment of such cutoffs can lead to different phylogenetic conclusions (Gift and Stevens, 1997). Of course, usually a rationale does exist and is not stated merely to save on space; however, if the rationale were based on a simple, untested assumption, it could be problematic. Wiens (2001) suggests that coding the character states as continuous variables may solve the problem. Partly, Wiens is confusing the ontological nature of characters as opposed to how we define them epistemologically: what do we really mean when we say such and such character exists as opposed to how we identify characters in general? To implement his suggestion, one can turn to the extensive literature on how to accomplish coding of continuous characters, including gap coding (Mickevich and Johnson, 1976), segment coding (Colless, 1980), divergence coding (Thorpe, 1984), simple gar coding (Almeida and Bisby, 1984), generalized gap coding (Archie, 1985), range-coding (Baum, 1988), m-coding (Goldman, 1988), gap-weighting (Thiele, 1993), finite-mixture coding (Straight et al., 1996), analysis of variance-multiple range test (Sosa and De Luna, 1998), and step-matrix gap-weighting (Wiens, 2001). Of five methods compared by Garcia-Cruz and Sosa (2006), Thiele's gap-weighting method yielded the most number of parsimony-informative data.

The goal of most of these methods is to produce qualitative states for continuous or meristic characters by searching for gaps of variation. The methods of Thiele (1993) and Wiens (2001) differ in that they incorporate information about the distance between state values based on the differences between mean values of the

characters, essentially weighting the changes based on these differences. Wiens' (2001) modified method, incorporating step-matrices, is designed to increase the number of possible codes from 13 to 999. (Garcia-Cruz and Sosa, 2006, did not evaluate this method.) This is an interesting proposal, but one concern might be that Wiens' modified method, with its emphasis on epistemology, might introduce certain artifactual effects. In particular, how do phylogenetics programs typically deal with such continuous characters and might biases in the broader phylogenetic results follow when such methods are employed? In short, is the cure worse than the disease? Ultimately, these types of questions need to be explored in greater detail using modeled and actual data. Our suspicion is that the emphasis on defining characters in qualitative terms does work (consider the long history of successful systematic research) and qualitative is not just a substitute for poor quantitative; however, it is also true that characters are ultimately the data used to identify phylogenetic signal, and thus we should define them carefully and recognize the consequences of one versus another type of character homology statement. Further, continuous, quantitative character data could be more fully used, especially when looking for phylogenetic signal within species or among closely related species. There are proposals for dealing with such data as shapes, as we discuss in the next section.

Morphometrics and Phylogenetics

Another approach to transforming continuous variation into discrete phylogenetic characters, particularly in terms of shape assessments, might be achieved through morphometric analysis. Morphometrics is the study of covariances of biological form, especially as these pertain to shape (e.g., Bookstein, 1991). The goal is to understand how variation in the size and shape of biological objects (which together constitute the objects' form) is patterned with respect to variation in other biological objects or various factors that influence form variation (e.g., environment, function, development, selection). Aspects of the form of interest are captured through a series of simple measurements and the relationship among the measures taken as a whole is explored with multivariate data analysis techniques such as singular value decomposition (a form of principal component analysis).

Traditional approaches to morphometric analysis applies bivariate or multivariate techniques to quantitative variables such as distances or angles between sets of landmarks that can be located on all individuals in a sample. The basic technique consists of picking comparable landmarks, taking the measurements (length, width, height) based on these landmarks, and then asking questions about those aspects of the specimens' bodies that were measured. These questions may concern variation within populations (principle components analysis) or discrimination between populations (e.g., discriminate functions analysis). Many of these traditional techniques continue to be useful in systematic studies. However, some have argued that these are not the kinds of characters useful in phylogenetic analysis (Pimentel and Riggins, 1987).

In the 1990s, morphometricians turned to techniques that attempted to capture the geometry of morphological structures, reinvigorating the field to such an extent that Rohlf and Marcus (1993) described the changes in concept and practice as a

revolution. This new approach has come to be called *geometric morphometrics*. A brief review of the history of the geometric morphometrics movement is provided by Adams et al. (2004).

How morphometrics in general and geometric morphometrics in particular might relate to phylogenetics is still controversial. Before the "revolution," and as mentioned above, Pimentel and Riggins (1987) argued that the variables obtained from principle components analyses (eigenvectors, matrices of covariation, etc.) do not correspond to anything that can be related to a biological concept of homology that is necessary in order to enter into a phylogenetic analysis. MacLeod (2001) agreed, but suggested that this deficiency can be fixed by restoring to the analysis a sense of the topological relationships among the structures the measured variables are describing that provides the framework for recognizing homology irrespective of whether the concepts are applied to qualitatively or quantitatively defined characters. MacLeod (2001:197–198) also argued that incorporation of geometric morphometric tools into phylogenetic analysis was delayed by politics, as the geometric morphometric movement grew out of phenetics. However, this did not inhibit two proposals for incorporating geometric morphometrics into phylogenetics.

The first proposal came from a group at the University of Michigan (e.g., Zelditch et al., 1995) that includes both morphometricians and phylogeneticists. Basically the proposal was to directly use the deformational morphometrics that grew out of the work of Thompson (1917; see Bookstein, 1991) by isolating deformational differences between specimens and using these parameters directly in a phylogenetic analysis. The proposal called for using homologous landmark points collected over the entire body to discover localized structural homologies. By reducing the description of shape variation to a mathematical result they hoped to discriminate between alternative localized deformational forms of the same structure. They could then code them for use in phylogenetic analysis. This approach depends on the landmark points being biologically homologous.

This approach met with criticisms from a variety of other morphometricians (Bookstein, 1994; Naylor, 1996; Rohlf, 1998). Two important and relevant criticisms are (1) the extracted variables are not independent, thus failing the expected properties of phylogenetic characters as being semiautonomous properties of the organism and (2) geometric homology and phylogenetic homology are not the same concepts. In a perceptive review of their own methods, Zelditch et al. (2004) have retreated from their original position and now doubt that geometric morphometrics can be used directly in phylogenetic analysis. Instead, they suggest that morphometric protocols such as PCA can be used to discover characters that might be coded in a manner that would allow phylogenetic analysis (but see Pimentel and Riggins, 1987). However, Zelditch et al. (2004:380) stated: "Until we can define 'character' precisely, in terms just as comprehensible to mathematicians as to systematicists, we will make no further progress towards a mathematical solution."

The second proposal came from workers such as Naylor (1996) and MacLeod (2001, 2002) and is more optimistic. In this approach, we isolate parts that have been determined to be good candidates for being localized structural homologs, use morphometric techniques to document patterns of shape variation of these a priori homologs, and then use morphometric analysis to search for discrete clusterings of different shapes that could be coded in a qualitative manner for phylogenetic

analysis. MacLeod (2001) suggests that there are two relevant questions concerning the application of morphometrics to phylogenetics. (1) Can morphometrically defined variables exhibit a hierarchical structure that can be used to define nested sets of taxa? That is, can they be ordered into a transformation series and used in a phylogenetic analysis? MacLeod (2001) answers "yes," based on the simple fact that phylogeneticists have been doing this for years, as have other systematists. (2) Are groups circumscribed as monophyletic when using morphometrically defined variables the same as monophyletic groups circumscribed by other kinds of data (in MacLeod's papers, other "traditional" morphological data)? In other words, do we observe congruence that allows us to suspect that the morphometrically defined characters (variables) are acting just like nonmonphometrically defined characters? MacLeod argues that they are.

MacLeod (2001, 2002) suggests that part of the problem is the fact that geometric and phylogenetic homology are different concepts. For example, there may exist a geometric homology between the point location of the anterior base of the dorsal fin of a shark and a whale, but there is no phylogenetic homology between these point locations (called landmark points) because the structure themselves are not homologous. This point is also clearly recognized by Zelditch et al. (2004:177), whose examples are a scapula, a potato chip, and a chocolate chip cookie. But MacLeod questions the very concept that biologically homologous landmarks are necessary to discover morphometrically generated characters. In this respect, he differs from the approach used by Zelditch et al. (1995), Fink and Zelditch (1995), and Swiderski et al. (2002). To MacLeod, biological homology resides in the structure itself and not the points used by morphometricians as convenient ways to represent aspects of the shape of that structure. In other words, the character is the structure, and the shape, as defined by alternative configurations of the landmark points, is the character state. Similar point configurations (=shapes) may be considered primary homology statements, candidates for taxic homology; different point configurations may be candidates for transformational homology, as in the plesiomorphic condition of a square frontal bone and the apomorphic condition of a rounded frontal bone. MacLeod suggests that if we restrict ourselves to the analysis of shape variation by comparing localized structural components of organisms that are thought, a priori, to be phylogenetically homologous, then we can use the results of morphometric analyses to discover and discriminate between alternative shapes of these homologous structures. Further, MacLeod suggests that this is exactly what phylogeneticists do when they code the shapes of homologous structures qualitatively in a phylogenetic analysis, only with the added rigor of a geometric analysis of the shapes involved to test the proposition that the shapes are indeed separated in the relevant geometric space by discontinuities in shape variation.

An example of how the approach suggested by MacLeod (2001) might work in phylogenetics is shown in Fig. 5.15, which is a modified version of the example used by Zelditch et al. (2004), the simple case of two triangular bones. We observe that the same homologous bone comes in two seemingly different forms (Fig. 5.15a). In the nonmorphometric world, we might code the isosceles condition as "a" and the equilateral condition as "b." Or, because we know that the outgroup has the isosceles shape, we might code it as "0" and the other as "1" if we are using the convention that the presumed plesiomorphic condition is scored "0." Are the shapes really distinct? That is, is there a discontinuity in shape or is our value judgment incorrect?
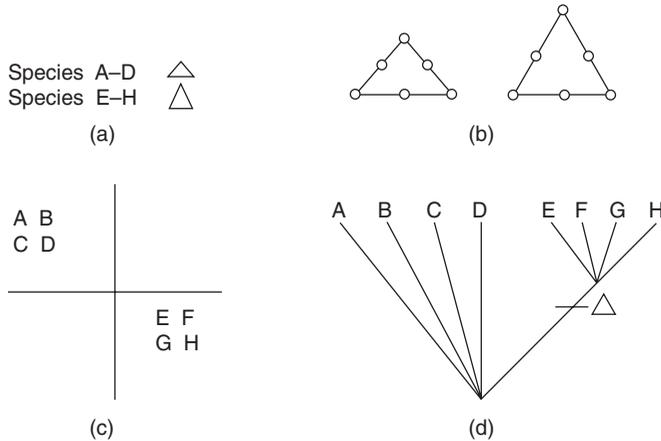
Figure 5.15. A simple example of morphometrics applied to phylogenetics. (a) Mean shapes of eight species for a hypothetical structure. (b) Landmarks (dots) picked to characterize the shapes of the structures. (c) The results of relative warp analyses showing that species with similar shapes cluster together in morphometric space and are statistically different, implying that variation is discontinuous. (d) A phylogenetic analysis assuming that taxon A is the outgroup and can be used to root the tree.

We can describe the shapes mathematically by selecting a series of landmark points (Fig. 5.15b). Some of these might meet the criterion of biologically homologous points as defined by Zelditch et al. (1995), but others might simply be picked to ensure that the bone (triangle) outlines are adequately captured. We then perform a relative warp analysis and see two distinct clusters, well separated in morphometric space (Fig. 5.15c). We conclude that the two clusters of shape are distinct and separated by a discontinuity. Thus, we feel confident (given our present sample) in our coding and the shape characters meet our criterion of being biologically interpretable as different states that can be used in a phylogenetic analysis (5.15d).

There are many terms and concepts that must be mastered if one wishes to apply morphometrics to discover phylogenetic characters. They are not that difficult if one has had a course in multivariate analysis. For example, the relative warp analysis mentioned above is a principal components analysis (eigenanalysis of the covariance matrix) of Procrustes-aligned shape coordinates. For the details, see MacLeod (2001, 2002) and for different views of both uses and values of applying morphometric analysis to the discovery or tests of shape characters for phylogenetic analysis see individual papers in Adrain et al. (2001) and MacLeod and Forey (2002).

We see a productive future in a partnership between geometric morphometrics and phylogenetics. It will not be the laudable goal of extracting phylogenetic characters directly from the mathematics of geometric morphometric analysis as envisioned by Zelditch et al. (1995). Instead, it will be along the lines suggested by MacLeod (2001), where the morphometric techniques are used as a tool to test the proposition that there are discontinuities of shape among homologous characters
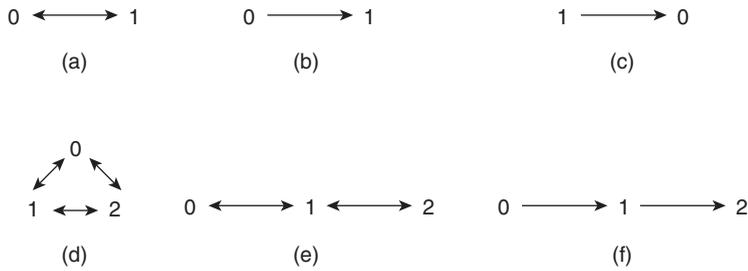
Figure 5.16. Relationships between characters: (a) binary, unpolarized; (b–c) binary, polarized; (d) unordered, unpolarized; (e) ordered, unpolarized; (f) ordered, polarized. Redrawn from Wiley et al. (1991), used with permission, Biodiversity Institute, University of Kansas.

in geometric morphometric space. As technologies such as 3-D imaging become widely available and enhance the ability to discriminate among shapes, geometric morphometrics will aid in our efforts to avoid vaguely described characters.

Characters, Transformation Series, and Coding

Although continuously variable characters can be analyzed phylogenetically (e.g., Goloboff et al., 2006), most characters appearing in the literature are coded qualitatively or are discovered through means such as morphometric analysis and then coded qualitatively. In a phylogenetic analysis, the matrix is composed of a number of data columns, which represent primary homology statements. There are usually at least two character states in a transformation series of morphological characters, and there may be more than two. In molecular analyses, the unit of analysis is an entire gene segment, with character states being expressed as base residues (ATCG), or amino acids (20 standard amino acids plus some rare nonstandard amino acids). Alignment may introduce other characters, such as gaps, and gaps may be coded as missing or real information.

Discrete character states are usually coded with a number or a letter. The usual convention is to code the presumed plesiomorphic homolog as "0" or "a" and the presumed apomorphic homolog as "1" or "b," with longer transformation series numbered in one-step integers. Modern computer programs can easily handle polymorphisms, so species having more than one character can be coded with both. Transformation series with more than two character states can be coded in a number of ways, depending on what one thinks one knows about the relationships among the character states. The most common kinds of coding are listed below.

Characters with two and only two states are binary. The relationship among binary character states is automatically ordered, but not necessarily polarized. Ordering specifies a particular path, but not a particular direction. The ordered binary in Fig. 5.16a may be polarized in two directions (Fig. 5.16b, c).

Characters with three of more states may have more complex relationships among character states because they are not necessarily ordered. Three unordered states (Fig. 5.16d) specify no particular direction of transformation while three ordered states specify a particular route along which evolution can occur one step at a time, although it does not specify the direction of charge (Fig. 5.16e). Three

Figure 5.17. Binary versus linear coding. (a) A transformation series polarized by outgroup comparison. (b) A binary matrix. (c) A linear matrix. Note that additive binary coding is additive because summing the rows in (b) results in the linear matrix. Redrawn from Wiley et al. (1991), used with permission, Biodiversity Institute, University of Kansas.

states that are ordered and polarized specify both a route and a direction (Fig. 5.16f). It is also possible that very complicated state relationships might exist (or be inferred to exist), such that one could construct a hierarchical character state tree. Such characters require more complex coding, such as step matrices (Maddison and Maddison, 1992), which are one form of generalized parsimony (Sankoff and Cedergren, 1983; Swofford and Maddison, 1992) where unequal probabilities of change are assigned to certain transformations (Ree and Donoghue, 1998).

General discussions on the distinction between ordered and unordered character states can be found in Farris et al. (1970), Fitch (1971), Mickevich (1982), Pimentel and Riggins (1987), Swofford and Olsen (1990), Mickevich and Weller (1990), Mickevich and Lipscomb (1991), and other papers summarized by Wilkinson (1992). Synonyms for unordered states include "nonadditive" and "maximally connected" character states. Synonyms for ordered states include "additive" or "minimally connected" character states (see Mickevich, 1982, and Slowinski, 1993, for different terms).

Additive binary coding. Additive binary coding (Sokal and Sneath, 1963; Kluge and Farris, 1969; Farris et al., 1970) is an alternative way of ordering a linear multistate character that preserves specific hypotheses of character transformation by segregating each transformation into a separate column (Fig. 5.17). This kind of coding is not used much in character analysis because it has the same effect as specifying Farris optimization for character states, but it is useful as an introduction to more complex methods and does have its use in biogeographic analysis. We can see that the relationships are additive because if the columns are summed they result in a linear coding. In addition, it is also true that if the character states have a branching relationship, then nonadditive or mixed coding can be applied.

Nonadditive binary coding. Nonadditive binary coding is used to capture relationships among states that are hypothesized to have a nonlinear relationship (Fig. 5.18). Each hypothesis of transformation is given a separate column, but because some transformations are isolated in the character tree, the rows do not sum to the equivalent of a linear transformation series. This coding method and mixed coding can be traced back to papers by Pimentel and Riggins (1987), O'Grady and Deets

Figure 5.18. An example of nonadditive binary coding. (a) A character–state tree expressing the apriori relationships among states. (b) Hierarchial information captured by a binary matrix. (c) A linear matrix coding each state separately (hierarchial information lost). (d) A "mixed" matrix capturing the hierarchical information in more compact fashion. Redrawn from Wiley et al. (1991), used with permission, Biodiversity Institute, University of Kansas.

(1987), and O'Grady et al. (1989) but apparently originated in a unpublished manuscript by M. Mickevich (O'Grady and Deets, 1987).

Mixed Coding. Mixed coding is a form of nonadditive binary coding. One data column includes states for several hypotheses of transformation in a linear fashion while other columns cover the branches. Mixed coding is a space-saving technique that eliminates at least one extra data column to capture the hypotheses of transformation (Fig. 5.18d).

Although sometimes the usage of particular character coding is clear, for instance additive binary versus unordered multistate, other times it may not be. For instance, often it is easier to posit homology relationships than it is to posit an order of homology transformation. Further, when there is a fair degree of missing data, the implementation of binary coding may be less problematic with respect to how phylogenetic algorithms deal with ambiguity. However, it is also true that creating separate binary characters from single multistate characters allows a computer algorithm to treat these characters as potentially separate and independent when, in fact, they may not be. The extent to which this happens can be investigated by mapping characters back onto the tree on the resultant phylogeny. Thus, each method and approach has its strengths and weaknesses. Ultimately, with the creation of binary characters, the focus is more on a taxic definition of homology, whereas with the creation of multistate characters the focus is more on a transformational definition of homology.

Complex Characters or Separate Characters?

While the four methods outlined above are common coding schemes that will work for many characters, they are by no means the only methods of coding. Wilkinson (1995a) has drawn attention to coding decisions involving characters that may covary, i.e., those characters that might be treated as independent or dependent characters depending on decisions made by the investigator. He recognizes two kinds of character constructions.

Reductive coding treats character states as quasi-independent characters that are decoupled in evolutionary processes and, thus, are naturally placed in different data columns. This is the usual decision made by the investigator and results in one of the four strategies outlined above. While some coding methods place a single transformation series in different columns, this is a limitation of the ability to code rather than a hypothesis that the characters are independent, and does not affect the overall tree length. For example, an ordered linear transformation series will result in the same number of steps as the same transformation series rendered in additive binary form, and because a linear vector cannot represent bifurcations, we are forced to adopt mixed or nonadditive binary coding for such hypotheses of character transformation.

Composite coding treats two character states as coupled and represents them in the same transformation series as a series of composite character states. Wilkinson (1995a) uses the example of Wake's (1993) coding scheme, rendering a series of eye muscle characters in caecilian amphibians as a series of composite states based on presence and absence of certain muscles. Wilkinson (1995a) makes the point that the decision to use composite coding is not entirely clear, but one rational decision would be to consider that the composite states were not quasi-independent. Unfortunately, because we do not know much about the genetics and epigenetics of characters, this may simply be a guess. Yet, if the character states are truly dependent, composite coding carries the burden of overestimating homoplasy. Neither decision is correct a priori, but the decision can influence the phylogenetic results.

Missing Data

Investigators are frequently faced with the fact that some taxa in an analysis have incomplete character information. This problem, the "missing data" problem, has been recently reviewed by Wiens (2003a) who summarized much of the literature that forms the basis for this section. Kearney and Clark (2003) also review the problem. Missing data can lead to multiple equally parsimonious trees and collapsed consensus trees (Gauthier, 1986; Nixon and Davis, 1991; Nixon and Wheeler, 1992; Novacek, 1992a, b; Maddison, 1993; Wilkinson, 1995a, b; Wilkinson and Benton, 1995; Gao and Norell, 1998; Wiens, 2003b). In fossils, this may be caused by incomplete preservation. When combining information from different studies (molecular, morphological, paleontological) in a total evidence matrix, the challenge becomes combining taxa for which different kinds of data might be available (e.g., Anderson, 2001). Avoiding entering missing data in a combined character analysis has led some to either exclude taxa with missing data (e.g., Wiley, 1976; Patterson, 1981) or practice taxonomic congruence in the form of construction of supertrees (Sanderson et al., 1998; Liu et al., 2001). However, whenever possible one should seek to

include as broad a range of taxa as possible. In particular, the inclusion of fossil taxa is often critical as they retain character states not seen in living taxa, and thus can yield insights into the synapomorphies of major clades. A compelling example comes from tetrapod phylogenetics. Analyses of only living species of amniote vertebrates results in recognition of a group, "Homeothermia," that includes mammals and birds. However, when fossils are included, birds group with other archosaurs (i.e., crocodiles, dinosaurs) (Gauthier et al., 1988; Donoghue et al., 1989), which seems more plausible, and is also now supported by the molecular data. Inclusion of fossils can also make a difference in how we view character evolution.

For example, if we consider only living teleost fishes, there are 27 hypothesized synapomorphies (de Pinna, 1996), which might suggest to some saltatory evolution. But if we include all known relevant fossils, the list shrinks to a single unique synapomorphy and about 8 synapomorphies that show homoplasy, depending on outgroup choices (Arratia, 1999). That inclusion of fossil taxa should decrease the potential list of synapomorphies should come as no surprise given that most systematists do not embrace saltatory evolution; that is, characters evolve piecemeal as a group differentiates via cladogenesis, and what today seems to be a host of characters that define a node in a tree of extant taxa likely did not crop up all at once. Each one of those characters may have been acquired in a separate cladogenetic event; however, because those intervening, missing link taxa are extinct, they might not be sampled in a neontological study.

The problem of missing data certainly arises when paleontological and neontological studies interdigitate, but the same situation can occur in recent taxa, for example, when we are unable to observe certain characters due to the rareness of specimens precluding dissection or other methods of preparation or the inability to sequence a gene region for some taxa (Wiens and Reeder, 1995). In a broader respect, the extant biota is actually a highly pruned sampling of the diversity of life given that more than 99.9 percent of all species that have ever lived are extinct. Neontological studies, to the extent that they do not or cannot sample these taxa, must remain incomplete. The question as to how this affects phylogenetic accuracy is not fully known, although the aforementioned studies by Gauthier et al. (1988) and Donoghue et al. (1989), and also the modeling-based study of Wheeler (1992), suggest that it is likely a significant effect.

There is, however, another aspect to the debate about including missing data. That is, that missing data can cause phylogenetics programs to produce artifactual, or spurious, results. In particular, taxa for which there is a large amount of missing data can serve as wildcards that roam around the tree, either mapping in particular places without significant, true support, or causing the overall support of the tree to decline. For this reason, Wiens (2003a) noted that sometimes taxa with a large amount of missing data can be excluded from the study (e.g., Wiley, 1976; Wiens' examples are Rowe, 1988; Grande and Bemis, 1998) or characters that are coded as missing for many taxa can be excluded (Wiens' examples are Livezey, 1989; Smith et al., 1995). There is a risk in employing such strategies because many studies have shown that increasing the number of characters and increasing the number of taxa improve the probability of obtaining an accurate phylogenetic tree (e.g., Wiens, 2003a). Simulations include analyses of data from laboratory produced phylogenies of viruses (Hillis et al., 1992, 1994; Wiens and Reeder, 1995), congruence analysis (Miyamoto and Fitch, 1995; Cunningham, 1997; Wiens, 1998b), and other types of

computational studies (Huelsenbeck, 1991a, 1995; Huelsenbeck and Hillis, 1993; Hillis et al., 1994; Hillis, 1995; Graybeal, 1998; Wiens, 1998a, b, c; Wiens and Servedio, 1998; Wiens, 2003a).

Wiens (2003a, b) suggests that there are actually two missing data issues. The first is the absolute or proportional number of missing character states (cells with no data). The second is the distribution of those missing data cells. A matrix with the absolute or proportional number of missing data cells that is randomly distributed among taxa with missing data may lead to inaccurate phylogenies (under simulated conditions where the phylogeny is known). However, if the same character states are scored for each taxon with missing data, accurate results may obtain even if the proportion of missing data is the same. This suggests that the accuracy of a phylogenetic analysis can be increased by sampling more characters that are available for all taxa and that the amount of missing data in absolute terms may not be so important. In contrast, adding characters that cannot be scored for all taxa might not help.

Wiens (1998b, 2003b) draws a number of conclusions. The number of missing data cells in a matrix is not the problem, but the number of incomplete columns of data is. In other words, sometimes too few characters (columns of data) have been sampled from the incomplete taxa to allow the algorithm to place them accurately on the tree. Thus, adding additional characters that are observable in such taxa increases phylogenetic accuracy. Accurate placement is easier if the incomplete taxa have the same missing data cells rather than a random assortment of missing data cells. By contrast, deleting characters that are not available for all taxa does not help. In fact, adding sets of incomplete data may be either neutral or even beneficial to the accuracy of the analysis in the absence of long-branch attraction. However, if the characters in question are those involved in long-branch attraction, adding such taxa will not solve the problem. If there is long-branch attraction among the taxa with complete data (e.g., taxon sampling is small), adding taxa with incomplete data may be beneficial, depending on the level of data completeness and their relevance to breaking up the long branches. The extent to which long-branch attraction is important either in a particular data set or in general determines how important adding additional taxa will be for obtaining accurate results. We note that building on the early work of Gauthier et al. (1988) and Donoghue et al. (1989) more and more studies have successfully combined fossil and recent taxa in matrices, including not only morphological but also molecular character data (e.g., Eernissee and Kluge, 1993; Wheeler et al., 1993; O'Leary, 1999; Gao and Shubin, 2001; Sun et al., 2002; Hermsen and Hendricks, 2008). Wiens (2003b) concludes that the limiting factor for successful analysis is the number of relevant characters that can be scored for the fossil taxa: if the fossil taxa can be placed on a tree of morphology, then they should be accurately placed on a tree that combined molecular and morphological data. Fossils are less likely to solve long-branch attraction problems if the numbers of recent taxa are low, but again, this reflects the critical importance of adequate taxon sampling (be they fossil or extant) in phylogenetic studies.

Homology and "Presence-Absence" Coding

We have argued above that homology takes two forms: taxic and transformational. However, some workers have asserted that only taxic homology exists and that

phylogenetic analysis can be undertaken by considering taxic and only taxic homology. This approach takes the form of presence-absence matrices and "three-taxon" analysis *sensu* Nelson and Platnick (1981). In this chapter, we will only consider issues of homology relative to the approach, we will briefly cover three-taxon analysis in Chapter 6.

We were struck by a recent paper by Mooi and Gill that made a curious statement: "Conversion of characters to matrices of 0s and 1s has changed the way we think of character states. In such a matrix, if 1s are considered apomorphic, all taxa with 0s are often considered to 'share' a state, when in fact they do not. 0s among taxa are not equivalent—having 0 only means 'not having 1'—which means: we have no further information" (Mooi and Gill, 2010:5).

At first, we were taken aback, our 0s are most often simply codes for the plesiomorphic state of a transformation series, for example: 0 codes for pectoral fin while 1 codes for foreleg. But then we realized that Mooi and Gill were using the Nelson–Platnick method of coding for three-taxon analysis, which considers character matches homologous only if they confirm a particular monophyletic group that is actually included in the analysis. Frankly, we had not considered three-taxon analysis particularly phylogenetic since reading the critiques of Kluge (1993) and Farris and Kluge (1998) as it seems to violate some of the basic tenets of parsimony analysis as practiced by phylogeneticists who apply Hennig's principles strictly. However, the attitude expressed is an interesting one that goes to the heart of whether you strictly try to separate pattern and process (as we believe Nelson and Platnick, 1981, advocate) or admit some modicum of evolutionary thinking into the process (as advocated by Kluge, 1993; and with which we agree). We will revisit this argument in Chapter 6 as we end our discussion of parsimony. Suffice to say, transformation, for evolutionists at least, is a necessary fact of nature (Farris and Kluge, 1998) and transformation is not captured by presence-absence coding when there is a reasonable plesiomorphic homolog available for a particular data column. The code "0" can be thought of simply as a code (it might as well be "a" and "b" rather than "0" and "1") and "0" does not have to imply "no information."

CHAPTER SUMMARY

- Characters are quasi-independent properties of organisms.
- Characters have part–whole relationships with the organisms of which they are properties.
- Shared characters have part–whole relationships with groups of organisms.
- Homologs that share an identity diagnose a monophyletic group at some level in the tree of life and may be termed taxic homologs.
- At any one level in the tree of life, transformational homologs comprise at least one plesiomorphic and one apomorphic state.
- One necessary (but not sufficient) property of transformational homologs is that they diagnose nested monophyletic groups.
- Testing homology is a multistep process that involves both Remane's and Patterson's criteria.
- Characters may be either qualitative or quantitative in nature.

- Although phylogeneticists usually prefer qualitative characters, modern morphometric methods offer an alternative for dealing with quantitative characters of shape.
- There are various ways to code characters, and whether to code complex characters in a single data column or separate them is not always clear.
- Missing data may or may not effect a phylogenetic analysis.

6

PARSIMONY AND PARSIMONY ANALYSIS

In the preceding chapter, we explored the concepts of characters and homology. In this chapter, we will use these concepts to demonstrate how phylogenetic problems can be analyzed using the principle of parsimony. This will be followed in the next chapter with a discussion of likelihood and Bayesian methods, which are usually described as statistical methods. We begin with a general consideration of parsimony, but will make mention of those aspects of parsimony analysis that are similar to likelihood, as the two are closely connected. We will then place this general discussion within the context of parsimony analysis.

PARSIMONY

The usual definitions of parsimony one encounters in English dictionaries concern money: extreme stinginess, extreme care in spending money, reluctance to spend money unnecessarily. Scientists and philosophers use a different version, usually attributed to William of Ockham (1285–1347) but in fact found in the works of Aristotle. The principle of parsimony is a methodological principle that posits simpler explanations of data relative to hypotheses are to be preferred over more complex explanations. This idea of simplicity relative to scientific hypotheses has been explored in some depth by Sober (1975), and the link between simplicity and parsimony has been discussed extensively in the phylogenetics literature (e.g., Wiley, 1975; Beatty and Fink, 1979; Farris, 1983). Both Farris (1983 and earlier works cited therein) and Sober (1983a) have linked simplicity with phylogenetic parsimony. Farris (1983), in particular, provides an extensive discussion of phylogenetic parsimony as a principle that leads to greater explanatory power of the resulting

phylogenetic hypotheses and argued (correctly in our view) that the role of parsimony was to minimize the number of ad hoc explanations embedded in the preferred hypothesis in the form of hypotheses of homoplasy. This provides a direct link with Hennig's auxiliary principle. We leave the philosophical justification of phylogenetic parsimony to the end of this chapter. For now, we are concerned with how phylogenetic parsimony works.

Parsimony: Basic Principles

1. Among the many possible phylogenetic trees that graphically portray the descent of three or more taxa, only one of these trees is correct, given that the taxa are natural entities. This principle is shared with other approaches (e.g., statistical approaches) and is a simple statement that the problem is historical.

2. Characters originate and become fixed over evolutionary time such that it is possible for ancestral species to pass on such characters to descendant species. Again, the principle is shared with other approaches.

3. The sharing of character states is always evidence that those taxa that share a character state are related unless the weight of other evidence dictates that they are of independent origin. This principle is unique to parsimony approaches, although it is possible to assign different confidence in the ability of other characters to influence the decision by assigning relative weight to the character.

4. Once a character appears and is fixed, there is no reason to postulate that it will change unless the weight of other characters dictates that change is necessary because tree topology changes. This principle differs from statistical approaches where there is always a probability of change built into a model of character evolution for each class of characters.

5. Characters (data columns, transformation series) are treated as independent in any analysis (Kluge's auxiliary principle; Brooks and McLennan, 2002) for purposes of testing character hypotheses. This principle is shared by most approaches for computational reasons.

6. The result of parsimony analysis consists of placing character states on a tree where they are thought to have originated or become fixed. Parsimony analysis is particularly "transparent" in this principle, but we can place states on trees using statistical approaches if we wish to do so.

7. The tree with the fewest number of independent origins of shared characters is the preferred solution. This is the maximum parsimony principle. Parsimony differs from other approaches because trees are evaluated based on minimum length—the minimum number of changes in characters that are hypothesized to have occurred for any particular tree hypothesis. Trees of minimum length fulfill the principle.

Parsimony, then, is built around the proposition that the "best tree" is the tree that describes the evolution of any particular set of characters using the smallest number of evolutionary changes of the characters analyzed. The question is: how do we do this? Below we will review parsimony methods, beginning with classic Hennigian

argumentation and proceeding to current optimality-driven algorithms. The progression of the chapter reflects the historical development of the methods used today.

Kinds of Parsimony

In Chapter 5, we discussed the various relationships that might obtain between character states (ordered, unordered, etc.). Two common forms of parsimony are directly related to how we treat the relationships between character states within a transformation series. Characters with only two states are treated the same in both common forms of parsimony, Fitch and Wagner parsimony, but may be treated differently in the uncommon forms of parsimony, Dollo and Camin-Sokal parsimony. We will review each briefly.

Fitch parsimony (Fitch, 1971). Fitch parsimony takes all characters as unordered (see Fig. 5.16a, d). When three or more character states exist (Fig. 5.16d), a reversal from two to zero or a transformation from zero to two is counted as a single step. This type of parsimony is commonly implemented when analyzing DNA base-pair data and multistate morphological characters.

Wagner parsimony (Farris, 1970). Wagner parsimony treats binary character states identically to Fitch parsimony. However, characters with more than two states are considered ordered (Fig. 5.16e). Thus a transformation from two to zero is counted as two steps because the only route from two to zero is to pass through state one.

"General" parsimony (Swofford and Olsen, 1990). General parsimony allows mixing of different kinds of parsimony in a single analysis following a generalized Sankoff approach. For example, Fitch parsimony might be used for some characters, Wagner parsimony for others, and a step matrix for others. "Informed parsimony" (Goloboff, 1998) is a form of general parsimony.

Two uncommon forms of parsimony are also recognized. Camin-Sokal parsimony (Camin and Sokal, 1965) imposes the constraint that evolution is irreversible. Once state 1 has appeared, subsequent transformation from 1 to 0 is not allowed. However, state 1 can evolve as many times as needed. Dollo parsimony (as implemented by Farris, 1977) allows a single change from state 0 to state 1 and as many reversals from 1 to 0 as needed to explain the data. Both of these methods work on rooted trees, in contrast to Fitch and Wagner parsimony, which can work on either rooted or unrooted trees.

With these distinctions in mind, we will examine different analytical approaches to parsimony analysis. The first, classic Hennigian argumentation, does not speak directly to the different forms of parsimony outlined above because the classical approach was performed in the absence of the data matrix and a specific numerical algorithm. Nevertheless, it is directly connected to more modern methods by Hennig's auxiliary principle.

CLASSIC HENNIGIAN ARGUMENTATION

Classic Hennigian argumentation was practiced long before the advent of computer-assisted analysis. It is founded on the proposition that the investigator makes a priori decisions on relative synapomorphy and groups based on those decisions. As

such, it is in the class of algorithmic approaches. Although issues such as parsing out different kinds of parsimony were not part of the discussion prior to the advent of computer-assisted analyses, we can think of Hennig Argumentation as a form of nonexplicit general parsimony. It is based on three rules (see Brooks and McLennan, 2002).

1. *The Grouping Rule.* Characters deduced as synapomorphies are evidence of unique common ancestry while symplesiomorphies and homoplasies are of no use in determining unique common ancestry (Hennig, 1966). How one determines which of two homologous characters is apomorphic is the process of "polarizing the transformation series" and is discussed in the next section.
2. *The Inclusion/Exclusion Rule.* The information from two transformation series can be combined into a single hypothesis of relationship (unique common ancestry) if the valid evidence (i.e., the synapomorphies) implies the identical grouping or allows for the complete inclusion or exclusion of groups implied by the valid evidence. This rule is illustrated in Fig. 6.1.
3. *The Homoplasy Rule.* If the information from two transformation series results in groupings that overlap or conflict, then one and possibly both putative hypotheses of synapomorphy are false at the level used. Either one or both are not homologs, or one or both are incorrectly polarized.

It is rare to see a phylogenetic paper these days that employs classical, precomputer analysis. This rarity does not mean that classical analyses are an invalid approach, but it does signal that complexities of analysis are usually greater than the ability of an investigator to consider all of the possible phylogenetic hypotheses that might be inferred from the data. In spite of this, it is worthwhile to understand phylogenetics at a level where we can consider the meaning of such practices as character polarization and phylogenetic argumentation on a character-by-character basis. Whether one uses computer-assisted analyses or not, the quality of the initial
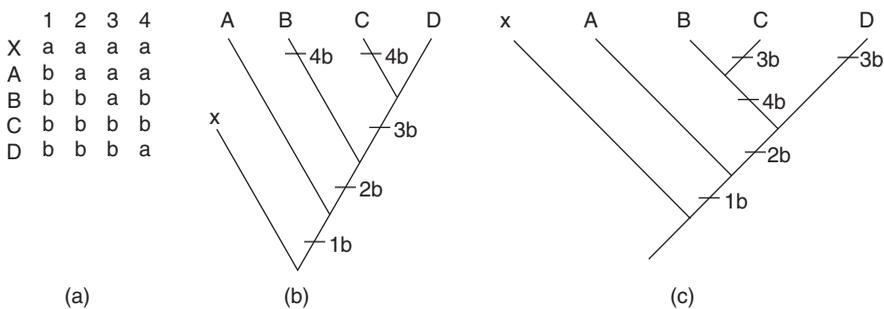
Figure 6.1. Simple examples of the inclusion/exclusion and homoplasy rules. (a) A data matrix with states coded "b" polarized as apomorphic based on the presence of states coded "a" in the outgroup X. (b) The argument that character states 2b and 3b confirm the monophyletic groups BCD and CD but excludes BC as a monophyletic group confirmed by 4b (4b in B and C is homoplastic). (c) If we accept the argument that 4b confirms the monophyletic group BC, then we must exclude 3b as a synapomorphy confirming CD and consider 3b homoplastic. Note that in both (b) and (c) the state 2b confirms the monophyletic group BCD and thus can include either hypothesis.

characters brought to the analysis is critical to the final result and a priori methods depend on such quality.

Polarization

Two character states are said to be polarized when we have determined which evolved first and which one came after. Thus, polarization refers to determining which one of two or more hypothesized states is plesiomorphic and which one (or ones) is apomorphic. The initial assumption is that all instances of the states are actually homologous, but we may find, using congruence, that they may not be homologous when we accept the homologies of other characters and their states. In fact, we may find nonhomology among instances of a state; parallel or convergent appearance of characters that share a nonhomologous identity. Thus, we are not committed to claiming that the characters and their states will turn out to be homologous, only that we will assume so for purposes of testing that very proposition.

In the example presented in Fig. 6.1, we assumed that we knew the homology and polarity of the states in advance, and then applied our three rules. In this section, we will explore ways of polarizing character states. This activity lies at the heart of the phylogenetic method, whether it is done by hand, a priori, or by rooting, a posteriori. Phylogeneticists rarely polarize characters a priori these days. Instead, an investigator relies on various computer algorithms to polarize characters by using one or more outgroups designated by the investigator to perform this task. However, understanding the reasoning behind polarization is important relative to the history of the discipline and also because it shows why computer-assisted analysis can arrive at a robust hypothesis only when given the best outgroup information possible. Further, if one is going to order more than two character states, one is performing a priori polarization, and thus the principles are vital. Many criteria for polarizing character states have been proposed, including several by Hennig (1966) himself. There is a general consensus (with some significant dissenters) that there is only one general criterion, outgroup comparison. We will discuss this criterion and then some of the alternatives.

Polarization by Outgroup Comparison. Consider a character with states distributed such that there is variation in the group under analysis. For example, among land plants, mosses and tracheophytes have xylem tissue while hornworts have undifferentiated parenchyma cells. Which is the apomorphic character? The closest relatives of hornworts, mosses, and tracheophytes are the liverworts. Liverworts have undifferentiated parenchyma cells. If certain assumptions, detailed below, are met, this observation leads to the conclusion that xylem is apomorphic relative to undifferentiated parenchyma cells. This is reinforced by the observation that the closest relatives of land plants, groups such as stoneworts (e.g., *Chara*), also have undifferentiated parenchyma cells rather than xylem. If we examine a tree of land plant evolution, we see that the most likely hypothesis of polarity is that xylem evolved sometime after the origin of hornworts but before the origin of mosses (Fig. 6.2).

This kind of reasoning is deductive, and the validity of the conclusions depends on certain assumptions. First, one must accept the monophyly of the group comprising hornworts, mosses, and tracheophytes, establishing the ingroup, which is the group one wishes to analyze. Second, one must accept the monophyly of all land plants (liverworts and above) to establish a rational sister group. Third, one must
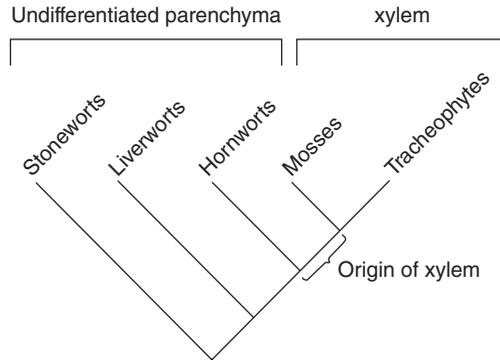
Figure 6.2. A hypothesis of plant relationships. Given this topology, the transformation of undifferentiated parenchyma cells to form xylem tissue is more parsimonious than the transformation of xylem to undifferentiated parenchyma because the close relatives of mosses and tracheophytes have undifferentiated parenchyma cells.

accept the hypothesis that stoneworts (and brittleworts and other filamentous green algae) are related to land plants. In other words, such deductive reasoning is accomplished with the acceptance of prior information. Because it depends on prior information, the conclusions will be valid only if the prior information is correct. With this in mind, the outgroup rule of polarization may be simply stated.

The Outgroup Rule. Given two (or more) homologous character states within a group studied, the state found outside this group in close relatives is the plesiomorphic state and the character found only within the group is the apomorphic state.

An explicit statement of the outgroup rule is, curiously, missing from Hennig (1966). However, and at least for binary characters, it is apparent to us that Hennig used outgroup comparison, as evidenced by the following quotes.

> Recognition that species or species groups with common apomorphous characters form a monophyletic group rests on the assumption that these characters were taken over from a stem species that only they have in common, and which already possessed these characters prior to the first cleavage (Hennig, 1966:90).

> [I]f it is a question of determining the relationships between different species groups, then it is of primary importance to show that each group has apomorphous characters, characters that are present only in it (Hennig, 1966:90).

Both of these statements imply Hennig used a comparative outgroup method. One could hardly reach the conclusion that a character was only found in the stem species of a group without examining species outside of the group. And no one could claim that a character is unique to a group without looking at other groups. Hennig (1966:95–116) discusses "accessory criteria" when considering "morphoclines," characters of more than two states. The fact that he characterizes them as "accessory" relative to the "scheme of argumentation of phylogenetic systematics" suggests that what Hennig considered strong evidence of monophyly were characters unique to a given group, which can only be deduced if one looks outside the group. This emphasis on sister groups sharing unique homologies is common in early

phylogenetic literature (e.g., Brundin, 1966). Uniqueness can only be accessed by looking broadly across groups. The criterion of what we now know as outgroup comparison was also well understood by early quantitative phylogeneticists, forming one of three criteria used by Kluge and Farris (1969).

The logic of implementing the outgroup rule was discussed by Wiley (1975) within a Popperian framework, but Wiley did not characterize looking outside the group with the formal designation of "outgroup" for those taxa consulted, but characterized the addition of new taxa (those we now think of as outgroups) as raising the level of universality of the problem (which is exactly what outgroups do). The actual origin of the term *outgroup* is not of particular consequence, because it is the principle, not the name, that is important. It is possible that the term originated in print with Wiley (1976:11):

> Hennig's (1966) method differs fundamentally from a purely phenetic method in that all the shared characters are not used to refute a given relationship; only synapomorphous characters are used. Such testing can only be accomplished in an open system, that is, by considering taxa outside the three (or more) taxon system. Such considerations may be termed *outgroup* [emphasis added] comparisons. The one condition placed on this procedure is that the three (or more) taxa must form a monophyletic group. The designation of outgroups for comparison permits an investigator to sort out which of the observed characters are unique to the three-taxon system and which characters have a more general distribution. The *outgroup* [emphasis added] comparison automatically raises the level of universality of the phylogenetic hypothesis to a new level. And, it allows the investigator to put his three-taxon statement in context with a hypothesis of a higher level of universality.

By the early 1980s, specific descriptions of character argumentation using the term *outgroup* were appearing (c.f., Eldredge and Cracraft, 1980) and specific forms of the Outgroup Rule were published (Wiley, 1981a; Watrous and Wheeler, 1981). However, the complexities of polarization using outgroups were best demonstrated by Maddison, Donoghue, and Maddison (1984). They demonstrated that the simple rules formulated in earlier works were not adequate. Their work also demonstrated that criteria such as "common is primitive" could be dismissed as fallacies. Maddison et al. (1984) begin by defining terms, illustrated in Fig. 6.3.

Ingroup. The group under analysis. In Fig. 6.3a, the ingroup is shown as a polytomy, suggesting unresolved relationships.

Ingroup Node. The trees used by Maddison et al. (1984) are node-based trees (vertexes are taxa), not stem-based trees (edges are taxa), so the internal nodes are ancestral species and the edges are relationship statements. The ingroup node represents the character states of the ingroup ancestor, that is, the ancestor of the group under analysis. Some of these character states will be synapomorphies for the ingroup, others will be plesiomorphies, and some cannot be polarized. Polarization depends on the distribution of the character state and its homolog(s) among the outgroups.

Outgroup and Sister Group. Any clade that is attached to the edge leading to the ingroup node is an outgroup. That clade immediately below the ingroup node is the relative sister group of that particular analysis.
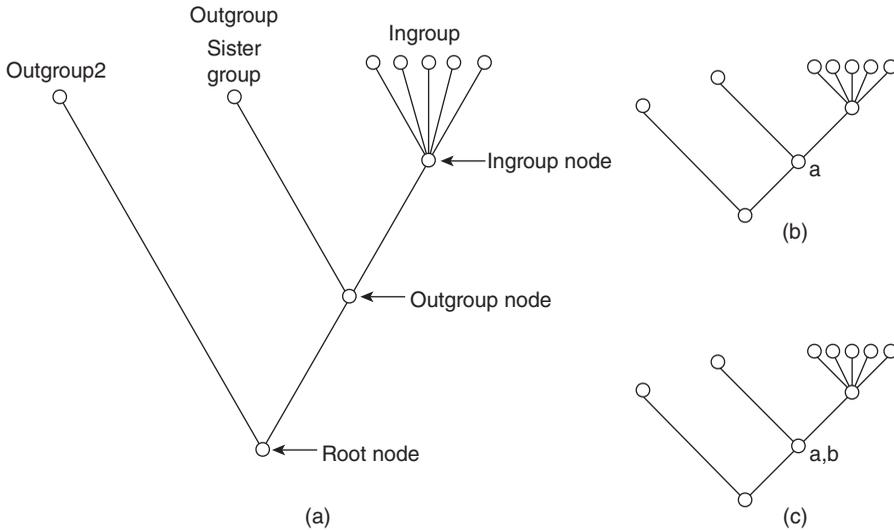
Figure 6.3. Basic terminology of parts of a Hennig tree following Maddison et al. (1984). (a) The ingroup node represents the ancestral species of all members of a group under analysis, the ingroup. The outgroup node is the node that refers to the ancestor of the ingroup node while the root node represents the most basal, or ancient, ancestral species. The sister group constitutes the closest outgroup to the group analyzed that is known or included in the analysis and constitutes the first outgroup (labeled outgroup 1). It is composed of one to many species, and its character states represent the character states inferred for the ancestral species of all members of the sister group. The outgroup 2 is simply the next known closest relative. In general, a minimum of two outgroups are needed to polarize a character *a priori* unless the sister group is considered entirely plesiomorphic (a bad assumption to make). (b) A decisive character decision in favor of state "a" of a character. (c) An equivocal decision for a two-state character. Note that these decisions are made at the outgroup node, not the ingroup node.

Outgroup Node. The node immediately below the ingroup node is the outgroup node. A character assigned to the outgroup node would be the character hypothesized to be present in the ancestor of the ingroup and its sister group.

Root Node. The most basal node in the tree.

Maddison et al. (1984) frame the polarity problem as a quest for the assignment of characters to the outgroup node. Why this is so is immediately apparent if we give it a bit of thought. We wish to arrive at two classes of hypotheses in our analysis. First, we seek evidence that the ingroup is monophyletic. Second, we wish to uncover evidence for monophyly of subgroups within the ingroup. Evidence that the ingroup is monophyletic can only be gained by accessing the character states present in the immediate common ancestor of the ingroup and its sister group. By determining the character states present at the outgroup node, we are able to either make this decision or know that the information is not adequate to make this decision, as we shall see below. Decisions at the outgroup node can be of two kinds. Given two homologous character states, if only one is assigned to the outgroup node, the polarity
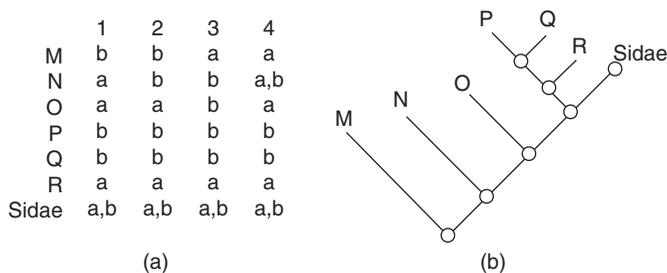
Figure 6.4. The Maddison et al. (1984) method of character polarity I. (a) The character matrix. (b) The Hennig tree of relationships of the outgroups to the ingroup. Note that this tree topology is given *a priori*; confirming characters for this topology may be totally missing in the character matrix. Also note that the ingroup is polymorphic for each character, a necessary condition for analyzing relationships among the ingroup. Redrawn from Wiley et al. (1991), used with permission, Biodiversity Institute, University of Kansas.

decision is **decisive** (Fig. 6.3b). If neither can be confidently assigned to the outgroup node, the decision is **equivocal** (Fig. 6.3c).

A simplified example of how the algorithm works is shown in Figs. 6.4 and 6.5 (taken from Wiley et al., 1991). We will use the binary transformation series, although Maddison et al. (1984) provide a general (and more complicated) algorithm for more than two character states.

Consider the Sidae, its outgroups, and character variation (Fig. 6.4a). Prior knowledge from other studies hypothesized a specific outgroup structure (Fig. 6.4b). This tree is not justified by the characters in the matrix because the taxon of interest is Sidae and the analysis of Sidae to its sister group and other outgroups is not a matter for testing (this may or may not always be a wise choice). Proceed in the following manner.

1. Proceeding from the most distant branches, label nodes on the tree according to the following rules. If all terminal taxa have the state "a," then label the node decisive "a." If all terminal taxa have the state "b," then label the node decisive "b." If one or more terminal taxon has "a" and one or more terminal taxon has "b," then label the node equivocal "a, b."
2. Proceeding, again, toward the outgroup node, label the next node with the majority character derived from the state that lead to that node. For example, if a node has the equivocal decision "a, b" and a terminal has "a," then the assignment of the next node is decisive "a." If both have "a, b", then assign "a, b."
3. Proceeding from all parts of the tree to the outgroup node, make decisions for each node until the outgroup node is reached.

These calculations are carried out for the first character in Fig. 6.5a and for the second character in Fig. 6.5b. If you perform these operations on a sufficient number of trees, you will notice that the sister group is the most influential group in the entire analysis, unless it is polymorphic. If the sister group has a single character, or if the decision at the most basal node within the sister group is decisive, then this character will appear at the outgroup node (Maddison et al., 1984). Maddison et al. (1984)

Figure 6.5. The Maddison et al. (1984) method of character polarity II. (a) An example of a decisive decision for character 1 of Fig. 6.4a. (b). An equivocal decision for the states of character 2 of Fig. 6.4a. Redrawn from Wiley et al. (1991), used with permission, Biodiversity Institute, University of Kansas.

noted two other general outcomes. First, if the sister group and the next sequential outgroup have the same state, then that state will always be decisive at the outgroup node (the *First Doublet Rule*). Second, if states alternate down the outgroup topology and if the sister group has the same state as the most basal outgroup, then that state will be decisive, but if the most basal group has a different character, then the decision will always be equivocal (the *Alternating Outgroup Rule*).

You can use such reasoning in a traditional phylogenetic analysis by preparing a matrix of ingroup and outgroup taxa, adopting an outgroup tree topology, reasoning through each polarity decision and following the grouping, inclusion/exclusion, and homoplasy rules. As we shall see in later sections, computer-assisted phylogenetic analysis does not make a priori polarity decisions. So, why in the modern age do we cover this topic? Although computer-assisted analysis does not make a priori character decisions, phylogenetic computer-assisted studies call for a priori designation of the outgroup(s) to be used to polarize the states once direction of transformation is specified through designation of an outgroup (always included in the analysis). Thus, the Maddison et al. (1984) paper is very applicable to general phylogenetic reasoning using more modern techniques of phylogenetic analysis (parsimony and statistical algorithms) for three reasons:

1. It demonstrates the importance of careful attention to identifying or discovering the sister group.
2. It calls attention to the fact that a single sister group is not sufficient to unambiguously polarize a character; the minimum for analysis is the sister group and one additional relevant outgroup, hopefully the next sister group down the tree.
3. It destroys the notion that if you do not know the sister group you can simply make a decision that the character state commonly found in some array of possible outgroups is the plesiomorphic character. This third point may call for the imposition of a particular outgroup topology prior to analysis or to the inclusion of characters that are not particularly relevant to the ingroup problem per se, but that give structure of the relationships of the outgroups to each other and to the ingroup and are possible synapomorphies of the ingroup

itself (as suggested by both J. S. Farris and D. L. Swofford to Maddison et al., 1984:99).

What if you don't know the relationships of the outgroups to the ingroups? Indeed, what if you have no a priori evidence that the ingroup is even monophyletic? Such a case calls for the solving of a larger problem and may call for a community global approach. For example, the ichthyological community has been working on the teleost tree of life in a phylogenetic framework for some forty years. Although many studies of smaller clades have been successfully pursued, the emphasis has been on working from the root of the teleost tree toward the tips. This approach creates outgroup structure *with* the flow of evolutionary time rather than *against* the flow (e.g., Greenwood at al., 1973; Stiassny et al., 1996, 2004).

The analysis of *Leysera* presented below illustrates a relatively simple application of classic Hennigian argumentation, but with considerable attention paid to the problem of identifying a suitable outgroup. This will be followed by an account of more current approaches of analyses using computer algorithms where character polarity is determined *a posteriori* using an optimality criterion.

Example 1. The Phylogenetic Relationships of *Leysera*

Leysera is a small group of four species of composite shrublets. Three species (*L. gnaphalodes*, *L. tenella*, and *L. longipes*) are found in southern Africa. One species (*L. leyseroides*) is found in the Mediterranean region. As a continuation of his study on other closely related genera, Bremer (1978a) analyzed this group.

Background Information. *Lysera* (Fig. 6.6) is a member of Compositae, tribe Inuleae. Merxmüller et al. (1977) placed *Leysera* into the *Athrixia* genus group (eight genera) within the subtribe Athrixiinae (23 genera total).

Bremer (1978a) supported the monophyly of four of the eight genera of the *Anthrixia* group on the basis of leaf and involucre characters: all have ventrally furrowed and pubescent leaves and wide, yellowish brow, and scarious involucral bracts. These characters are "a most uncommon feature" in Athrixiinae, uniting *Leysera*, *Antithrixia*, *Relhania*, and *Rosenia*. Of the four, *Antithrixia* has a pappus with many barbellate bristles compared to the three other genera, which have a reduced number of bristles as well as a complete loss of bristles on the ray-floret pappus (Fig. 6.7). Finally, Bremer (1978a) observed that only species of *Leysera* have a solitary capitula on a long peduncle (Fig. 6.6) whereas the other three genera have sessile capitulas with the exception of some species of *Relhania* (which Bremer interpreted as homoplasy based on the monophyly of *Relhania*). At this point, Bremer has established the following background information (Fig. 6.8):

1. *Antithrixia*, *Leysera*, *Relhania*, and *Rosenia* comprise a monophyletic group. Justification is via outgroup comparison and character rarity.
2. *Antithrixia* is the sister genus to the remaining genera, which form a monophyletic group (outgroup comparison).
3. *Leysera* is monophyletic (outgroup comparison).

Bremer has assumed that the *Anthrixia* genus group is monophyletic. He has also assumed that rarity in morphological characters among outgroups is evidence of
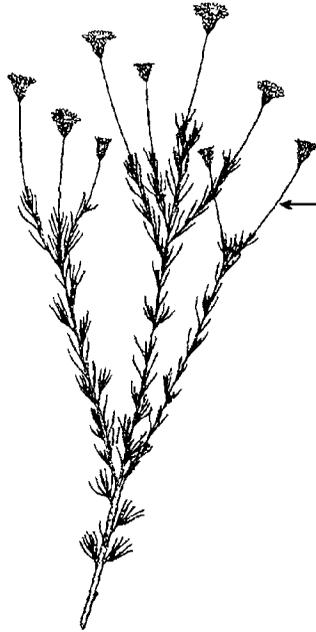
Figure 6.6. The composite plant *Leysera gnaphalodes*, illustrating the long peduncle typical of the genus (arrow). From Bremer, 1978a. Used with permission of Botaniska Notiser.
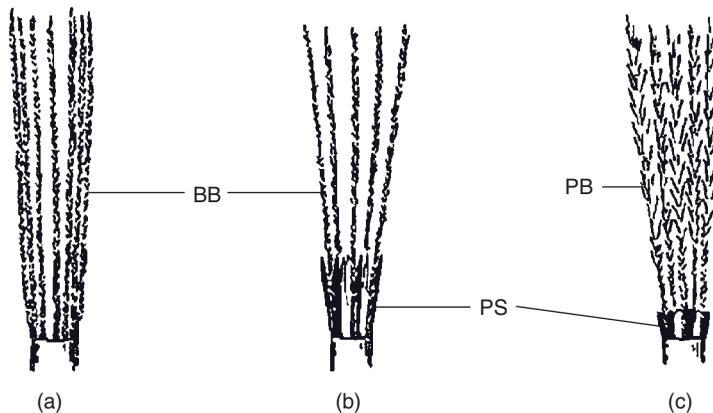
Figure 6.7. Features of the disc-floret in (a) *Antithrixia*, (b) *Leysera longipes*, and (c) *Leysera tenella*. Abbreviations: BB, barbellate bristles; PB, plumose bristles; PS, scales. Transformation of states in both characters proceed left to right. From original drawings by Kåre Bremer included in Bremer, 1978a. Used with permission of the author and Botaniska Notiser.

homoplasy for certain characters. (Note: this is pre-Maddison et al., 1984, and such assumptions were common.)

Bremer's (1978a) analysis involves two possible sister groups (*Relhania* and *Rosenia*). Further, there is a problematic taxon, "*Leysera*" *montana*. This species has solitary capitulae on long peduncles, like *Leysera*, but has a pappus with many
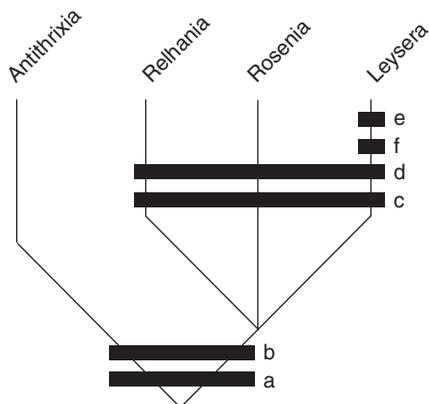
Figure 6.8. The phylogenetic relationships of *Leysera* and closely related genera. Synapomorphies are (a) leaves ventrally furrowed and pubescent; (b) involucral bracts wide, yellowish brown, and scarious; (c) floret pappus with scales but no bristles; (d) disc-floret pappus with reduced bristles and no scales; (e) solitary capula on long peduncle; and (f) chromosomes 2N = 8. Adapted from Bremer, 1978a.

barbellate bristles, like *Antithrixia* and other Athrixiinae outside the clade. Bremer removed this species from *Leysera* and later described the monotypic *Oreoleysera* for it (Bremer, 1978b) because *O. montana* did not have the synapomorphies that would place it in the monophyletic clade containing the three genera, even though it had the character that unites *Leysera*, forcing Bremer to conclude that the match was a homoplasy.

Given that there are four species of *Leysera*, a total of 15 rooted bifurcated trees are possible. Bremer (1978a) analyzed 13 characters for the four species using what is now referred to as outgroup comparison relative to the two possible sister groups, *Relhania* and *Rosenia*. Among the states were five autapomorphies, four in *L. longipes* and one in *L. leyseroides*, which will not be discussed further. In addition to the analysis presented above that demonstrates the monophyly of *Leysera* and its relationships to its relatives, two additional levels of synapomorphy analysis are required to complete the analysis.

1. Establishing the basal member of the species group.
2. Breaking up the remaining trichotomy.

One character from each level will be discussed. The complete table of characters is shown in Table 6.1, and the tree of relationships is shown in Fig. 6.9.

Level 1. Character 1; Receptacle smooth versus rough. Within *Leysera* there are two character properties of the receptacle. In *L. longipes* the receptacle is more or less smooth, without scale-like growths. In the remaining three species the receptacle is rough, and this roughness is caused by scale-like outgrowths. Receptacles with scalelike growths are not known in species of *Relhania* or *Rosenia*, nor found in any other members of the *Athrixia* genus group. They are known from less closely related genera of composites. To argue that the rough receptacle of the three species

TABLE 6.1. Characters used by Bremer (1978a) to analyze the phylogenetic relationships of *Leysera*. Autapomorphies are not listed. Characters are shown in the hypothesis presented in Fig. 6.9. All determinations were made by outgroup comparison.

Character	Plesiomorphic	Apomorphic
1. Receptacle	Smooth	With scalelike growths
2. Floret tubules	Glands present	Hairs present
3. Pappus	Barbellate	Plumose
4. Achenes surface	Smooth	Cells imbricated
5. Pappus scales	Subulate	Wide and flat
6. life cycle	Perennial	Annual
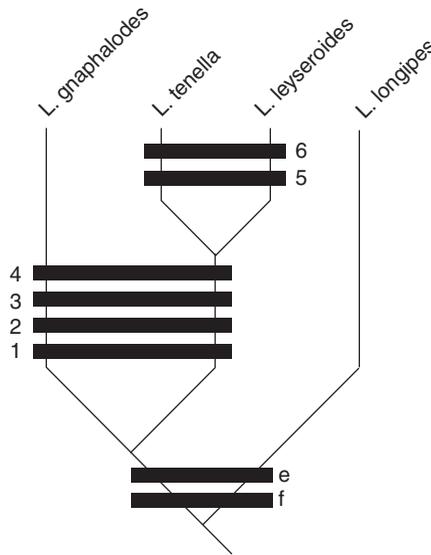
Figure 6.9. Bremer's (1978a) hypothesis of the relationships among species of *Leysera*. Character numbers correspond to the apomorphic state of transformations in Table 6.1. The two synapomorphies of the genus correspond to the states in Fig. 6.8. Autapomorphies for each species are not shown.

was convergent because a similar condition was found in distant relatives is a violation of the auxiliary principle. To argue that these three species are not members of the *Athrixia* genus group would require rejection of all synapomorphies that place *Leysera* within the group and united to *Relhania* and *Rosenia*. Bremer (1978a) concluded that the rough receptacle was a synapomorphy uniting *L. leyseroides*, *L. tenella*, and *L. gnaphalodes* by outgroup comparison and parsimony.

Level 2. Character 5; pappus scales (Fig. 6.7b, c) subulate versus wide and flat. Wide and flat pappus scales are found in *L. tenella* and *L. leyseroides* while subulate scales are found in *L. longipes* and *L. gnaphalodes*. Given the four synapomorphies that unite *L. gnaphalodes* with *L. tenella* and *L. leyseroides*, and given the monophyly of *Leysera*, we can use *L. longipes* as a "functional outgroup" to polarize the

transformation series. The use of functional outgroups is discussed by Watrous and Wheeler (1981). In short, once the monophyly of a group is established, an investigator can create "functional outgroup comparisons" by comparing the basal member(s) of the group with more apical members of the group (see also Wiley, 1981a:175–176).

A POSTERIORI CHARACTER ARGUMENTATION

There is another way to argue characters, and it is a basic aspect of more modern phylogenetic analyses. If you consider that all characters are freely reversible, and that they are not fated to ratchet ever forward, then it turns out that we can assemble a tree without a root and without polarization whose topology is logically consistent with a rooted topology determined by a priori character argumentation. The characters on a rootless tree have no phylogenetic interpretation, but we can give them such an interpretation if we specify the starting point, an activity termed *rooting the tree*. This is useful, because it provides a bridge between computer-assisted phylogenetic analysis and traditional phylogenetic analysis. Examine the character matrix and unrooted tree in Fig. 6.10a, b. Note that we have made no judgments of character polarity, we have simply plotted the characters coded "b" along branches where they occur. If we root the tree along the edge leading to E (Fig. 6.10c), then all of the characters coded "b" appear on the rooted tree as apomorphies. But if we root on the edge leading to D (Fig. 6.10d), then most of the apomorphies are those characters coded "a." If you count the total number of possible changes, you will note that both trees are the same length, TL = 7 steps. Rooting changes both the topology and the character polarity interpretations, but it does not change the tree length.

ALGORITHMIC VERSUS OPTIMALITY APPROACHES

Swofford and Olsen (1990) and Swofford et al. (1996) discuss a useful distinction between two approaches to phylogenetic inference. Algorithmic approaches combine tree inference and definitions of the preferred tree into a single operation that defines a sequence of steps that lead to the determination of a tree. Evolutionary assumptions are embedded in the algorithm and used for the analysis. In contrast, optimality approaches define an objective function such as minimum tree length (parsimony) or maximum likelihood (ML), use an algorithm to generate trees, and then sort trees with a preference for that tree that meets the objective function. In this approach, the objective function embodies the evolutionary assumptions and any algorithm that generates trees can be used because the result is not dependent on evolutionary assumptions embodied in the algorithm, but rather, the result is dependent on whether one tree (or set of trees) meets the objective criterion better than another tree (or set of trees).

In algorithmic approaches, the algorithm is important because it defines the selection criterion and combines the inference and the criterion for the preferred tree into a single operation. Examples include classic Hennigian argumentation (Hennig,
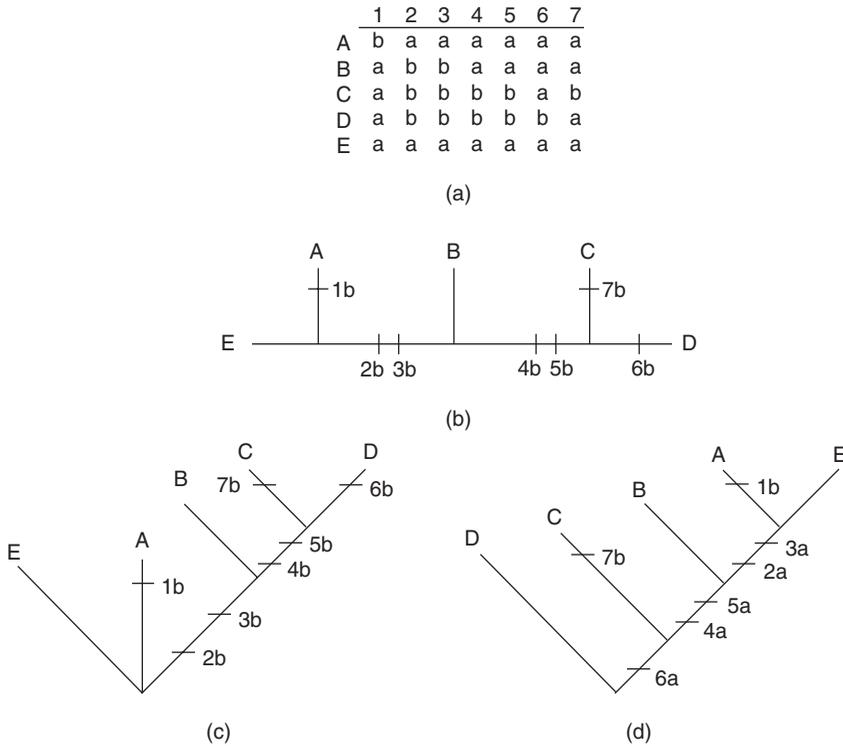
Figure 6.10. Rooting an unrooted tree. (a) A hypothetical data matrix. (b) The unrooted tree that minimizes the number of transformations needed to account for the changes shown in the matrix. (c–d) Two rooting decisions. Note that while the character polarities are much different, the lengths of the trees are the same (7 steps).

1966), Wagner ground plan divergence analysis (Wagner, 1961), most first-generation computer algorithms such as the Wagner Algorithm (Kluge and Farris, 1969), and some distance algorithms in current use such as neighbor joining (Saitou and Nei, 1987; Studier and Keppler, 1988). Algorithmic approaches are fast in terms of computation time and are likely to find trees that are close to optimal (or even optimal if the data are fairly clean). Their speed and efficiency make them excellent for building a tree hypothesis, but they can become stuck in local optima depending on the starting conditions and the nature of the data.

   In optimality approaches, the investigator specifies an objective function and then uses an algorithm to compute that function for a particular tree topology. It then computes that same function for another tree and compares the trees. In parsimony the shortest tree "wins." How one obtains trees to compare is not relevant to the process, but simply to the efficiency of the search. Examples of computer packages that implement the optimality approach include all modern parsimony programs; several generations of Wagner programs (i.e., Hennig86, Farris, 1989a; PAUP, Swofford, 2001; PHYLIP, Felsenstein, 2007; NONA, Goloboff, 1999a; TNT, Goloboff et al., 2000).

Although fast, algorithmic-driven programs suffer from a problem. They compute a tree well enough, but the investigator might miss other trees that are just as good or very close to the tree computed. And the investigator has no ready way to compare the robustness of the results relative to other possible outcomes. In contrast, optimality-driven programs are slower, but can search many trees and return the results for all of the trees that fit the objective function and even those trees that might not meet the objective but are some specified distance from it. For example, if the objective function was maximum parsimony in terms of steps and there were 5 trees of length 100 steps, an exhaustive optimality search would return all 5 trees. Further, if the investigator wishes to also examine trees within 10 steps of the shortest trees, then all trees from 90–100 steps would be returned.

OPTIMALITY-DRIVEN PARSIMONY

In most current computer packages, one might begin the analysis by constructing a tree using, for example, the Wagner Algorithm or neighbor joining. However, most of the actual computing time is spent evaluating different tree topologies (branching patterns) to recover the tree(s) that meet a criterion of optimality given the data. How the tree is actually generated may be irrelevant. For example, you can evaluate all of the possible trees for a three-taxon problem by simply mapping the character distributions on the four possible trees in the most efficient manner (i.e., maximizing the number of synapomorphies and minimizing the number of homoplasies needed given the tree). You don't have to build a tree; all of the possible trees are given. Under the criterion that the shortest tree is the optimal tree (the objective function in parsimony analysis), all you have to do is count the changes and pick the shortest tree(s) among the four possibilities. Polarity is not determined a priori, but a posteriori through the designation of one or more outgroups. This is because the algorithm first computes an unrooted tree and then roots the tree at the point designated by the investigator. Although this may sound strange to classical phylogeneticists, the trick is to understand that solutions involving freely reversible characters yield a network that is logically consistent with a rooted tree that could be found using a priori character polarization in reference to the same outgroup. In essence, this is why parsimony analysis using computer algorithms is the same research program as parsimony analysis using classical Hennigian argumentation. The difference is this: as more and more taxa are analyzed and as homoplasy levels increase, the less the chance that classical Hennigian argumentation will yield all of the equally parsimonious solutions, or even the single most parsimonious solution. The order of taxa added to the analysis might lead the investigator into a local optimum. Some possible solutions for dealing with suspected homoplasy might be missed. There can also be problems for optimality approaches using computers, but they cover more ground in the hunt.

In parsimony analysis, the optimality criterion is tree length. The tree topology that minimizes the number of evolutionary steps needed to explain the evolution of characters in the matrix is the optimal tree, given the data. Other sorts of optimality criteria are possible. For example, ML also has an optimality criterion: the tree topology and evolutionary model applied that maximizes the probability of observing the data is preferred (see Chapter 7). Any particular algorithm is a method for

estimating the optimal tree given a particular criterion. As Swofford et al. (1996) stress, algorithms change and improve while optimality criteria may not. Thus, modern parsimony methods concentrate on:

1. Fitting characters on particular tree topologies such that the number of evolutionary steps is minimized within each transformation series.
2. Comparing the results obtained among tree topologies to determine which tree (or set of trees) is the shortest.
3. Visiting many possible trees in an effort to avoid locally optimal solutions.

DETERMINING TREE LENGTH

Kluge and Farris presented the algorithm for tree length with ordered characters. Fitch (1971) presented the algorithm for determining tree length in the case of unordered characters. Sankoff (1975) generalized the algorithm for general parsimony. There are two basic algorithms for determining tree length in the absence of a step-matrix or other weighting schemes. Each follows one of the two common parsimony approaches: Wagner parsimony (ordered characters) and Fitch parsimony (unordered characters). Each requires a single down-pass through the tree. Note that tree length is computed on unrooted trees. Also, although we begin with taxon A in our example, tree length can be computed from any starting taxon, as discussed above.

Tree Length under Ordered (Wagner) Parsimony. We will present the example used by Swofford et al. (1996) informally; that is, avoiding as much set theory and formal algorithms as possible to show the general method for a single transformation series. Consider an unrooted tree with five taxa, and a single ordered transformation series (Fig. 6.11a).

1. Root the tree with a terminal node. For each terminal node, assign the characters it has based on the input matrix (Fig. 6.11b). This is the taxon's character set.
2. Proceed from the tips toward the terminal node, and assign characters to each interior node (labeled X, Y, and Z in Fig. 6.11c) according to two rules.
   2a. If the intersections of the character sets of descendants is not empty, then let the character set of the ancestor equal the intersection as a closed interval.
   2b. If the intersection of the state sets is empty, let the character set of the ancestor equal the smallest closed interval containing an element from each set. Increase tree length by the length of the interval (the difference between the end points of the interval).
3. If the internal node is adjacent to the root node of the tree (immediate descendant of the root node), then go to step four, otherwise return to step two.
4. If the character of the root node is not contained in the character state of its descendant node, then increase tree length by the shortest distance between them.
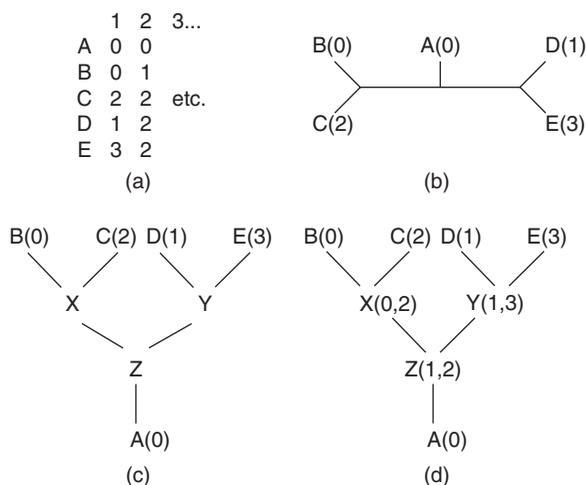
Figure 6.11. Calculating tree length for a single character. (a) A polarized transformation series with "0" as the plesiomorphic state. (b) An unrooted tree showing the distribution of transformations of character 1. (c) The tree rooted with taxon A. (d) Assignment of states to the interior nodes.

For the interior nodes in Fig. 6.11d, the values shown are computed below. Note that Z is the node at the basal fork and A is the root node.

1. X: $[0] \cap [2] = \emptyset$. Thus X = $[0, 2]$. English translation: the intersection, $\cap$, of "0" and "2" is empty, thus the state set of X is the interval $[0, 2]$. This follows Rule 2b. Increase the length by two steps.
2. Y: $[1] \cap [3] = \emptyset$. Thus Y = $[1, 3]$. This follows Rule 2b. Increase the length by two steps. Tree length now equals 4.
3. Z: $[0, 2] \cap [1, 3] \neq \emptyset$. Thus Z = $[1, 2]$. This follows Rule 2a.
4. The state set of A = $[0]$, while that of Z = $[1, 2]$. Thus we increase tree length by one step.

Tree length under Wagner parsimony is TL = five steps. We would then perform the same operations on the next column of data and add the results to our count, adding columns until we reach the end of the matrix.

Tree Length under Unordered (Fitch) Parsimony. If you wish to calculate tree length under Fitch parsimony, we modify the algorithm slightly:

2a. If the intersections of the state sets of descendants is not empty, then let the state set of the ancestor equal the union of the intersection.
2b. If the intersection of the state sets is empty, let the state set of the ancestor equal the union of the state sets and increase tree length by one step.
4. If the state set of the root node is not contained in the state assigned to the basal fork of the tree, then increased length by one step.

In our example, X is assigned the character set $[0, 2]$ and the tree length is increased by one step. Y is assigned the character set $[1, 3]$, and the tree length

is increased by one step. Finally, Z is assigned the intersection of [0, 2] and [1, 3], which is [1, 2] and because this does not intersect with A (with character 0), the tree length is increased by one step. Thus, tree length under Fitch parsimony is TL = three steps.

Tree length may be further modified if weight is given to an entire transformation series or if a step-matrix is used for one or more transitions within a transformation series (although the use of step matrices requires a general parsimony procedure because algorithms for both ordered and unordered characters do not apply). As a simple example, if we assigned a weight of 100 to the first data column in our example, then the length would be 500 under Wagner parsimony and 300 under Fitch parsimony.

FINDING TREES

Fitting characters to a tree is a relatively easy procedure. Equip yourself with a program, and input a tree for any particular matrix of characters. Then have the program optimize the characters on the specified tree using one of the optimization criteria discussed later in this chapter. It takes almost no time to accomplish this task. The harder trick is to find the optimal tree (in parsimony, the shortest tree). There are several strategies to do this, depending on the number of taxa and the size and complexity of the data matrix.

Strategy 1: Exhaustive Search. For fewer than 12 taxa, one can simply optimize the characters on all possible tree topologies and pick the shortest tree(s). This strategy is preferred given a small number of taxa. It guarantees that the shortest tree(s) will be found.

Strategy 2: Branch-and-Bound. For up to about 20–22 taxa, one can employ a branch-and-bound algorithm that will be guaranteed to find the shortest tree(s). Above 20 taxa, the algorithm, as implemented on most computer platforms, is too slow.

Strategy 3: Heuristic Search. Above around 20–22 taxa, or in situations where the data matrix is "messy" (i.e., contains a high level of homoplasy), the number of possible tree topologies becomes so great that exact solutions are no longer possible. In such cases, heuristic search routines must be implemented.

Heuristic searches are common to all mathematical problems for which an exact solution is unobtainable or impractical. We will meet "heuristic searches" in parsimony analysis, ML analysis and Bayesian analysis under different names and with different algorithmic strategies. In parsimony and likelihood analyses, the investigator is equipped with an optimality criterion and what may be described as a "landscape" of trees with different values for that criterion.

In the case of parsimony, we might imagine a landscape where a plane surface is defined by the average length of all possible trees. This surface is interrupted by valleys and hills. The valleys are filled with trees of longer-than-average length while the hills are filled with trees of shorter-than-average length. Other metaphors describe a sea with islands of shorter trees (Maddison, 1991). Any rational metaphor works if one gets the idea that some hills are higher than others, or some islands

contain shorter trees. Hills are separated from each other by the "inhospitable" landscape of average trees or by valleys of long trees. Or the islands are separated by long stretches of the barren ocean of long trees.

If our problem is simple and there is only one hill, then we can find and climb it. Perhaps a simple step-wise parsimony analysis or classic Hennigian argumentation will be efficient. If our problem is complex and there are many hills, some taller than others, we may ascend a low hill and feel we have found the most parsimonious tree when we have only found a locally optimal solution, not a globally optimal solution.

Local Optimum. A local optimum is achieved when the search finds the shortest tree(s) at the top of a particular hill in the parsimony landscape. Because searches always accept shorter trees and reject longer trees, it is possible to achieve a local optimum that is globally unparsimonious if there is no mechanism for exploring other hills. Locally optimal and globally unparsimonious means that you are on the top of a hill but there are higher hills that you have not found.

Global Optimum. A global optimum is achieved when the search finds the shortest tree(s) on the highest peak(s) on a particular parsimony landscape. Global optimality is never guaranteed in a heuristic search, but may be approached if strategies are adopted that allow exploration of the landscape and in a manner that allows discovery of multiple hills.

Searches, however implemented, are designed to keep the analysis from being trapped in locally optimal solutions by random perturbations of a given tree topology. If the perturbations result in a shorter tree, that tree is retained and it is perturbed; the process continues until the program cannot find any shorter trees. There are a variety of strategies to accomplish such perturbations, ranging from modest to radical, and we describe each of these more fully.

1. Random addition searches.
2. Rearranging tree topologies and analyzing isolated parts of a larger tree.
3. Parsimony ratchet.
4. Simulated annealing.

Random Addition Searches

All modern phylogenetic methods begin with a starting tree, built by some method (or randomly assembled). In parsimony analysis, the order in which taxa are added to the tree can affect the initial tree topology (as was the case with sequence alignment discussed in Chapter 5) and this, in turn, affects all subsequent manipulations. The usual strategy is to employ random addition searches (RASs). An RAS is a strategy of running the analysis many times (10s to 100s) and varying the initial tree by adding taxa randomly during the initial tree-building process. In the metaphor of the plane of parsimony, an iteration of the RAS algorithm allows the tree to land on a different part of the landscape. If the initial tree generated is close to optimal (close to the shortest tree possible), then it may (remember the search is heuristic) land on or near a hill and quickly find the most parsimonious tree on that hill. Increasing the number of RAS iterations increases the possibility of finding more hills. Using an initial algorithm to obtain a starting tree that is close to optimal

ensures a more efficient search. Combining this strategy with rearrangement of tree topologies allows the program to explore many of the trees on a hill or even jump to a new hill.

Rearranging Tree Topologies

The idea behind rearrangement of tree topologies is the exploration of tree topology space. For any particular phylogenetic problem, there is a large number of alternative trees. As the number of taxa increases, the number of possible topologies increases (see Felsenstein, 1978a). The purpose of the exploration of tree topology space is to visit as many tree topologies as possible, compare the new rearrangement to the older result(s) and determine if the new rearrangement results in a shorter tree or group of trees. If it does then the new result is accepted and another round of rearrangements is performed in an attempt to find another group of shorter trees. The process continues until no shorter trees are found or until the investigator terminates it.

One implementation of rearrangement is *branch-swapping* (Fig. 6.12). Most programs allow the investigator to perform one of a variety of branch-swapping routines, in concert with RAS (i.e., RAS + branch swapping). Three common branch-swapping routines are listed below. The terminology is that used in PAUP (Swofford, 2001), but the routines are available in all modern parsimony programs.

1. Nearest-neighbor interchanges (NNI).
2. Subtree pruning and regrafting (SPR), global rearrangements (Felsenstein, 2007).
3. Tree bisection and reconnection (TBR), branch-breaker (Farris, 1988).

Branch swapping, when used in concert with RAS, is an efficient method for finding short trees when the number of taxa is relatively small, say 100 taxa or less. However, if a large number of taxa or "messy" data are analyzed, the computer time used to find the shortest set of trees may be prohibitive. For example, the "Zilla" data set of 500 plants and 1428 DNA base pairs (bp) (Chase et al., 1993) ran for 3.5 months on three Sun workstations without finding shortest trees using RAS + TBR (Rice et al., 1997). Soltis et al. (1998) found similar problems using a 2800 bp data set—a set they had hoped would cut computation time because fewer equally parsimonious trees are likely to be present if the data set is larger.

One reason for the long computation time was the fact that then current implementations of heuristic search routines concentrated on finding all of the most parsimonious trees on each island of trees. Each RAS attempts to find all of the most parsimonious trees. If there are many islands of short trees, each RAS might go through a great number of trees (hundreds of thousands or millions). As Farris et al. (1996) and Goloboff (1999b) point out, when data sets are large and complex, finding a significant number of shortest trees from *different* islands of optimality will result in a consensus that is likely to be identical to that produced by finding *all* of the most parsimonious trees and then computing a consensus. Farris et al. (1996) used the jackknife to discover strongly supported groups—groups that would appear well supported in any analysis, while eliminating poorly supported groups that appear only on a minority of shortest trees. In essence, they reasoned that
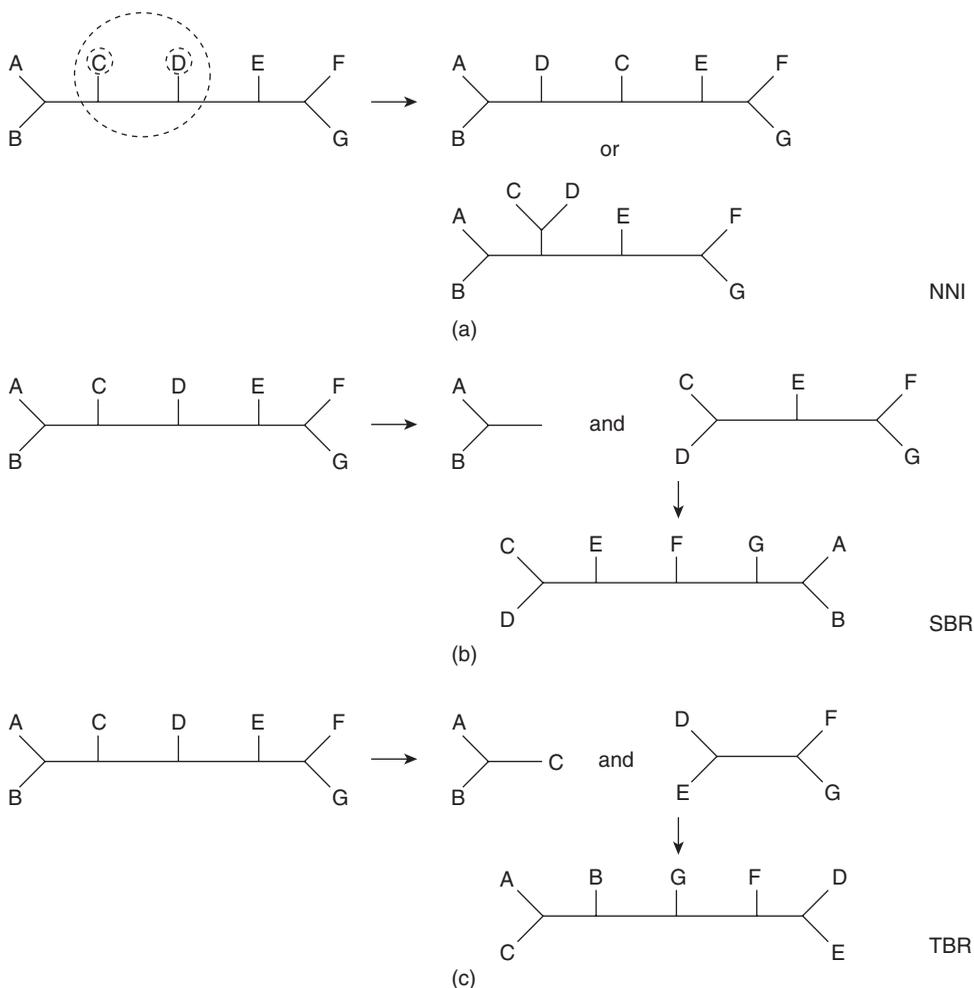
Figure 6.12. Branch swapping. (a) Nearest-neighbor interchanges, NNI. (b) Subtree bisection and regrafting, SBR. (c) Tree bisection and reconnection, TBR.

laboriously computing trees that would contain groups that disappeared when a consensus was computed was a waste of computer time.

Goloboff (1999b) demonstrated that simply increasing the number of RASs, using TBR, and keeping only a few trees for each search could dramatically decrease computation time. Using NONA, he was able to find shortest trees for the "Zilla" data set in 24–48 hours compared to 2.5 months of exhaustively finding the shortest tree on each island of trees. However, as Goloboff (1999b:417) pointed out, large data sets have "composite optima" that interfere with the quest for globally optimal solutions. Large trees of more than 50 taxa tend to have sectors, defined as local groups of taxa. A tree of 500 species, such as the "Zilla" tree, might comprise 10 sectors of 50 taxa each. The problem is: each sector may have its own local optima and whether it is placed on the tree may be partly independent of the placement of the other sectors. (If they are truly independent, then the problem is simplified; if

very dependent, then the problem is much harder because changing one affects the other(s).) This is exactly what RAS + TBR does; each iteration might break up a local optimum, and unless a great number of RASs are performed, no globally optimal solution will be found. The problem, then, is to find a tree where the sectors are in proper configuration with each other. Goloboff (1999b:417) states:

> Thus, the solution requires sectors be improved separately, one at a time—that those sectors which are suboptimal are improved without worsening the ones that are already optimal. For this, there are four basic methods: ratchet, tree fusing, tree-drifting, and sectorial searches. These methods do not attempt to find multiple trees during swapping, but simply concentrate on finding trees as short as possible.

Goloboff (1999b) suggested a number of swapping techniques built around the central idea stated above that the consensus tree produced by brief visits to many islands of most parsimonious trees would be identical to the consensus produced by laboriously calculating all of the most parsimonious trees. One strategy was simple: even with RAS + TBR, it is possible to cut computation time by saving only a few trees with each RAS and performing many RASs.

Speeding Up Rearrangements. SPR and TBR rearrangements can be speeded by recalculating only the part of a tree (a sector/window) that has been changed (Goloboff, 1999b). This uses a method outlined by Ronquist (1998b) to cut down on computation time by looking at the sectors nearest the connection point of the recalculated sector.

Tree Fusing. Tree fusing consists of exchanging subgroups of the same taxa between trees. The subgroups that are exchanged are present in the consensus of both trees and not dichotomously resolved in a consensus of the two trees. (Exchanging subtrees of dichotomously resolved taxa between consensus trees is unproductive because both trees have the same dichotomous relationships for the taxa exchanged.) If an exchange results in a shorter tree, then this tree is saved. This strategy is built around the idea that the formed subgroups might be optimal, but that relationships within them and to other subgroups might not be optimal for a particular tree.

Sectorial (Tree Window) Searches. Sankoff et al. (1994) and Goloboff (1999b) suggest that isolating certain clades and then performing an analysis might improve the resolution of the isolated subclade. Because fewer taxa are involved, the analyses are faster, and thus, the computational burden of attempting to escape local optima is less. Sankoff et al. (1994) isolated subclades of 20 or fewer nodes and performed branch and bound analyses on the isolated subclade. Goloboff (1999b) prefers the quicker method of TBR and thus analyzes more nodes (35–55). If the result improves the length of the tree, then the analysis moves to another subclade (another window or sector) and performs an analysis on the new subtree. Felsenstein (2004) suggests that if the purpose is to escape local optima, then the less exact method of Goloboff (1999b) might be preferable.

The Parsimony Ratchet

Nixon's (1999) parsimony ratchet is a technique that escapes local optima by emphasizing a limited number of characters within the data matrix to see if these characters

lead to shorter trees. From the original data matrix, some percentage (5–15 percent) of the characters are selected and weighted more heavily than the other characters. An analysis is then performed, and this will favor the weighted characters. The resulting tree topology is then evaluated using the entire data matrix with all characters equally weighted to determine the length of the tree. If a shorter tree results, the tree is saved. Many reweightings, searches, and evaluations are carried out, and the shortest trees are retained. Ratcheting is related to techniques that explore tree space by analyzing only part of the data (e.g., the Jackknife as used by Farris et al., 1996). Although Nixon (1999) implemented the ratchet specifically for parsimony analysis, Felsenstein (2004) calls attention to the fact that ratcheting can be used on any number of other approaches.

Simulated Annealing

Parsimony analysis of large data sets is one of many kinds of complex combinatorial problems for which exact solutions are not possible in practice. The solutions for such problems can be estimated using simulations of statistical mechanics using the Metropolis algorithm (Metropolis, 1953; Kirkpatrick et al., 1983). We shall hear much more about this approach in the chapter on statistical phylogenetics where it is used in Bayesian and (rarely) likelihood analyses. In short, the algorithm usually accepts a shorter tree, but it might accept a longer tree under certain, specified conditions. As the simulation proceeds, it wanders through the tree landscape usually favoring shorter and shorter trees until it settles on a peak (valley/island) from which it cannot escape. As implemented by Goloboff (1999b) under the name *tree-drifting*, suboptimal trees may be accepted during branch-swapping if they meet a criterion based on the relative fit difference between the trees.

OPTIMIZING CHARACTERS ON TREES

Character optimization is an initial step in understanding the evolution of characters. It can be applied to any tree, not just the shortest tree(s). You will find this useful if, for example, you wish to see the interpretation of the evolution of a particular character state on your preferred tree as compared to rival, less parsimonious, trees.

Both Farris (1970) and Fitch (1971) suggested strategies for optimizing characters on trees, each based on their own algorithms and neither providing formal proofs. Swofford and Maddison (1987) provided a proof for ordered optimization routines, and we use this to give an example of how to calculate the length of a tree in an earlier section. They also provided a proof for finding other equally parsimonious interpretations of character evolution based on most parsimonious resolutions (MPR) sets. The results are alternative ways of interpreting character evolution when one has more than a single most parsimonious character reconstruction. If we accelerate character transformation, then the effect is to push the time of transformation down the tree. This is commonly called ACCTRAN (accelerated transformation). If we delay character transformation, then the effect is to push transformation up the tree. This is commonly called DELTRAN (delayed transformation). These alternatives are easy to visualize with some examples (see also Wiley et al., 1991). We will begin with ACCTRAN, which is the original Farris (1970) optimization.

ACCTRAN Optimization

In Fig. 6.13a, we have a small matrix of two characters and four taxa. We will optimize the first character on the tree in Fig. 6.13b.

For a group of taxa, select the outgroup and root the tree with the outgroup and label all the terminal nodes/labeled taxa (Fig. 6.13b). Unlike computing tree length, you cannot pick any taxon to root the tree; it must be the outgroup because character polarity will vary with outgroup selection. Proceeding from the tips to the root, apply the following rules to the internal nodes.

Rule 1a. If the intersection of the state set is empty, then let the character set of the ancestor be the smallest interval from each set. (For binary transformations, simply label the ancestor with both characters, [a, b], [0, 1], etc.; Fig. 6.13c.)
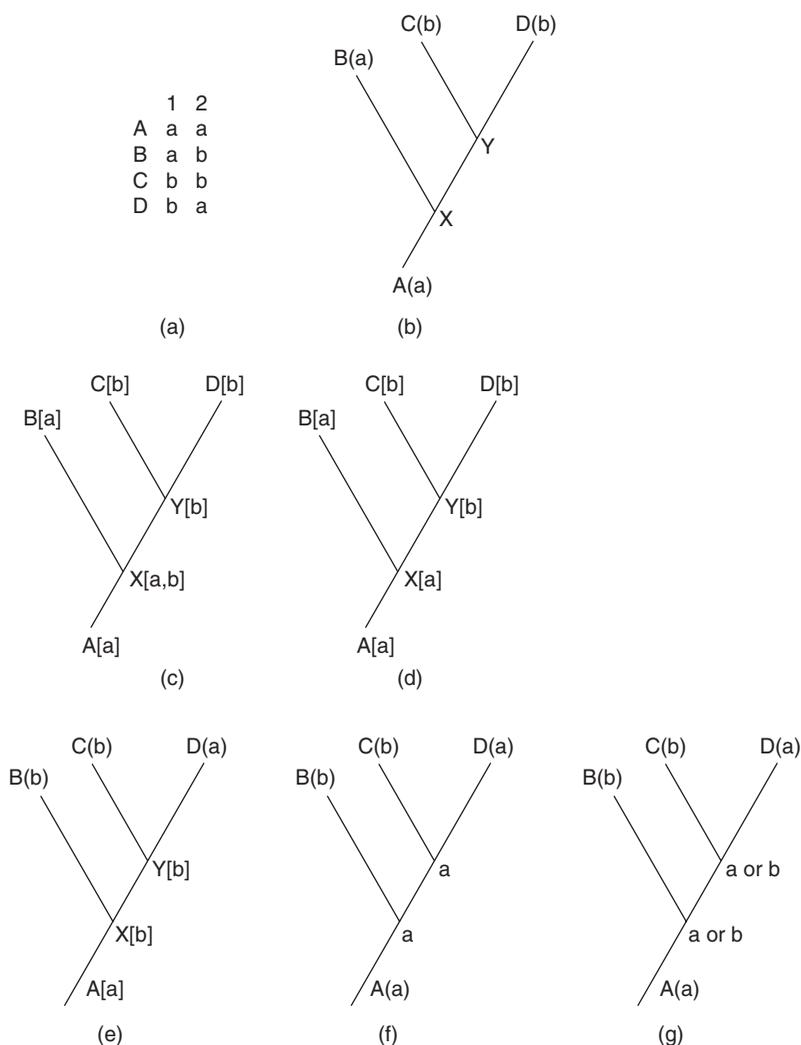
Figure 6.13. ACCTRAN and DELTRAN optimization. (a) A data matrix. (b–d) ACCTRAN optimization of character 1. (e–f) Two different but equally parsimonious optimizations of character 2 on the tree. (g) The MPR sets for all ancestors on the tree.

Rule 1b. If the character set is not empty, let the character set of the ancestor be the intersection as a closed interval. (For binary characters, this is simple: [0] or [1], [a] or [b], etc.; X in Fig. 6.13c.)

Once you reach the root, you then traverse up the tree toward the tips and apply these rules.

Rule 2a. If the descendant node has a character set with a single element, then it remains unchanged.

Rule 2b. If a descendant node has a closed interval, assign the interval with the smallest distance from the ancestor to the descendant.

For our very simple tree (Fig. 6.13b), we can see that the down-pass results in Y being assigned [b] and X being assigned [a, b] (Fig. 6.13c). As we move from root (A) to the tip, because A[a] and X[a, b], we change X to [a] (Fig. 6.13d). Because Y[b], we do not change it, even though its ancestor has the state set [a].

DELTRAN Optimization

For some character distributions, there are other possibilities besides ACCTRAN. Let us look at the second character column in Fig. 6.13a. ACCTRAN interprets character evolution as the accelerated transformation of A[a] to X[b]. Then it interprets another transformation from Y[b] to D[a] (Fig. 6.13e). Note that tree length is two steps.

However, there is an equally parsimonious tree with a different optimization, shown in Fig. 6.13f. In this interpretation, transformation from [a] to [b] is delayed; state [b] evolves independently in taxa B and C. This tree is also two steps in length.

Obviously, X and Y have two possible elements in their character sets, [a, b] for this transformation series (Fig. 6.13g), but ACCTRAN finds only a single element [b].

Swofford and Maddison (1987) presented a formal proof for finding all of the possible elements of the node's character set, not just some as found in ACCTRAN. This character set is termed the MPR set. We will illustrate the process of finding the MPR set using the binary characters in Fig. 6.13a and show that the MPR set is exactly those states shown in Fig. 6.13g. DELTRAN is then simply implemented with the same type of upward traversal as ACCTRAN, but results in character states optimized as in Fig. 6.13f rather than 6.13e.

1. Beginning with the unrooted tree (Fig. 6.14a), we pick an internal node and root the tree with this node (Fig. 6.14b; although we used taxon X, we could have actually used any internal node).
2. We perform a downward pass, assigning characters to the internal nodes just as we did in ACCTRAN (Fig. 6.14c).
3. We then reroot the tree with the next internal node, and we perform a downward pass, assigning character states to each internal node that has not been previously optimized (Fig. 6.14d). Note that we would not change the state set of X as it has already been determined.
4. Once we have rerooted the tree with all internal nodes, we have completed the assignment of the MPR sets for each of the nodes. We then root the tree
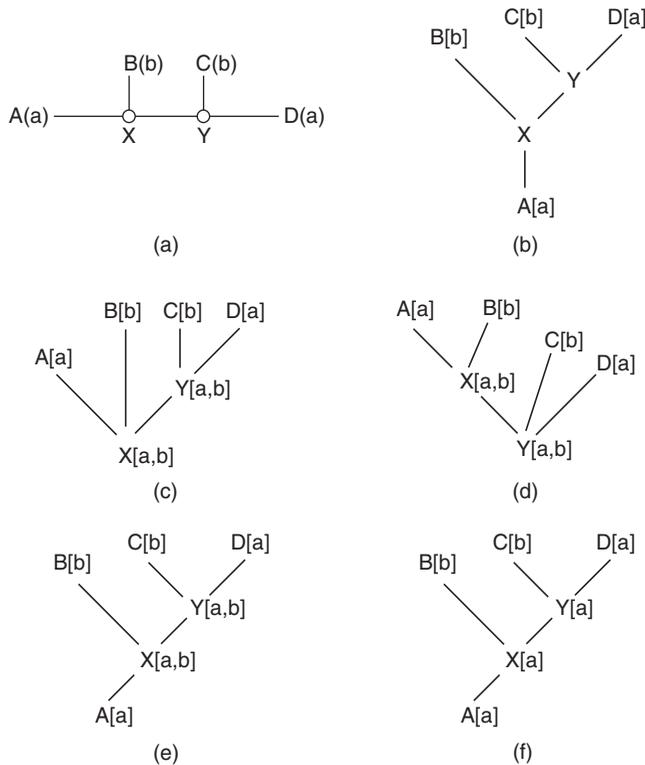
Figure 6.14. DELTRAN optimization. (a) Distribution of character states for character 2 of Fig. 6.13. (b–f) Sequential steps of optimization.

with one of the terminal taxa (presumably the outgroup) and perform an upward traversal, assigning characters to the internal nodes using the same rules we used in ACCTRAN, resulting in the optimization shown in Fig. 6.14f, which is identical to Fig. 6.13f.

The formal algorithms of Swofford and Maddison (1987) were built on informal optimization models by Farris (1970) and Fitch (1971) and applied only to dichotomous trees. W. Maddison (1989) extended MPR algorithms to polytomous trees.

SUMMARY TREE MEASURES

Once one obtains a tree or set of trees, there are various character performance measures that can be used to summarize the analysis and compare the tree(s) obtained with other possible solutions. Current computer packages provide this information, either automatically or upon request. They are not useful in evaluating the results of two data sets of the same taxa; but rather, to compare the results between trees for the same set of characters.

   Tree Length. We have already described how tree length is calculated. Because the optimality criterion of a parsimony analysis is the minimum path of evolution that explains the data, tree length is a fundamental measure. Tree length is simply

measured by summing the number of changes that occur on the tree, as detailed above. There are two kinds of most parsimonious trees. First, there is the set of trees that have the same length but differ in topology, that is, parts of them contain different hypotheses of common ancestry. Second, there is the set of trees that have the same topology but differ in their interpretation of character evolution. These two types of trees have different qualities.

If an analysis results in a large number of equally parsimonious tree topologies, this reflects conflict among characters. The areas of conflict may be explored by performing a strict consensus analysis (discussed later in the chapter), which will result in polytomies where the conflict occurs. Because the number of possible trees increases quickly, it is possible to obtain many most parsimonious trees in a large matrix where conflict is confined to relatively small local regions (as, for example, among terminal species that belong to only one of many groups), but sometimes the entire consensus tree in such a situation can be unresolved. If an analysis results in a set of trees that are identical in topology but contain different interpretations of character evolution, the difference might be interesting from an evolutionary perspective. For example, it might provide possible tests of evolutionary mechanisms of character change that could be explored.

Consistency Indices. Kluge and Farris (1969) introduced measures of the performance of both individual characters and entire matrices relative to particular tree topologies. Consider a single character. If, on a particular tree, the states of this character could be mapped in such a way that there were no instances of homoplasy, then these states have "perfect" performance relative to the topology. However, if the topology was such that the only way to map the states was to invoke some level of homoplasy, then performance is less than perfect. Writ large, if an entire data matrix was composed of characters with states that required no homoplasy on a particular tree topology, then the performance of the entire data matrix would be "perfect" relative to that particular topology. The more conflict required to map the states, the greater the deviation from perfect performance. Various measures of character consistency can be generated to explore the performance, both of individual characters and entire data matrices. We can use the example provided by Wiley et al. (1991) to see how such measures are generated. We begin with data shown in Table 6.2.

Consistency index of a single transformation series (ci, or c). The ci of a single character is the ratio of the minimum number of steps or changes it might undergo

TABLE 6.2. Data matrix for the hypothetical clade A–E and its sister group OG. From Wiley et al. (1991).

Taxon	Transformation series							
	1	2	3	4	5	6	7	8
OG	0	0	0	0	0	0	0	0
A	1	0	0	0	0	1	0	1
B	1	1	1	0	1	0	1	0
C	1	0	1	1	1	0	0	0
D	1	1	1	0	1	1	1	0
E	1	1	1	1	1	1	1	1

Figure 6.15. Two trees showing the distribution of synapomorphies of taxa A–E based on the matrix in Table 6.2. (a) A stem-based tree with synapomorphies mapped along inferred ancestral lineages. (b) A node-baesd tree with synapomorphies mapped at ancestral species nodes. Redrawn from Wiley et al. (1991), used with permission, Biodiversity Institute, University of Kansas.

and the number of changes or steps it actually undergoes on a particular tree topology:

$$\text{ci} = m/s$$

where m is the minimum number of steps, and s is the actual number of steps.

For binary characters, $m = 1$. For more than two states, m = the total number of steps necessary to minimally account for the evolution of the homologies (i.e., for three character states, $m = 2$; for four character states, $m = 3$, etc.).

Examine the matrix and consider character one in Fig. 6.15. Because 1 is binary, the number of minimum steps is $m = 1$. Note that there has been a single transformation from 0 to 1 at the internode leading to the clade ABCDE (labeled 1-1). Thus, $s = 1$ and the ci for this character is ci = $m/s = 1/1 = 1$.

Now consider character 8. Again, $m = 1$, but state 8-1 has evolved twice on the tree, so $s = 2$ and the ci for this transformation series is 0.5. Now consider character 6. If you do the calculations, you will find that the ci of character 6 is the same as character 8 (0.5). Note, however, that the quality of the two characters is different. In Fig. 6.15, the hypothesis that states 8-1 is a synapomorphy uniting taxa A and E is rejected in favor of the interpretation that each is an autapomorphy. As such, they contribute nothing to the resulting topology of the tree. In Fig. 6.15, the hypothesis that 6-1 is a synapomorphy is confirmed for the clade DE, but is disconfirmed for a clade containing D, E, and A. Thus, the character shared by D and E contributes to the topology of the tree, yet the ci in both 6 and 8 is identical.

The Rescaled Consistency Index (rc). To overcome this problem, Farris (1989b) introduced the rescaled consistency index (which appeared in Hennig86; Farris, 1989a). It is the product of the original consistency index and the **retention index (ri)**. We will use the characters in Fig. 6.15 to illustrate calculating the retention index and the rescaled consistency index.

The **retention index (ri)** measures the fraction of apparent synapomorphy to actual synapomorphy. To calculate the retention index, we need a new parameter, the **g-value (g)**; it is a measure of the "best of the worst" possible performance of each character relative to the actual performance of that character. The "best of the

TABLE 6.3. Some values* used to calculate rescaled consistency indices form Wiley et al. (1991).

TS	m	s	g	ci	ri	rc
1	1	1	1	1.00	0/0	0/0
2	1	1	3	1.00	1.00	1.00
3	1	1	2	1.00	1.00	1.00
4	1	2	2	0.50	0.00	0.00
5	1	1	2	1.00	1.00	1.00
6	1	2	3	0.50	0.50	0.25
7	1	1	3	1.00	1.00	1.00
8	1	2	2	0.50	0.00	0.00
Totals	8	11	18			

* m = no. changes a character might show on a tree; s = no. changes a character does show on a tree; g = minimum no. steps for each TS given a polytomy; ci = character consistency index; ri = character retention index; rc = character rescaled consistency index.
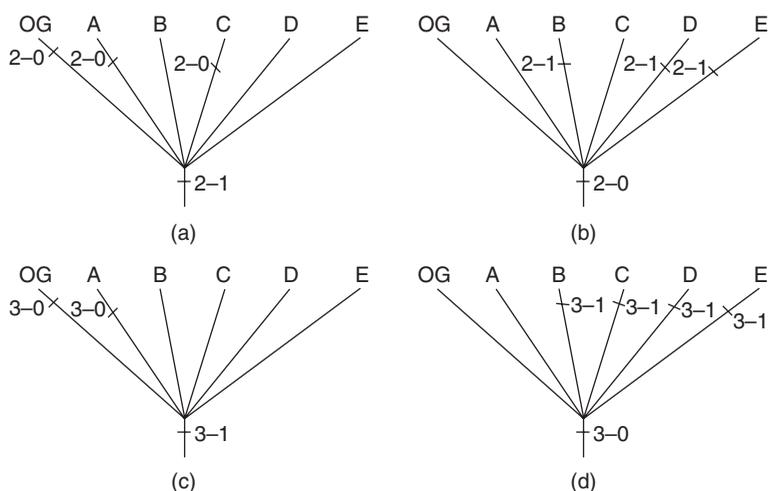
Figure 6.16. Performance of character states under the parsimony criterion. (a, b) Performance of character 2. (c, d) Performance of character 3. Table 6.3 shows the minimum number of times each is allowed to evolve independently on the polytomy. The fewest number of times is used to derive the metric "g" in Table 6.3. Redrawn from Wiley et al. (1991), used with permission, Biodiversity Institute, University of Kansas.

worst" performance would be the performance (in steps) of each transformation on a polytomy of all taxa by considering two scenarios in the binary case, one that assigns one state to the root and evaluates the performance of the other state at the tips and vice versa. The smaller value is the "best of the worst." Here is how it works.

Refer to Fig. 6.15 for the characters and Fig. 6.16 and Table 6.3 for how this is calculated. Consider character two. The worst possible performance of any character would be its performance on an unresolved tree. The "best of the worst" would be a contrast, in the binary case, between two states on a polytomous topology, with the state showing the fewest changes being better than the one showing more changes. (We are using binary transformation series for simplicity.) Now consider two cases, characters two and three (Fig. 6.16a–d). If we set the state 2-1 to the root

of a polytomous tree of all taxa, we can see that the tree requires 2-0 to evolve three times (Fig. 6.16a). Conversely, if we set 2-0 to the root, the tree requires 2-1 to evolve three times (Fig. 6.16b). It is a tie. Neither performs better in the polytomous case, so the *g*-value is $g = 3$.

Now consider character 3. If we set state 3-1 at the root, the tree requires 3-0 to evolve twice (Fig. 6.16c). If we set 3-0 to the root, we would require 3-1 to evolve four times (Fig. 6.16d). Thus, the "best of the worst" is the g-value of $g = 2$.

We can now define the rescaled consistency index, using the g-value and the s- and m-values that formed part of the original ci:

$$ri = (g - s)/(g - m)$$

where *g* is the best performance on the unresolved tree, *s* is the actual number of steps of a transformation series on the resolved tree, and *m* is the minimum number steps of a transformation series on the resolved tree. The rescaled consistency index, **rc**, is simply the **ri * ci** (see Table 6.3).

For character 2, the retention index ri = $(3 - 1)/(3 - 1) = 1.0$. For 3, the ri = $(2 - 1)/(2 - 1) = 1.0$ (Table 6.3). This makes sense; both transformation series have perfect ci-values and show no homoplasy. Thus, they are contributing the maximum possible to the tree topology.

Now, consider characters 6 and 8 (Table 6.3). They have identical consistency indices (0.5), but 8 contributes nothing to the structure of the tree while 6 acts as a synapomorphy in one place and an autapomorphy in another. If we calculate the g-value for character 8, we see that $g = 2$. Character 6 has a *g*-value of $g = 3$. Calculating ri, we find the following (Fig. 6.16e):

$$ri\,(8) = (2 - 2)/(2 - 1) = 0.0$$
$$ri\,(6) = (3 - 2)/(3 - 1) = 0.5$$

The result is straightforward. The most parsimonious tree interprets the state 8-1 as two instances of homoplasy resulting in two autapomorphies. Its contribution to the structure of the tree is zero. In contrast, the most parsimonious tree interprets the state 6-1 as one instance of homoplasy, one instance of synapomorphy, and one instance of autapomorphy. The instance of synapomorphy contributes to the structure of the tree. Incidentally, all instances of unique characters (autapomorphies or characters that map on at the ingroup node) will also have ri = 0.0, removing unique characters from contributing to the ri-value. In contrast, the ci-values of unique characters are ci = 1.0. This becomes important when we consider the next set of measures: ensemble values.

Ensemble consistency indices can be used to examine the relationship between an entire matrix and a given tree topology. One commonly reported index is the ensemble consistency index (CI) (Kluge and Farris, 1969). For a binary matrix, this index is simply calculated by taking the ratio of the number of data columns and the length of the tree. If there is no homoplasy, then the CI-value will be CI = 1.0. Deviations from this value indicate that homoplasy is present.

The CI-value suffers from some problems. The CI is artificially inflated by unique characters that contribute nothing to the structure of the tree. One solution is to calculate CI after eliminating all transformation series that contain unique characters (Carpenter, 1988). This result can be calculated in most parsimony programs.

Another problem is the fact that there is a negative relationship between CI-values and the size of the data set (Archie, 1989). Large data sets have small CI-values simply as a function of the size of the data set, not the contribution of the data to the tree structure. To address these problems, Farris (1989b) suggested that ensemble values be calculated using rescaled values, not raw values. The various formulae needed are shown below.

$$CI = M / S$$

where CI is the ensemble consistency index, M is the sum of m-values for each character, and S is the sum of s-values for individual characters ("totals" row, Table 6.3).

$$RI = (G - S)/(G - M)$$

where RI is the ensemble retention index and G is the sum of individual g-values for all transformation series ("totals" row, Table 6.3). Then:

$$RC = (CI)(RI)$$

where RC is the ensemble rescaled consistency index. Values calculated for the example (Table 6.3) are shown below.

$$CI = M / S = 8 / 11 = 0.727$$
$$RI = (G - S)/(G - M) = (18 - 11)/(18 - 8) = 0.7$$
$$RC = (0.727)(0.7) = 0.509$$

EXAMPLE 2: OLENELLOID TRILOBITES

The basal lineages of trilobites, apart from the controversial Agnostida, comprise what Fortey (1997) recognized as the order Redlichiida (early to middle Cambrian). In a series of papers, Lieberman (1998, 1999, 2001, and 2002a) analyzed the relationships of this group and came to the conclusion that it is actually a grade group leading to Eutrilobita (the crown trilobites). The 2002a paper is the one we shall consider here.

Fortey (1997) recognized the paraphyletic nature of Redlichiida, and he divided the grade into two suborders: Olenellina and Redchiida. Among olenellines, Fortey recognized two superfamilies: Olenelloidea and Fallotaspidoidea. Lieberman (1998) analyzed these basal lineages and concluded that the olenelloids were monophyletic, but that Fallotaspidoidea was not; some fallotaspidoids were more closely related to olenelloids while others were more closely related to redlichiids. Further, the redlichiids are related to the rest of Eutrilobita (Fig. 6.17). Thus, olenelloids and a set of taxa traditionally assigned to the Fallotaspidoidea (specifically, judomioids and nevadioids) are the sistergroup of all other trilobites, including the other fallotaspidoids; he referred these to a monophyletic Olenellina. Lieberman (1999, 2001) subsequently analyzed relationships among the Olenellina, returning to the problem of the remaining fallotaspidoids and their relationships to other trilobites in Lieberman (2002a).
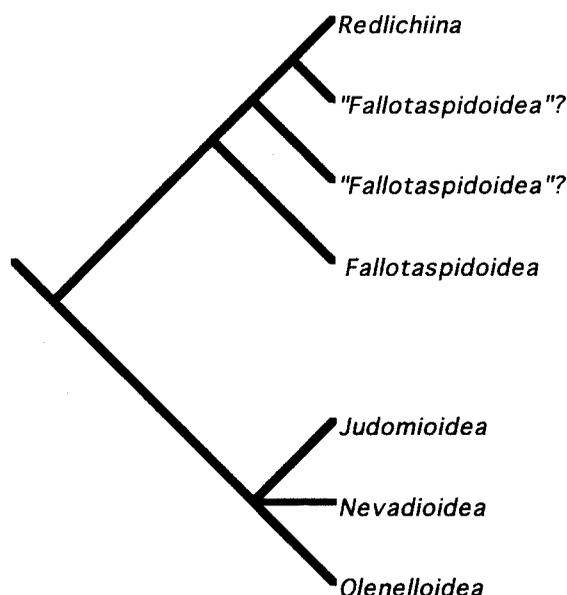
Figure 6.17. Relationships of basal trilobites as resolved in Lieberman (1998). From the *Journal of Paleontology*, used with permission of the Paleontological Society.

Lieberman (2002a) analyzed 16 species for which well-preserved specimens were available (see Table 6.4). This included all known genera of nonolenellinid fallotaspidoids and three species representing the basal members of the rest of Eutrilobita (the apical trilobites). These basal members (one species of *Bigotina* and two species of *Lemdadella*) were in effect used to root the fallotaspidoids to taxa up the tree. The analysis also added an outgroup; the outgroup used was a synthesis created by reconstructing the character vector for the inferred common ancestor of Olenellina *sensu* Lieberman (1998) using MacClade (Maddison and Maddison, 1992–2008). Twenty-nine transformation series were analyzed, all taken from the exoskeleton of holaspids. (In redlichiids and olenelline trilobites, and indeed in most trilobites, holaspids are specimens were subsequent molts do not result in an increase in the number of thoracic segments, and thus presumably represent adults.)

The resulting data matrix (Table 6.4) was analyzed using PAUP* 4.0 (Swofford, 2001). Lieberman ran 100 RASs and TBR branch-swapping. Transformations were optimized using ACCTRAN. A single most parsimonious tree (Fig. 6.18) of 106 steps was recovered. Summary tree measures were CI = 0.53, RI = 0.58, and RC = 0.30.

Character support for this tree ranges from a number of unique and unreversed transformations (e.g., data columns 3, 4, 7, 24) with $ci = 1.0$ to characters that are highly homoplastic (e.g., column 19, $ci = 0.20$). This analysis points out how ci values can be misleading. In particular, sometimes a character may have $ci = 1.0$ but there is missing data. For example, having a pygidium with its axis distinctly separated from the pleural field (column 27, character state 1) appears to be unique and unreversed, but there are several taxa that could not be scored for this character. Perhaps in reality these taxa possess state 0 such that the ci really should no longer $= 1.0$. Here we consider three character matches in greater detail.

TABLE 6.4. Character state distributions for taxa used in phylogenetic analysis of Lieberman (2002a) Selected characters and character states are discussed in text Missing data are indicated by "?". Character numbers are listed at top of table. Character states listed as "V," "W," "X," "Y," and "Z" are polymorphic, where "X" = (0&1), "Y" = (1&2), "Z" = (0&2), "W" = (0&1&2), and "V" = (1&3).

	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2
	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9
Olenellina Node	0	0	0	0	0	0	0	0	0	0	0	0	W	0	0	Z	0	0	0	0	0	0	0	0	Z	0	0	0	0
Repinella sibirica	0	0	0	0	1	V	0	1	0	0	?	0	0	1	?	1	1	0	?	0	0	1	0	0	0	1	?	?	?
Profallotaspis jakutensis	0	0	0	0	1	3	1	1	0	0	1	1	1	0	1	1	1	0	1	1	2	2	0	0	?	?	?	?	?
Pelmanaspis jurii	0	0	0	0	1	3	1	1	0	0	1	1	1	0	1	1	1	0	1	1	2	1	0	0	?	?	?	?	?
Eofallotaspis tioutensis	0	0	1	0	1	3	1	1	0	0	1	1	0	2	?	0	1	1	0	1	1	1	0	0	?	?	?	?	?
Daguinaspis ambroggii	0	0	1	0	1	3	1	1	1	0	0	1	X	2	2	0	0	0	0	1	2	2	1	0	0	0	0	0	0
Choubertella spinosa	0	1	1	0	1	2	1	1	1	0	1	1	0	2	2	0	0	0	1	1	2	2	1	0	0	1	0	1	1
Fallotaspis typica	2	1	0	0	0	0	1	1	0	0	0	1	1	0	2	0	1	1	1	1	1	0	0	0	Y	1	?	?	?
F. bondoni	1	1	0	0	X	X	1	1	0	1	1	0	0	0	1	Z	1	1	1	1	1	0	0	0	2	1	0	0	0
Parafallotaspis grata	1	0	0	1	1	2	1	1	0	0	2	1	0	2	0	2	1	1	0	1	1	0	0	0	?	?	0	0	1
Archaeaspis hupei	2	0	0	1	1	1	1	0	1	0	0	1	0	0	1	0	0	1	1	1	1	1	0	0	?	?	?	?	?
A. nelsoni	1	1	0	1	1	3	1	0	0	0	1	0	0	0	2	0	?	?	1	1	1	0	0	0	0	1	?	?	?
A. macropleuron	1	1	0	1	1	1	1	0	0	1	0	0	0	0	2	0	?	?	0	1	1	0	0	0	1	0	?	?	?
Fallotaspidella musatovi	1	0	0	1	1	1	1	1	1	0	0	1	1	0	1	1	0	1	0	0	0	0	0	0	?	?	?	?	?
Bigotina bivallata	1	1	0	1	1	V	1	X	1	0	1	1	Y	2	0	2	1	0	0	0	0	1	0	1	?	?	1	1	1
Lemdadella antarcticae	1	0	0	1	1	1	1	X	0	0	1	1	2	0	1	0	0	0	0	0	0	X	0	1	?	?	1	1	1
L. linaresiae	1	1	0	1	1	1	1	1	0	0	X	1	1	2	1	2	0	0	0	0	0	1	0	1	0	1	1	1	1

Figure 6.18. A phylogenetic tree from Lieberman (2002a). Only characters from Table 6.4 that are discussed in the text are mapped on the tree. Modified from the *Journal of Paleontology*, used with permission of the Paleontological Society.

Character 1. The relative length of the anterior border of the head shield is either shorter than the length of the occipital ring (L0) as shown in Fig. 6.19a, about the same length as the occipital lobe, or much longer than the occipital ring (Fig. 6.19b). Although this character has a $ci = 0.667$, the first state (short to equal) is actually unique and unreversed and diagnoses a major clade of trilobites. The homoplasy is found in the second state, where the longer length of the anterior border is homoplastic in two species formerly assigned to Fallotaspidoidea and contributes nothing to the structure of the tree.

Character 5. The cranidium (central head region) of trilobites is composed of the glabella (segmented middle part) and a complex broader shield-like structure, the fixigenae. A furrow runs across the anterior part of the shield medially, and the glabella either contacts this furrow (coded 0, Fig. 6.19a) or does not (coded 1, Fig. 6.19b). This character also has a $ci = 0.667$. However, this is due to a reversal in only one taxon (*Fallotaspis typica*). Another taxon has missing data. Thus the similar ci-values conceal very different kinds of character evolution.

Character 19. Returning to the occipital ring, there are two kinds of ornamentation found on the medial surface of the ring, a faint node (0, Fig. 6.19a) or a spine (1, not figured). This character presents a typical worst case scenario for characters, with a ci only $= 0.20$. Although it diagnoses at least two smaller clades, we can safely conclude that homoplasy makes it of little use in distinguishing between alternate
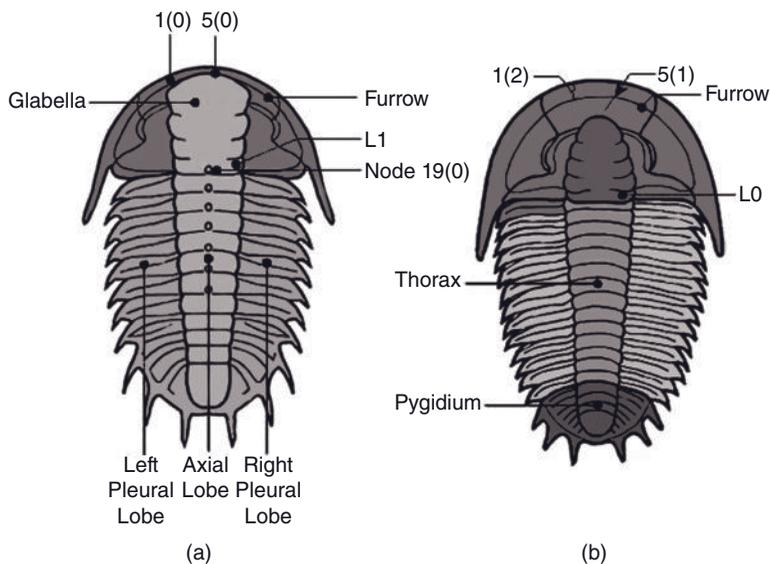
Figure 6.19. Two diagrammatic trilobites illustrating some of the characters used in the Lieberman (2002a) analysis. Key to characters: Anterior border of head shield narrow [1(0)] versus broad [1(2)]; glabella contacts furrow [5(0)] or not [5(1)]; a faint node on the occipital ring [19(0)] versus a spine [19(1)]. Used with permission of the Paleontological Institute, University of Kansas. See color insert.

tree hypotheses. When considering characters 1, 5, and 19 (in the context of all the other character data), we can see how that each character state has sorted itself out via parsimony. We are able to discern which character states ultimately provide the most phylogenetic utility and, by contrast, which are relatively uninformative.

EVALUATING SUPPORT

While the various consistency measures might yield information about the characteristics of particular transformation series on a tree and the amount of relative homoplasy, they do little to illuminate the actual support for individual monophyletic groups. There are, however, several ways of addressing this second, also important, question. At the most basic level, this question involves considering how robust is the tree(s) obtained; are all or most of the clades well supported, and which clades are suspect?

Count and evaluate the synapomorphies. The most direct method of evaluating support for a monophyletic group is to count and evaluate the synapomorphies, a largely qualitative assessment. Monophyletic groups supported by synapomorphies with high individual ci-values may be judged to be robust and the more of these kinds of characters the better. Monophyletic groups supported by synapomorphies with relatively low ci-values may be judged relatively weak and are likely to be rejected if new evidence emerges. The quality and complexity of the synapomorphies might also be judged. Highly complex synapomorphies, ones judged unlikely to have evolved twice, get high marks. Frequently specialists in a group will avoid

certain kinds of characters because they have the a priori notion that these characters are subject to homoplasy, but in principle, this should not be necessary because such characters should sort themselves out via the test of congruence when parsimony is applied.

This simple approach, as one might expect, has some problems. One problem is that the connection with some more general principles seems lacking. History is a contingent phenomenon, and evolutionary theory does not predict particular classes of characters that might fit our a priori judgments in these matters. For example, should losses be counted less than gains? Perhaps they should in the case of restriction sites, maybe not in the case of teeth. It is hard to generalize. Nevertheless, we are impressed with trees full of tick-marks that have high ci-values and suspect, with reason, that the clade is robust.

Another problem, shared by all methods of evaluation, is the problem of the independence of characters. Consider a hypothetical case in which one clade is supported by 10 unique and unreversed synapomorphies and the alternative is supported by two unique and unreversed synapomorphies. What if the 10 unique and unreversed synapomorphies corroborating the first clade are not independent of each other while the 2 synapomorphies supporting an alternative clade are independent? Then the 10 synapomorphies really represent a single synapomorphy and the alternative group would represent more support. But how do we evaluate phylogenetic independence? If synapomorphies lie on different parts of the tree and support different or nested monophyletic groups, then we can deduce independence. But if they appear as support for the same clade, then no such deduction follows and research would have to be undertaken to prove that they are truly independent (e.g., developmentally or genetically independent).

This simple approach also has an advantage, a concern for the evidence. Hennig (1966) stressed that critical phylogenetic inquiry is not simply a process of building a tree; it is a process of reciprocal illumination where the investigator is constantly questioning both the data and the results. More attention paid to the kinds and quality of evidence is never a bad goal and lays the foundation for more complicated measures of support. However, by the same token, obtaining a particular phylogenetic result should not necessarily then motivate subsequent targeted character search strategies by a worker. In particular, *a posteriori*, it would be invalid to spend a disproportionate amount of time trying to obtain additional character evidence to bolster support measures for a particular group to the expense of actually considering evidence that the group might not be monophyletic.

Bremer Support. Bremer (1988) suggested that a useful measure of support for a particular clade might be the difference in the length of a tree where it appeared as a monophyletic group and the length of the tree where it did not. We can easily run an analysis constraining the results to include the monophyletic group in question, and then run an analysis constraining the results to not include the monophyletic group in question using options commonly available in program packages. This can also be implemented using the package TreeRot (Sorensen and Franzosa, 2007). The difference in length between the two trees is a measure of how many steps longer a tree would be in order to overturn the monophyly of the group. For example, we run an analysis and obtain the group XYZ within a tree that is 100 steps in length. We can then run an analysis and have the analysis find the shortest tree that *does not* contain the group XYZ. Let us say that the resulting tree is 115

steps. The difference, 15 steps, is the Bremer support for the clade XYZ. If we do this for all groups found in our tree, we can obtain Bremer support values for each node containing a monophyletic group. Interpretation: we would have to accept a tree that is 15 steps longer than the most parsimonious tree in order to break up the clade XYZ.

Bremer (1988) argued that a way to quantify total Bremer support for a tree would be to simply sum the values at each node and then divide by the total tree length. However, partly because of the way trees are constructed, especially large trees based on complex data sets, it is not necessarily the case that different individual support values at each node should be thought of as additive across the tree (Faith and Ballard, 1994). This is because the presence of a particular clade within a tree might thereby constrain the appearance of other groups (Gatsey, 2000). Much depends on the distribution of homoplasy. Only if homoplasy is more or less evenly distributed across the tree will Bremer values be additive. If homoplasy is bunched in local regions of the tree, then they are not additive. Gatsey (2000) discusses measures of linked branch support which lead to better descriptions of tree stability than Bremer support alone.

High Bremer-values are almost always associated with strongly supported monophyletic groups and frequently correspond with other measures of nodal support such as the jackknife and bootstrap resampling, discussed below. The major problem is that no one has any idea what does or does not constitute a significant Bremer support value. Like the first method, we feel relatively confident when our clades have high Bremer support and not so confident when they have low Bremer support values. But the problem remains; what is a *significant* Bremer support value?

Jackknife and Bootstrap. Statistical measures of tree support are built on the statistical proposition that some parameter can be estimated from samples drawn from a population and that the result can be evaluated by drawing a new sample from the population. For example, if we estimate the mean body length of a population of mice, we would measure the body lengths of a sample drawn from the population and calculate the mean and standard deviation derived from the measurements. The true mean of the population would be expected to fall within some interval of length. We can check this hypothesis if we go back to the population, select another sample of mice, measure their body lengths, and find that the newly estimated mean falls within the interval we have calculated using the first sample. We can also test the hypothesis that another population of mice has mean body lengths similar to the first population. The problem in applying this strategy in phylogenetic analysis is that we have no new sample on which to draw. In such cases, we can simulate the statistical approach by subsampling the original characters and see if the subsample re-creates the result. If we draw many subsamples and reanalyze our problem with each subsample, we will create a set of trees whose number is the number of times we have subsampled. We can determine the frequency with which groups in the original analysis reappear over the course of subsampling by determining the frequency with which the groups appear in the set of trees.

For example, a strongly supported clade characterized by many synapomorphies might be expected to appear in all of the trees while a weakly supported group might appear only in a few. These are expressed as probabilities, usually by subject-

ing the set of trees to a majority consensus analysis (see below), with the probabilities expressed as the percentage that a particular clade appears on the majority consensus tree. There are two common methods for accomplishing this strategy in phylogenetic analysis, jackknifing and bootstrapping. The jackknife is the older approach and will be discussed first.

The Jackknife. The usual implementation of the jackknife is to rerun the analysis some predetermined number of times while deleting one or more observations without replacement. The trees resulting from this process are saved, and the frequency of appearance of each clade over all the trees constitutes its jackknife frequency. The variability among the trees generated by the analysis depends on the number of observations (data columns) deleted. Although any percentage of the original matrix can be subsampled, two common strategies are employed. The "half-delete jackknife" (Wu, 1986; Felsenstein, 1985a; Felsenstein, 2004) randomly samples half of the characters in each iteration, without replacement. The "parsimony jackknife" (Farris et al., 1996) deletes fewer characters. The half-delete jackknife apparently has properties similar to the bootstrap and, naturally, is favored by Felsenstein (2004) who introduced bootstrapping to phylogenetics (Felsenstein, 1985a). The parsimony jackknife, preferred by Farris et al. (1996), favors strongly supported groups and finds these groups with greater frequency when they are present. That is, such groups will have higher jackknife scores under parsimony jackknife than under half-delete jackknife.

Another approach is to jackknife taxa rather than characters (Lanyon, 1985). The question asked would take the form of seeing what effect the removal of species might have on the subsequent tree. Felsenstein (2004) points out that such a jackknife has no easy statistical interpretation. But, it might have its uses. For example, consider the scenario of the investigator who is analyzing 100 species in a group of 1000 species. It might be interesting to note that removal of 10 species drastically affects the topology, yielding groups not observed in the original analysis. Another application might be a scenario in which some groups were represented by many species and other groups, just as speciose, were represented by few species. Random subsampling would tend to pick taxa from the groups containing large numbers of species. Would this have an effect on the resulting topology that would call into question taxon sampling?

The Nonparametric Bootstrap. The bootstrap was first used in phenetic studies (Mueller and Ayala, 1982; see Felsenstein, 2004) before its introduction to phylogenetics by Felsenstein (1985a). There are two versions, parametric and nonparametric, of which the nonparametric is commonly employed to assess the fit of data to a tree. The idea behind the nonparametric bootstrap is that the matrix is a sample of the true underlying distribution of characters. If we knew the true underlying distribution of characters, then we could assess the degree of support inherent in the data for any clade. Of course, we do not know this, but bootstrapping is a way to simulate the variability in the underlying pattern of character distribution. It is a method for estimating the unknown and presumed true distribution by using the known empirical data.

A nonparametric bootstrap analysis begins by creating a number of pseudoreplicate data matrices by subsampling the original data matrix, with replacement. Each being a matrix of the same size as the original and composed of a random sample

of characters from it. Any one transformation series might be represented in any pseudoreplicate matrix once, twice, many times, or not at all. Each of these matrices is analyzed and the trees collected into a set of trees. The frequency that particular clades appear over the entire set of shortest trees generated by analyzing all of the pseudoreplicate matrices constitutes its bootstrap score (summarized using a majority consensus technique). We can even generate confidence intervals such that if we run a large number of similar bootstrap analyses, we would expect the score for a particular clade to fall within that interval. Intuitively, if a particular clade has a high number of characters supporting its monophyly, and there are few characters that refute its monophyly, then chances are that at least some of these characters will appear in each pseudoreplicate matrix and the group will appear in many or all of the sets of trees. Conversely, if evidence for the monophyly of the group is weak or if there is a high level of homoplasy in the original matrix, the group might not appear at all, or at low frequency.

There is considerable literature on biases in jackknife and bootstrap analyses as applied to phylogenetic analysis, and the consensus of opinion is that the probabilities obtained are usually low relative to the perceived reality of the clade given the data. For example, Hillis and Bull (1993) suggest that bootstrap values as low as 70 percent may indicate well-supported clades, in contrast to the usual statistical threshold of 95 percent (but see Newton, 1996; and see Felsenstein, 2004, for additional literature and discussion of various methods used to reduce bias). The probabilities obtained are not probabilities of the reality of the clades per se, but reflect the relative support of the clades in the data matrix given the assumptions of the analysis.

Permutation Tests. Permutation tests have been used in phylogenetics, but every application has been to a greater or lesser extent controversial. Both Swofford et al. (1996) and Felsenstein (2004) provide examples of the application of these tests and the controversies surrounding them. In general, permutation tests fall into two categories.

Permutation tail probability tests (Archie, 1989; Faith and Cranston, 1991) are designed to test for hierarchical structure. The test works by shuffling characters in each data column and assigning them randomly to species. We would expect that if we analyzed any single shuffling of the data the result would be much worse than an analysis of our original matrix, if our original matrix contained real hierarchical signal. Alternatively, if our original data was itself comprised of randomized data that lacked any phylogenetic signal then we would not expect to see a difference between our original data and a randomized version of our original data in such parameters as tree length or ensemble consistency indices. Of course, it is possible, by chance, that a randomized version of our data might yield a tree with as much support as or original data, but we would not expect to see this very often if our tree was supported by "good" data. We can specify how often we might expect a randomized matrix to perform as well as our original data and that expectation is the usual statistical expectation of p = 0.05. If we do a great number of permutations, derive a tree from each permutation, and calculate its length, we can build up a distribution of tree lengths (or other measures). If only a small percentage of these trees are as good as our tree derived from the original data, then we reject the null hypothesis that there is no difference between our original tree and a tree derived from random data. We conclude that the data contain hierarchical signal. A variant

test, the topology-dependent permutation tail probability test (T-PTP), was developed by Faith (1991) to test in a similar fashion whether specific clades are supported.

Incongruence Length Difference. The second major use of permutation tests is to test the null hypothesis that two data sets are inferring different trees. Rejection of the null hypothesis implies that both data sets infer the same or highly similar tree topologies. Data are combined and permutations are conduced by permutating the data columns (not the order of characters). This test was suggested by Farris et al. (1994a, b) and independently introduced in PAUP by Swofford in 1995 as the partition homogeneity test.

Measure of Skewness. Imagine the situation in which we could determine the length of every possible tree and plot a histogram of tree length frequency. The peak would contain many less parsimonious trees, trailing off to fewer very long trees on one side and fewer very short trees on the other side. Hillis (1991) suggested that if this frequency distribution was skewed, it suggested strong phylogenetic signal because there would be far fewer relatively short trees than long trees (see also Huelsenbeck, 1991a). By contrast, trees based on low quality or effectively randomized data should show little skewness and instead symmetry of tree-length distribution. Hillis (1991) proposed using the g_1 statistic, a measure of tree-length frequency distribution, as a way of assessing phylogenetic signal in a data set. The resultant g_1 statistic could be compared to the distribution of g_1 statistics produced from randomized data to assess the degree of significance. The proposed test is an interesting one, but it has been suggested that the tree-length frequency distributions of some data sets possessing or lacking phylogenetic signal do not always behave in such a stereotypical manner (Källersjö et al., 1992). Because of this, and for other reasons, it has been suggested that tree-length frequency skewness "may be of limited power in detecting phylogenetic signal" (Felsenstein, 2004:363).

USING CONSENSUS TECHNIQUES TO COMPARE TREES

Topologically different trees can be combined to explore their common and unique features using consensus techniques. A consensus tree is a summary of the common topological features of two or more trees that contain the same taxa and differ in details of their topology. Usually rooted trees are compared; however, consensus techniques can also be applied to unrooted trees. It is entirely possible that two unrooted trees are indistinguishable, yet two rooted trees derived from them can be in conflict, simply by specifying the root in different locations.

There are several techniques that are used to derive consensus trees. Each has its strengths and weaknesses. Three kinds of consensus trees are commonly employed, and several other kinds have been used.

Strict Consensus (Rohlf, 1982). In phylogenetic analysis, strict consensus trees contain only those monophyletic groups that are common to all of the trees compared (Fig. 6.20). Strict consensus trees are a mechanism to convey a reduced tree that highlights the common branching patterns, common speciation events for which all most parsimonious trees agree or for which trees derived from different data sets of the same taxa agree. It asks the question: what groups are always monophyletic? This is the type of consensus tree most frequently employed in phylogenetic studies.
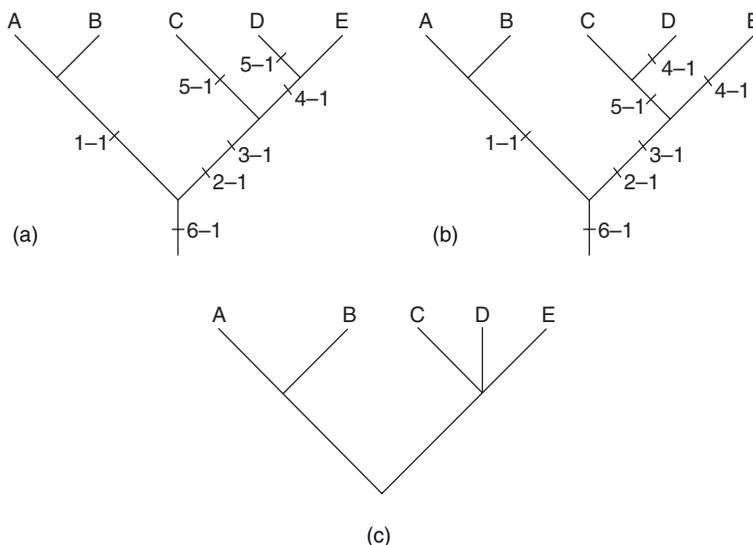
Figure 6.20. Strict consensus. (a–b) Two equally parsimonious trees. (c) The strict consensus of the two trees.

It does not give the best estimate of phylogeny and character evolution, because any of the most parsimonious trees provide a more parsimonious explanation of character change. However, the strict consensus provides an efficient summary of how the different most parsimonious trees agree. Thus, it is a way of reflecting common monophyletic groups.

Majority-rule Consensus (Margush and McMorris, 1981). Strict consensus trees are actually one extreme of the Mℓ family of consensus trees of Margush and McMorris (1981), where ℓ varies between 50 percent and 100 percent and is defined as the percentage of times a particular group appears among all of the trees compared. A commonly reported tree is the 50 percent-majority rule consensus tree, which reports all clades that appear in more than 50 percent of the trees. A strict consensus tree is simply a M_{100} tree, and a M_{75} tree would report all groups that appear in 75 percent of the trees compared, etc. The common output format is a tree with nodes that report the percentage of the occurrence of the resolved groups. The Mℓ family of consensus trees (including strict consensus) can be used either for rooted or unrooted trees.

Majority-rule consensus trees are sometimes reported in the parsimony literature, but usually only in cases where the strict consensus is poorly resolved and the scientist wants to put a more positive spin on the results. While it is tempting to consider them a "best estimate" of phylogeny, they are not so under the optimality criterion of parsimony because there remain alternative equally parsimonious trees. Further, there is no reason to suppose that the percentage of times a particular tree resolution appears in a result has any significance given that sometimes quite different results are equally parsimonious. The validity of this technique as a general way of presenting parsimony results is dubious.

These consensus techniques are more commonly used as output on statistical tests such as the jackknife and the bootstrap, as discussed above. In these cases, the

majority-rule consensus values are used to assess the fit of the data to the tree. They may be a reasonable proxy for qualitative degree of confidence in results.

Adams Consensus (Adams, 1972). Adams consensus is strictly for rooted trees (Felsenstein, 2004). Informally, Adams consensus seeks the highest resolution possible for two trees and accomplishes this by moving inconsistent taxa to basal positions where they do not conflict. The result may not necessarily reflect monophyletic groups supported in the original analysis, even those common to all trees (as in strict consensus).

Adams consensus has been used to find taxa that are "unstable" on a set of trees that are otherwise similar. They can be used to answer two kinds of questions. First, what is the most highly resolved tree that will identify "problem" taxa? An example of a problem taxon might be a species of hybrid origin with partial expression of synapomorphies of each group that contains its parental species (Funk, 1985). Second, are the trees logically consistent? Most computer packages can calculate Adams consensus trees. Wiley et al. (1991) provide some simple examples that can be calculated by hand.

There are a number of other consensus techniques, including semistrict or combinable component consensus (Bremer, 1990) and Nelson or Nelson-Page consensus (Nelson, 1979; Page, 1990; Swofford, 1991; Felsenstein, 2004), and consensus techniques that are based on branch lengths (e.g., Neumann, 1983) or path distances (e.g., Lapointe and Cucumel, 1997). Bryant (2003) provides a list of consensus techniques and their uses in phylogenetics and Felsenstein (2004) provides an overview.

STATISTICAL COMPARISONS OF TREES

A class of tests termed *paired-sites tests* can be used to test the null hypothesis that two tree topologies are statistically identical. For example, we may have the most parsimonious tree found in our own analysis and wish to see if it is statistically different from a tree of the same taxa that was previously published. Or we may wish to test the proposition that our most parsimonious tree is significantly different from a suboptimal tree that is 10 steps longer. There are parametric and nonparametric versions of the paired-sites test. Intuitively, if there is very little difference between the trees, there should be very little difference in the variation of site performance from one tree to the other and the null hypothesis would be confirmed. Alternatively, if there are large differences at many sites, then we might expect the null hypothesis to be rejected and conclude that there is a significant difference between trees.

We may wish to compare our data and results with the work of others. Or we may wish to see if our shortest tree is really that different from a tree that is almost as short. There are a number of ways of accomplishing such tasks. The most common technique is the Wilcoxon signed ranks test first introduced to parsimony analysis by Templeton (1983a, b; see also Felsenstein, 1985b), or its simplified version, the winning sites test (Prager and Wilson, 1988). Some programs, such as PAUP* and Mesquite, can be used to calculate the Wilcoxon signed ranks test and the winning sites test statistics. Other tree comparison tests are parametric and, thus, require models of evolution. We will discuss these in the next chapter.

WEIGHTING CHARACTERS IN PARSIMONY

Character weighting in parsimony can take several forms, including selecting a priori characters (what characters to include in an analysis) and assigning a particular cost of transformation to one or more characters. Character selection is common in morphological analyses and also present in molecular analyses in the form of gene selection. A priori weighting assigns the cost of a particular transformation prior to an analysis and usually is based on some implicit model of evolutionary changes envisioned by the investigator. A posterori weighting is performed after an analysis and is based on assumptions regarding character performance in that analysis.

Character selection. Character selection is an inevitable consequence of the inability to examine all of the characters of specimens. How characters are selected may profoundly affect the results of a phylogenetic analysis. We suspect, but cannot prove, that the reason many morphological analyses work well in a parsimony framework is that the investigator picks characters that show (1) a fairly low level of intra-taxon variation and (2) an interpretable level of inter-taxon variability. This would result in the informal adoption of a model of evolutionary change that favors parsimony analysis so long as the covariation of true synapomorphies is greater than the covariation of true homoplasies. However, because we do not know the distribution of true synapomorphies over that of true homoplasies, we are betting that the level of homoplasy is relatively low compared to that of homology. Note that this differs rather strongly from typical molecular analysis when entire contiguous regions of DNA or amino acids are analyzed and there is no a priori picking of properties that show these characteristics. In a DNA analysis, we may find characters that do not vary in their states and have no "control" over intra-taxon variability. In molecular analysis, character choice resides in the gene regions picked for the analysis.

A Priori Weighting

All forms of parsimony (Fitch, Wagner, etc.) are special cases of applying Sankoff optimization, with "Sankoff characters" being those characters where the cost of transformation between states is specified by the investigator (e.g., Sankoff, 1975; Sankoff and Rousseau, 1975; Sankoff and Cedergren, 1983; Swofford and Maddison, 1992; Goloboff, 1998). As Swofford has pointed out, treating all character states as equally weighted is assigning a weight of one to the weighting function. Character weighting is different from ordering. Although the investigator may think that no evolutionary assumptions are invoked when he or she decides to treat all characters as equally weighted and unordered, he or she has, in fact, made an explicit evolutionary assumption that each transformation has the same information content.

What we usually think of as weighting in parsimony is the activity of assigning different costs (in steps) to different kinds of transformation. The weights we assign are based on what we think are the probabilities of the transformation of one character state to another. Modern parsimony programs make it possible for the investigator to assign different weights to different characters and to assign different weights to the transformation of different states within a character.

A Priori Weighting: Parsimony-Specific Weighting Functions. Coding characters might imply different step-costs under certain conditions in a parsimony analysis and optimization. Some of these costs are bound up in selection of a parsimony

criterion. For example, Fitch parsimony introduces equal costs between transformations while Wagner parsimony may impose different costs (zero to one costs one step; zero to two costs two steps). Others costs fall into the class of a priori weighting as a function of general parsimony where differential cost is assigned to characters within the context of an overall parsimony analysis.

In modern computer packages, differential weighting of states may be easily accomplished by using one or more step or cost matrices and referring particular characters to these step matrices during analysis (Maddison and Maddison, 1992; and Ree and Donoghue, 1998). Maddison and Maddison (1992) discuss a simple example of two characters with five character states, one treated as with Farris optimization (Fig. 6.21a, b), and the other with Fitch optimization (Fig. 6.21c, d). The cost of transformation in each case is entered into the matrix according to the evolutionary model adopted, and these costs are added to the weight of the edges when the tree is calculated. Another example would be differential weighting of transversions and transitions among sequence data (Fig. 6.22).
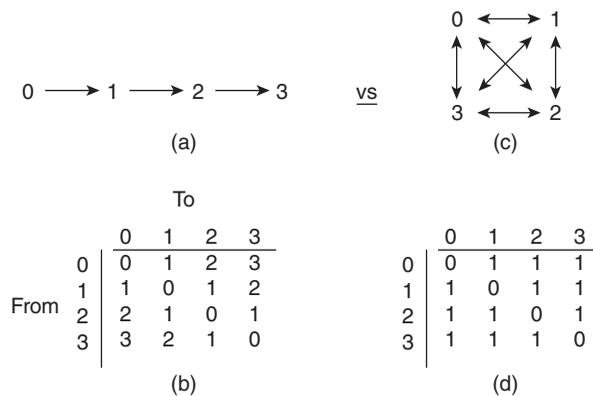
Figure 6.21. Character relationships and cost (or Sankoff) matrices. (a) An ordered and polarized series of character states. (b) A cost matrix expressing the cost of transformation between each state in (a). (c) An unpolarized and unordered series of states. (d) The cost matrix expressing the cost of transformation between each state in (c). From Maddison and Maddison (1992).
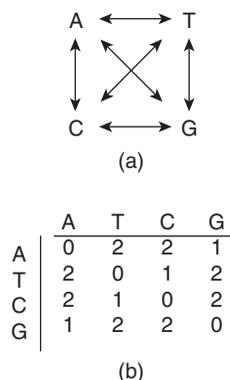
Figure 6.22. An example of a molecular cost matrix. In this case we have an unpolarized and unordered series of character states, but we have set the cost of transversions twice as high as the cost of transitions.

Weighting by Performance

Two common methods of performance weighting are successive approximations (Farris, 1969) and self-weighted optimization (Goloboff, 1997). Each is founded on the proposition that parsimony analysis can be refined by taking into account the performance (in units such as steps) of characters, given a topology and on the idea that the search for the final topology can be "informed" by character performance.

A successive approximation (Farris, 1969) performs an equally weighted analysis and then proceeds to a round of analyses where characters are weighted according to the consistency indices in the previous analysis. Some have suggested that this approach is circular (e.g., Cannatella and de Queiroz, 1989; Swofford and Olsen, 1990). Others have argued that it is recursive (Carpenter et al., 1993; Carpenter, 1994). Felsenstein (1981a) likens it to compatibility analysis. In such analyses, the investigator attempts to circumvent circularity problems by taking the average consistency index (for example) over all most parsimonious trees rather than from a single topology. This approach is probably the one that is most consistently applied in parsimony analyses that apply successive approximations.

Goloboff's (1997) self-weighted optimization takes a different approach based on Farris' (1969) concave fitting function (Fig. 6.23). In developing successive approximations, Farris (1969) developed a series of functions to illustrate the relationships between the relative weight of a character and its influence on the final tree. For equally weighted characters (in Fitch optimization), the relationship is linear because transformation of both homologous and homoplastic characters equally influences the tree topology. If some characters are weighted more than others, there are two possibilities. If characters with homologous states are weighted more than characters with homoplastic states, the fit function becomes concave. If the opposite obtains, then the fit function becomes convex. Goloboff's (1997) self-weighted optimization is built on his earlier work (Goloboff, 1993) of evaluating
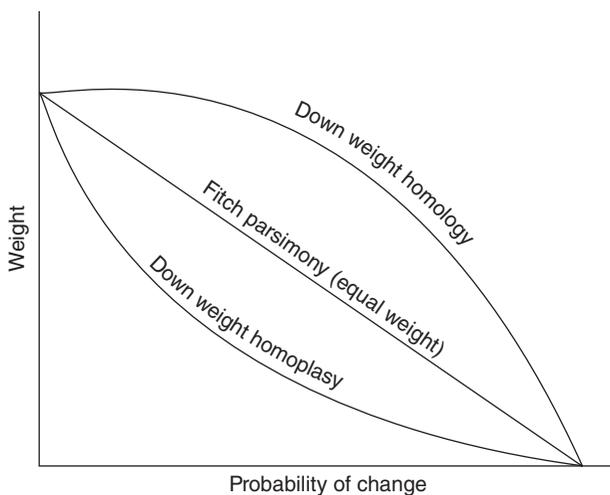
Figure 6.23. Three functions showing the effect of character weighting on the fit of probability of changes on a tree topology (from Farris, 1969).

trees relative to the weight they imply, and the weight of a tree is related to the concavity function.

Consider a matrix with no homoplasy. A tree mapping this matrix would have a linear function, and the lack of homoplasy would mean that there is no distortion. Distortion can be seen as a lack of congruence and the more the homoplasy, the greater the distortion. Now consider a matrix with some level of homoplasy. The linear fit is not perfect, creating distortion. However, if we weight the homoplasies less than the homologies for that particular tree, we minimize the distortion, minimizing the concavity that is created by the presence of homoplasy. The Goloboff (1997) method is to combine tree searches and weighting in an attempt to find the best concavity function for the tree and data as a whole.

Goloboff (1997) reviews criticisms of his technique. The most interesting charge is that the results are not parsimonious (Turner and Zandee, 1995). Goloboff's (1997:236) answer to this criticism is interesting: adopting an optimality criterion based on a concave rather than linear fit function is a "refined way to measure parsimony in trees."

Weighting by Character Elimination

In general, phylogeneticists are loath to eliminate characters once they gather them (which does not mean they are loath to not gather characters if past experience suggests they are bad performers). There are times when weighting by character elimination can be reasonably employed, and this usually involves some understanding of the strength of particular homology statements. For instance, certain regions of DNA such as some loops in 16S mitochondrial ribosomal DNA, introns between coding regions of protein-coding genes, etc. are so entropic (randomly evolving; see Brooks and Wiley, 1986) as to make homology matches impossible or meaningless. Therefore, such regions are often avoided in molecular systematic studies and this is a repeatable and valid weighting criterion.

Weighting: Concluding Remarks

Any form of weighting, including equal weighting, assumes certain things about the evolutionary process. Differential weighting appears to assume more than uniform weighting because it implies that the investigator knows something about the relative behavior of one kind of transformation relative to another kind of transformation. It is a form of model selection, albeit not a statistical form of model selection. We take up the statistical form of model selection in the next chapter.

PHYLOGENETICS WITHOUT TRANSFORMATION?

In the first edition of *Phylogenetics*, Wiley (1981a) took considerable time to discuss two alternative systems of systematic inference: Phenetics and Evolutionary Taxonomy. We largely bypassed that section in this edition because we thought that controversy was resolved. However, a new version of systematics has appeared that competes with Hennig's system that merits a short discussion, the idea that phylogenetic analysis can be performed without the assumption that characters transform

during the course of evolutionary history. This particular idea is not new; it can be traced back to what Platnick (1979) termed the "transformation of cladistics" and was first fully explicated in Nelson and Platnick (1981). What is amazing is that its recent incarnation, exemplified by a recent book by Williams and Ebach (2008), actually accuses those of us who practice traditional Hennigian principles of being "pheneticists."

The method of analysis is usually termed three-taxon analysis (3ta). It attempts to avoid the entire idea of transformation, and thus the entire idea of evolutionary descent, by analyzing presence-absence matrices where the state zero simply means the absence of the state one. Starkly, Scotland provides the following statement that exemplifies the basic idea of what he terms a "complement relationship":

> For example, paired appendages (fins+limbs) constitute a homology at the level of gnathostomes. Within gnathostomes, fins do not form a group and are therefore non-homology. Forelimbs diagnose a group (tetrapods) and are homologous at the level of tetrapods (Scotland, 2000:488).

In opposition to the kind term of *homology proposition* described above are what he terms *paired homologs* entertained by standard cladistic analysis. The unwary may be misled unless they pay careful attention to what Scotland is actually saying. Note that he uses the term *paired appendages* and includes in that category both fins and limbs. This obviates the need for transformation. Did the ancestral species of all other gnathostomes have both fins and limbs? This seems to violate Patterson's (1981) conjunction criterion: angels cannot have both arms and wings if arms and wings are homologous.

So far as we can see, paired homologs are simply hypotheses of transformational homology; for example, hyomandibular and stapes or pectoral fin and front leg. Three strong claims by Scotland are: (1) "Standard cladistic analysis" never tests the transitional proposition represented by "paired homologs." The relationship is "simply assumed" (Nelson, 1994; Pleijel, 1995; and Carine and Scotland, 1999 are cited as the source of this statement). (2) Because characters do not give rise to other characters (Sattler, 1984, 1994), the entire idea of the transformational view is in question. (3) The final claim seems to be that because characters and their states are hypotheses, they cannot be said to have participated in any real processes (Weston, 2000, is cited for this point).

Claim (1) is false when we consider the entire process of analyzing characters. Transitional propositions are tested in many systems and for day-to-day phylogenetic analysis these usually take the form of applying the various criteria discussed by Remane and by Patterson that take the final form of columns of data and the transformational hypotheses they contain (i.e., pairs of plesiomorphic and apomorphic homologs in the binary case). Further, transitional propositions can be refuted after one analysis and before another; that is one part of Hennig's reciprocal illumination idea. It is true enough that the relationship between plesiomorphic and apomorphic homology pairs (or triplets, etc.) are not directly tested during a phylogenetic analysis, but to say that they are not tested at all implies that no thought has gone into gathering empirical data concerning the plausible nature of the relationships before matrix construction and that no thought is given to the results obtained after the analysis. The idea that it is "simply assumed" that there is a his-

torical relationship between pectoral fins and forelegs or between the hyomandibular and staple is, in our opinion, not valid. We note that Scotland (2000) acknowledges what he terms "deep homology," which is simply the idea that it is information in the genome and epigenetic phenomena that are behind the structures we study, but yet he did not get the tetrapod limb correct. Yes, it is true that fin rays are not homologous to autopodial bones, but everyone already knew this, even before the work of Shubin et al. (1997), and the endochondral skeleton of the pectoral girdle includes more than radials, axials, and fin rays. It also includes a scapula and a coracoid and so do tetrapods. Further, it is a matter of rearranging the expression of certain genes that seems to be behind the transformation of fins to limbs, making the transformation understandable.

Claim (2) is dealt with in Chapter 5. But let us expand on it further. It is as nonsensical to claim that "complement relationships" (identity statement in our terms, hypothesizing that similar characters in different organisms are the "same" character) are an illusion as it is to say that paired homologs are an illusion given the reasoning of Sattler (1984, 1994). Why just single out paired homologs? Complement homologs do not give rise to complement homologs any more than plesiomorphies give rise to apomorphies. Instead, as we outline in Chapter 5, information that specifies how to build a complement is passed to each generation and the complement is built anew each generation from the previous generation. The difference between complements and pair homologs is that some of the information has changed during the transmission of that information. Nothing new here unless you wish to expunge all of biology from consideration instead of simply expunging evolution. Change, of course, is nothing but entropy in action (Brooks and Wiley, 1986), and change is as easy as falling off a log. It's stasis that is hard to explain, not change. In short, one can account for neither sameness nor transformation without attending to underlying processes. Far from providing a rationale for rejecting transformational homologies, the observation that characters do not give rise to other characters is cause for rejecting pattern analysis in general.

Claim (3) is the most curious of all. All that we recognize, whether complement or paired, are conjectures about the regularities of the world. Is Scotland arguing that there is no way of studying processes in the world at all? Characters and states are data about the world. All data about the world are hypothesis-bound. Complements are as much data (and identity-statement-theory-laden) as paired homologs. Can we not talk about the process of gravity because our data on falling objects are, in the end, data hypotheses?

Variation on methods of three-taxon analysis have been presented by Nelson and Platnick (1981), Carine and Scotland (1999), Scotland (2000), and Williams and Siebert (2000), as well as Williams and Ebach (2008). Criticisms of three-taxon analysis can be found in Farris and Kluge (1998), Farris et al. (1995), Kluge and Farris (1999), and a recent review of the Williams and Ebach volume by Farris (2010). Williams and Ebach (2008) claim that transformed cladistics and its attendant method, three-taxon analysis is the true phylogenetics, and that everyone not connected to transformed cladistics is actually practicing phenetics (!). We do not consider three-taxon analysis (or pattern analysis in general) a phylogenetic technique and rather than review its procedures in detail refer the reader to the works cited above, both pro and con. Our opinion: while the phylogenetic tent is big enough to include such diverse approaches as parsimony, likelihood, and Bayesian

approaches, it is not big enough to include three-taxon analysis with its need for a priori character ordering, reliance on irreversibility, rejection of reversals as synapomrophies, and other attendant methods and assumptions.

CHAPTER SUMMARY

- Parsimony analysis is performed under the assumption that the best estimate of phylogeny is that tree which is the shortest tree, measured by the number of evolutionary transformations among the characters.
- Phylogenetic analysis may be performed by polarizing the characters a priori and employing rules of character inclusion and exclusion. This is an algorithmic approach and the one used by early phylogenetists to reconstruct phylogenies.
- Computer-assisted phylogenetic analysis may take either the algorithmic or criterion-driven approach, but the criterion-driven approach is usually employed.
- A large number of increasingly sophisticated computer programs are available for parsimony analysis.
- Trees may be evaluated using certain data summaries such as tree length and consistency indices and fit of data to result, including Bremer support, jack-knifing, and bootstrapping. Parsimony trees may also be compared using other techniques such as the Wilcoxon signed ranks test.
- Systematics without transformation and outside the evolutionary paradigm is not phylogenetic whatever its other qualities might be.

7

PARAMETRIC PHYLOGENETICS

> Given that we select estimated evolutionary trees always according to the maximum likelihood criterion, the method for constructing an estimated evolutionary tree is in principle well determined once a stochastic model of the evolutionary process has been selected.
>
> —James S. Farris, 1973

Parsimony analysis is only one of several methods for analyzing phylogenetic relationships. In this chapter, we examine two other approaches: maximum likelihood (ML) and Bayesian analyses. Huelsenbeck and Crandall (1997) provide a general introduction to likelihood analysis. Lewis (2001a) provides a general introduction to Bayesian analysis, and Holder and Lewis (2003) provide a useful introduction to several approaches including ML and Bayesian and parsimony analyses. ML and Bayesian analysis are similar in using likelihood calculations as the basis for inference, but they are very different in their philosophical approach to problem solving. ML methods use a criterion-based approach: the preferred tree is the tree that has the highest probability of producing the data we observe given a specific model of evolution adopted by the investigator, the tree topology, and the branch lengths between nodes. The model is used to calculate probabilities of observing the data on a specified tree, one transformation series at a time. It differs from parsimony in taking branch length into consideration relative to an explicit model of change from one character state to another along the branch. Specifically, the longer the branch, the lower the probability that matches (e.g., identical base residues) appearing at the base of the branch and the tip of the branch are homologous. These probabilities

are then multiplied over the entire tree to produce a likelihood score. This is often expressed as a logarithm of the likelihood for purely computational reasons. When estimating the topology, the tree with the highest log-likelihood scores is accepted as the best estimate. Like parsimony, likelihood produces a point estimate of the best tree given the criterion such that the fitted model (data fitted on a tree topology and other parameters) is the "best model."

Bayesian analysis uses likelihood calculations, but stands the probabilities on their head by estimating the probability of the tree topology given the data and the model rather than the probability of the data given the model and tree topology. It does so by calculating a posterior probability, a probability that is conditional on what the investigator is willing to accept as true before the analysis. Most frequently, this prior probability states that all alternative trees are equally probable, a priori, but it is entirely possible to specify a particular tree topology as the most probable. Bayesian inference offers an entirely different approach to statistical inference and is controversial among statisticians, as we shall discuss below. As practiced by phylogeneticists, in Bayesian inference the criterion employed is that of maximizing the posterior probability of the tree given the data and model of inference. It uses algorithms that are designed to explore probability space to find the area(s) where the probability density is highest and, thus, expected to be the area where the tree(s) with the highest posterior probability reside. Unlike either likelihood or parsimony, Bayesian analysis does not produce a point estimate of the model, but rather, a probability distribution of models that may contain one to many tree topologies.

The use of explicit models of evolution to reconstruct phylogenetic relationships and accompanying statistical tests has increased tremendously in the past 20 years, due in part to the increased availability of computer packages capable of performing such analyses (e.g., PHYLIP, PAUP, MrBayes) and the application of the method of Felsentsein (1981b) in these programs that dramatically decreased the computational burden of the calculations. However, the idea that likelihood models could be used for phylogenetic inference dates back to at least Cavalli-Sforza and Edwards (1964).

A phylogenetic researcher might be motivated to incorporate evolutionary models for a number of reasons. For instance, investigators may believe that they understand the evolutionary dynamics of some kinds of data well enough to model their expected evolution over a particular phylogeny. This a priori belief is no different in kind than that of an investigator who weights the performance of certain characters in a parsimony analysis or even the a priori choice of one kind of character over another; any such activities may or may not be warranted. Both investigators are motivated to increase performance of the analysis, increasing what they believe is the veracity of the resulting phylogenetic hypothesis, by applying their experience to the problem. The outcome, for better or worse, is left to other investigators and subsequent researchers to evaluate. If the evolutionary dynamics of the model capture something real about the data, the veracity of the result is better. If led astray, the result may be worse.

The growth of this particular branch of phylogenetics is likely to continue at an almost exponential rate. In this chapter, we will approach statistical phylogenetics in a narrative fashion. We will cover the basics of likelihood-based approaches with simple examples and concentrate on how they work, keeping mathematical formali-

ties to a minimum. The goal is to understand how such approaches work and why we should choose, or not choose, to use them. We leave the details to others who are better qualified to describe the mathematical formalities in detail (e.g., Swofford et al., 1996; Felsenstein, 2004; Yang, 2006). We also recommend Sober's (2008) account of the philosophical basis for inference as it relates directly to issues of likelihood, Bayesian, and parsimony inference.

MAXIMUM LIKELIHOOD TECHNIQUES

Likelihood is the probability that an event that has happened in the past would yield a specific outcome. ML is a procedure for finding the value of one or more parameters for a given model that makes the likelihood of observing the data at hand attain its maximum value. Consider a very simple example of some data taken from a population of fishes, the length and height of the body. We have adapted this example from a similar example given in Forester and Sober (1994) and will return to it later in the chapter. We might attempt to fit a line to these data to investigate how they are related. Two such lines are shown in Fig. 7.1. It should be obvious that the bottom line fits the observations better than the top line. If we consider both lines to be "models," then we can intuit that the data are better explained by the bottom line than the top line and, thus, the bottom line would have a higher likelihood value than the top line. Indeed, we can fit any straight line to these data, calculate a likelihood value under the assumption of a normal distribution of errors, and compare it to our lower line. Some will be a very bad fit, but some will be a fairly good fit.

As revealed by this very simple example, the basic idea of ML is quite simple: the best line is the line that is most consistent with the observations. It "predicts" (postpredicts or "postdicts") the data better than other lines. Consistency is measured statistically, by the probability that the observations should have been made if the line-model is taken as true. We can adjust the model (including its parameters) and test whether a new line is better than the old line. If the likelihood goes up,
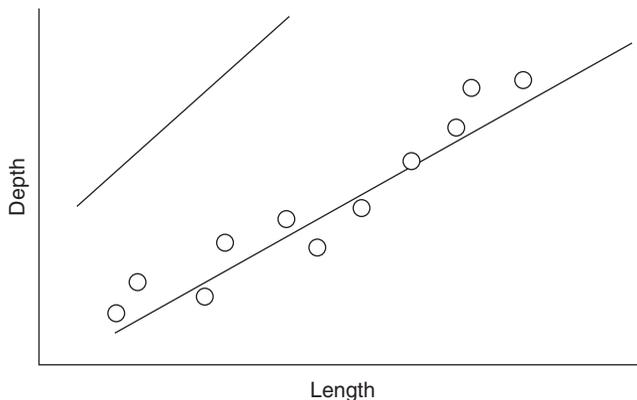
Figure 7.1. A scatter plot with two regression lines. Intuitively we can guess that the lower line fits the data (circles) better than the top line.

then the new line predicts the data better than the old line. We can continue to adjust the parameter values until such time as we fail to obtain a higher likelihood value. Or we can adopt an entirely new form of model, for example, a curved line rather than a straight line. It is actually more complicated, but this will get us started.

You might ask: why not simply calculate the regression line using the least squares method? Of course, you could do this and you probably would, given this simple example. If you simply calculated the regression, you would be using the algorithmic approach rather than the criterion-driven approach. Algorithmic methods work well in many areas of statistical analysis, and the algorithmic approach is preferred because it arrives at the ML solution in a faster and more efficient manner (just as algorithmic methods are faster than criterion-based methods in parsimony). Indeed, the regression line is the ML estimator. But phylogeny reconstruction is harder and the parallels of algorithmic versus criterion-driven approaches discussed in the last chapter are entirely applicable to likelihood-based analyses of phylogenetic problems.

Most explanations of ML methods begin with a demonstration of coin flipping. Read (2000) uses a very simple example that we believe illustrates the process in a way that can be followed through to more complex as well as Bayesian examples. Consider a population of organisms. The data consists of antennae lengths. We sampled 20 individuals. What we need now is a model. Where we get the model depends on our past experience. In this case, our experience leads us to assume that populations of measurements such as these have some unimodal distribution, central tendencies, and that the observations are expected to deviate from the central tendency in some predictable fashion given certain assumptions about variation. Thus, we will expect the population of measures at hand have some mean value and a standard deviation (SD), which is a measure of the scatter of the data points relative to the mean. The simplest case is to consider the central tendency of the population from which the samples were drawn is μ and the variation around this tendency is σ^2. If we assume that the data are distributed according to a normal (or Gaussian) distribution, then we can calculate the likelihood of the data given the model by calculating the probabilities of each piece of data for any chosen values for the mean and standard deviation. We can compare the fit of the data to the model and its associated parameters.

Consider Fig. 7.2a. This figure is a plot of the normal curve with a mean antenna length of 5 mm and a standard deviation of 1. The x-axis is the antenna length, and the y-axis is the probability density associated with finding a particular value of antenna length. We will tell you now that these parameter values are the true parameter values for this population. Note that the highest probability densities [p(x)] for observing particular antennae lengths are associated with the value of 5 mm and that the probability density decreases as antenna length increases or decreases from the mean value. Another way of saying the same thing is to say that our probability density is high (more measures have a higher probability of being observed). This is intuitively understandable; we expect to observe values closer to the true mean more frequently than values farther away from the true mean. In other words, we have the expectation that the probability of a measure close to the mean is larger than the probability of a measure farther from the mean. We have also labeled the particular lengths of two specimens. The line from the specimen value to the curve is a measure of the probability density of encountering the particular measurement
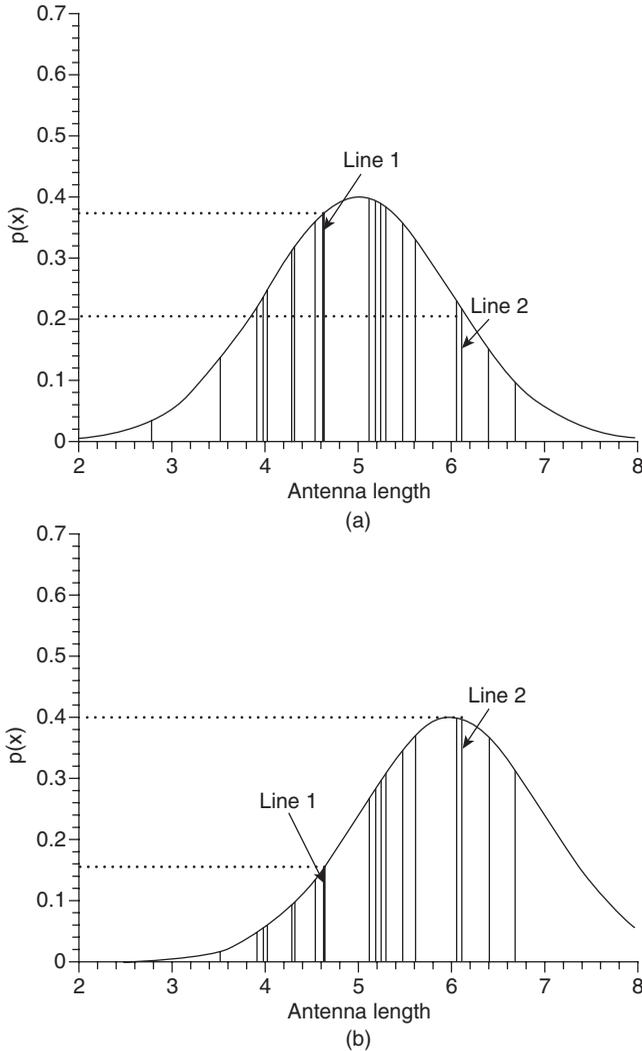
Figure 7.2. Two models of the mean and standard deviation of the antenna lengths for a species of insect. X-axis is antenna length. Y-axis is the probability of observing a measure. (a) Model with mean = 5 mm, standard deviation = 1 mm. (b) Mean = 6 mm, standard deviation = 1 mm. Each line drawn from the x-axis is a measure drawn to the length of its probability of occurrence in the population. The dotted lines are drawn from selected measures to the probability axis. Modified from Read (2000), used with permission.

given the model parameters (i.e., given that mean = 5 and standard deviation = 1). We can read off the values of each measure on the probability axis.

Now consider Fig. 7.2b. The same plots for the same measured antennae are shown. We have kept the same general model, but we have changed the model parameter values. In this case, the model has a mean of 6 mm and a standard deviation of 1. Note that Line 1 is shorter, indicating that the probability of encountering this measurement is less in this model than in the model shown in Fig. 7.2a. Note
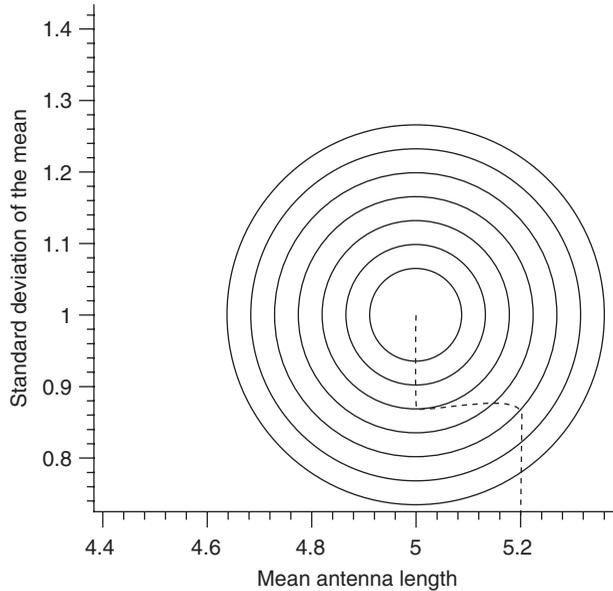
Figure 7.3. A contour map of the probability density space in mean and standard deviation for the example in Fig. 7.2. The dotted line is an imaginary route taken by fitting different values to a fitted model from less probable to more probable. Modifed from Read (2000), used with permission.

that Line 2 is longer indicating that the probability of encountering this value is greater if the mean of the population is 6 mm and the standard deviation is 1. Now, consider all 20 measurements and the two models. If you took the time to actually measure the lengths of each line, you would find that encountering these 20 measurements is more probable if the mean = 5 mm than if the mean = 6 mm, because the total length of all lines together is greater when the mean = 5 than when the mean = 6. The same is also true if you compare a model with mean = 5 and standard deviation = 1 with a model where mean = 5 and standard deviation = 0.6 or 2.0. Thus, we conclude that a fitted model with the parameter values mean = 5, standard deviation = 1 maximized the likelihood of observing the 20 antenna lengths. Although we have visualized these parameters as simple graphs, we can also visualize the relationship between the standard deviation and the mean as a contour map (Fig. 7.3). Such a map could be prepared by plotting the results of many likelihood models. It permits a look at the models relative to the data as compared to the data relative to the models.

The connection between ML and simple statistics in this example is transparent. The ML of the data corresponds to a fitted model where the mean and standard deviation of the sampled lengths follow a normal distribution and have values of mean = 5 mm and standard deviation = 1.0 (the standard deviation being biased relative to the population standard deviation).

In this simple example, the choice of using this particular model may seem unproblematic. But there are deeper issues. Models are simplifications, which means

that in all cases none of the models one is evaluating is a true model. Sober (2008) makes the point that the value of models is not whether they are true but whether they are good at predicting in the face of new data. If they are successful in predicting new data, then the fit of data to the model may lead to something approximating truth in nature. In this example, which is a simple problem of estimating means, the fitted model predicts that the next draw of samples will have estimated means and standard deviations within a range predicted by the fitted model mean standard deviation. If this prediction is fulfilled, then the fitted model will have led to a valid estimate of something that exists in nature.

Simplicity

This example leads us to an interesting feature of models and their parameters. It is embedded in philosophical controversies surrounding simplicity and parsimony. Philosophers have often wondered why simple explanations are to be preferred over more complex explanations (e.g., Sober, 1975; Forster and Sober, 1994; Sober, 2008). There are actually two questions. Given a model and some parameters, which values assigned to the parameters do the best job of explaining the data? This one is fairly easy. In Fig. 7.1, we see that our best-fitting regression line explains the observed data better. In parsimony analysis under a model that all changes are equally weighted, the shortest tree is to be preferred over a longer tree because it describes the data better (Farris, 1983). However, there is also the question of which parameters to choose for the model. Do we choose the smallest number of parameters that adequately "explain" the data? In the example involving means (Fig. 7.2 and discussion above), there are several kinds of models. We can distinguish them by their parameters. The first model contains the mean and standard deviation; a second contains mean, standard deviation, and skewness. The models are nested, because the model containing the parameters mean and standard deviation are included in the model containing mean, standard deviation, and skewness. In this case, it is a simple matter to determine if there is a significant difference in likelihood between the models. A similar scenario can be applied to phylogenetic analysis. Does adding a new parameter result in a significantly better fit to the data? If so, add the parameter. If not, the parameter is not contributing to the resulting estimate in any significant manner.

Akaike (1973) suggests that we can determine the need to add extra parameters in explaining the data at hand by taking into account the relative "burden" of these extra parameters in explaining the data. Given the data in our example, 20 measurements, and that the population is slightly skewed in its observed distribution of data values, we can intuitively (and statistically) see that adding a parameter for skewness does not significantly improve the fit. Akaike (1973) provides a formal way of determining the burden of adding extra parameters that takes into account the simplicity of models as well as their ability to fit the data. Forster and Sober (1994) provide an accessible description of these formalities.

In essence, here is the problem. Even if a model is true, this does not result in perfect data because observations contain error. Return to our first example of fitting a line to a series of points (see Fig. 7.4): Consider that the line is true (that is, that the model is true) and that the points were generated from this line, rather than the line being generated to fit the points. Because the points do not fall exactly
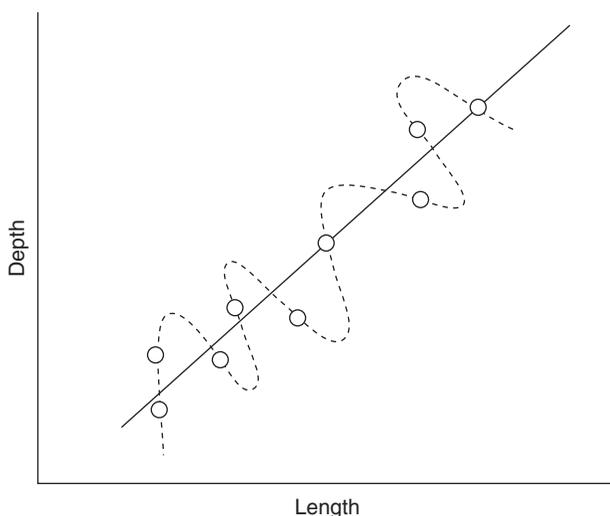
Figure 7.4. Graphic representation of two models that explain the relationship between length and depth of a hypothetical population with circles indicating actual measures. The straight line represents a fitted regression model that is most predictive of new data that might be collected. The dotted line is chasing error.

on the line, we can assume that the reason they do not is due to sample variation or measurement error. Now, we could fit a very complex line to these data, "explaining" all of the points (Fig. 7.4, dotted line). This line would contain a great number of parameters. But the extra parameters would actually be explaining the error, missing the truth. It would be "over parameterized." Akaike's criterion contains within its framework a mechanism for addressing this problem because it takes into consideration both the fit of the data and the simplicity of the model. The more complex the model, the more it must explain if it is to be the preferred model.

There is another thing we can learn from the example of slightly skewed data. Remember that we were using only 20 observations of antennae length. Recall that skewness does not help because the difference in likelihood found with and without considering skewness was not significant. But what if we had 10,000 measurements? It is quite possible that we would then need that extra parameter. This makes sense because our judgments are made relative to data, and not in a vacuum.

We can think of the model selection process as that area of parametric tree building where parsimony plays a part. Parameter-rich models can lead to the inability to discriminate between trees (Yang et al., 1995). The simplest model that can be implemented is preferred over more complex models.

Likelihood in Phylogenetics: An Intuitive Introduction

We introduced an intuitive sense of likelihood using a single correlation example (Fig. 7.1). Given what we have learned of parsimony and parsimony analysis, we
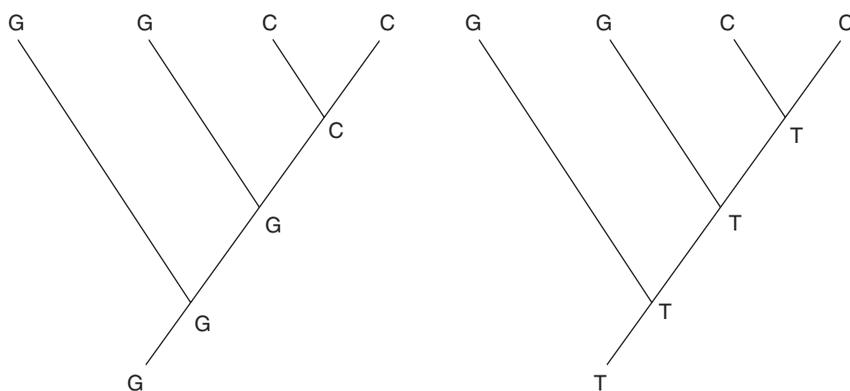
Figure 7.5. Two trees with identical topologies of bases assigned to the leafs but with different bases assigned to the nodes. Intuitively, the assignment of bases on the left tree has a higher probability than those on the right tree given either the parsimony or likelihood criterion.

can gain some intuitive sense of likelihood by considering the two trees in Fig. 7.5. These trees are identical in topology and observed data (shown at the tips) but are assigned different bases at the nodes. Note the distribution of the bases in the terminal taxa and the assignment of bases in the nodes. Which assignment of bases at the nodes do you think are more probable? Intuitively, we might agree that the assignment of bases on the left is more probable, given the data actually observed in the terminals. Parsimony gives us this same conclusion. The tree on the left requires but one change (G to C) while the one on the right requires four changes. As a parsimony problem, we would say that the character state assignments at nodes are more optimal in the tree on the left because the assignments on the left-hand tree require fewer steps. It is no wonder that we reach both conclusions: Sober (1983a, b, 1987) has argued that likelihood and parsimony converge to the same result if characters are considered one at a time under a simple model of character state change; Tuffley and Steel (1997) proved this result given certain conditions such as a Junks-Cantor-like model and different sets of branch lengths for each character.

However, we might ask another question: would the assignment of bases on the right tree be impossible? Answer: no, it is simply highly improbable given the topology and the distribution of bases on the terminals. If we were using parsimony, we might consider adenine to be allocated to the basal node, but we would never consider allocating adenine to the node leading to two terminal cytosines. We might concede that this is possible, but our character optimization algorithms do not allow for the result. Using likelihood, however, we can calculate a likelihood score for the distribution of bases on the right-hand tree as well as the bases on the left-hand tree, with a model of change and some math. In fact, we can calculate all of the various possible transformations that might occur no matter how improbable they may be if we have such a model. There would be 16 such possible base assignments for this particular problem of 4 terminals and 3 internal nodes, each specifying a particular scenario and each being relatively more or less likely, given the model. If

we added all of these calculations up, we could calculate a total likelihood for the entire tree over all possible transformations for this character. This is what a likelihood analysis accomplishes for any particular tree topology; it sums all of the likelihoods for all possible evolutionary scenarios of each site. It returns a value, that is, a sum of the log likelihoods for characters. The "best tree" is the one that maximizes this value.

As it turns out, we can do the same thing for parsimony. We could run all 16 permutations and evaluate the results in terms of tree length. The essential differences are that likelihood embodies a model of change that takes into account the length of the branch connecting two nodes and in likelihood we typically sum over all possible ancestral character states while in parsimony the score is based on the evolutionary pattern that yields the shortest length. Consider the C at the node leading to the two terminals (Fig. 7.5). The model specifies the probabilities of C remaining C or changing to A, T, or G over time or a surrogate of time, branch length. For relatively short branch lengths (definition: relatively short periods of time or relatively slow rates of mutation and fixation), the probability that a C in the terminal had an ancestral C at the node is relatively high, perhaps approaching 1. But over relatively long branch lengths, it becomes increasingly probable that the C at the node will have changed to some other base(s), which then changed back to C. This is under the evolutionary assumption that changes from C to some other base are governed by the model adopted.

The total likelihood for one tree can be compared to likelihood values obtained for other trees just as tree length values can be compared between parsimony trees. The tree with the highest likelihood score is preferred over alternative topologies because it fulfills the optimality criterion (ML) better than its competitors.

Likelihood in Phylogenetics: A More Formal Introduction

ML analyses in phylogenetics evaluate the likelihood of observing the data on a particular tree topology, given a particular model of evolution. If we consider phylogenetic analysis alone, we are searching for the best topology. Among the alternative topologies, that topology with a higher probability of giving rise to the observed data is preferred to topologies with lower probabilities, given the model adopted (Swofford et al., 1996). In certain respects, this is similar to how parsimony works. For instance, given an array of possible tree topologies, how well do the characters fit that topology relative to a specified criterion and the assumptions of the analysis? In parsimony, as mentioned previously, the criterion is minimum number of steps and the assumptions have been detailed in Chapter 6.

Likelihood is calculated for a given tree topology and a given set of branch lengths. Unlike parsimony, every possible transformation in each character is calculated over the topology and the branch lengths of that topology. Each topology has a potentially infinite number of associated branch lengths. An example, provided by Swofford et al. (1996) and redrawn in Fig. 7.6, illustrates the process. Note that while the most likely character state at node 5 is "C," the likelihood of all four bases at that node and all possible derivations of the C from the node below it also are calculated. Given a particular tree topology, the basic steps in performing an ML analysis are listed below.

Figure 7.6. The flow of likelihood calculations. (a) A data matrix of four taxa (1–4) and DNA bases (ATCG); we will use the data in column "j" for the example. (b) An unrooted tree with the characters from column "j." (c) A rooted tree with the same data; we will follow the process using only this tree, but a full account would require us to calculate likelihoods for other trees and compare them in some fashion. (d) Sample of the 16 various assignments that could be made to the nodes of tree (c); the sum of the likelihoods of all 16 constitute one of the elements of the next two figures (i.e., L_j would be the sum of the 16 possibilities for data column "j"). (e) The all "N" columns of data $[L_{(1)} \ldots L_{(N)}]$, the likelihood of the tree (c) is the sum of the likelihoods for all data columns. (f) The usually reported value is the log likelihood, which is the sum of the log likelihoods of each data column. Copyright Sinauer Associates, Inc., used with permission.

1. For each character of the data matrix, calculate the likelihood of character transformation at each node of a topology and its associated branch lengths, from the tips to the root. This is accomplished using the evolutionary model in concert with the observed character distributions to calculate the probabilities of observing character states at nodes.

2. Calculate the likelihood of the entire tree by multiplying the probabilities of each character together under the assumption that each character is evolving independently. That is, the likelihood of the entire tree is the product of the likelihoods of each character (each site) and this is usually expressed by summing the log of the likelihoods (because the product of likelihoods is very small).

3. Among the trees evaluated, pick that tree with the highest likelihood.

Note that the likelihood approach is a contrastive approach (Sober, 2008). Likelihood scores have no direct interpretation unlike the case in parsimony where steps have a direct interpretation. It is the difference between two likelihood scores that is important, not the scores themselves.

There are a variety of models that might be used to calculate the probabilities for character state change. For molecular data, these models are expressed in a matrix of change that may be overlain with assumptions about variation in rates over characters. Consider the following matrix of possible changes from one base to another:

	A	C	G	T
A	A→A	A→C	A→G	A→T
C	C→A	C→C	C→G	C→T
G	G→A	G→C	G→G	G→T
T	T→A	T→C	T→G	T→T

Consider → to specify some model of change from one base to the other (and the probability of not changing). We can scale the rates such that the branch lengths are in terms of the expected number of substitutions per site (column of data, character). The rate of "leaving" a state is reflected in the diagonal of the rate matrix, as in A→A or C→C. This rate is simply a negative number with a magnitude equal to the sum of the rates of changing to each of the other states, e.g. A→C or A→G. So, A→A is: –(A→C+A→G+A→T).

	A	C	G	T
A	–(A→C+A→G+A→T)	A→C	A→G	A→T
C	C→A	–(C→A+C→G+C→T)	C→G	C→T
G	G→A	G→C	–(G→A+G→C+G→T)	G→T
T	T→A	T→C	T→G	–(T→A+T→C+T→G)

Now consider one cell in the matrix: A→C. There are at least three components to the transition rate from A to C along any one branch of the tree: the relative frequency of base C in the model, the relative rate parameter for the A→C transformation, and the instantaneous substitution parameter. The second component, the relative rate, asks a question: how often does A change to C relative to a change of A to G or a change of T to C? We can write this relative rate into a matrix, discounting bases changing into themselves, as shown below:

	A	C	G	T
A	—	A	B	C
C	G	—	D	E
G	H	I	—	F
T	J	K	L	—

If we consider evolution to be time-reversible, as we usually do in parsimony calculations, then the relative rate of change of A to C would be the same as that from C to A, and we can simplify the matrix, because A = G, B = H, etc., as shown below:

	A	C	G	T
A	—	A	B	C
C	A	—	D	E
G	B	D	—	F
T	C	E	F	—

Let us consider this component of the substitution matrix. The values of the relative rate parameter depend on the assumptions of the evolutionary model. The simplest model is Jukes-Cantor (1969). This model assumes that there is an equal probability of any base changing into any other base. In short, the relative rates are equal. So the matrix is quite simple:

	A	C	G	T
A	—	A = 1	B = 1	C = 1
C	A = 1	—	D = 1	E = 1
G	B = 1	D = 1	—	F = 1
T	C = 1	E = 1	F = 1	—

Another relatively simple model, Kimura (1980), allows that there are differences between the rates between transitions and transversions. Thus, the model is a bit more complicated. For example, in the matrix shown below, κ determines how much more likely transversions are than transitions. If we set $\kappa = 2$, then the rate of a transition is twice the rate of a transversion.

	A	C	G	T
A	—	1	κ	1
C	1	—	1	κ
G	κ	1	—	1
T	1	κ	1	—

The next component, the instantaneous rate component (μ) is the probability that a change will occur at some very short time period along a branch. In time-reversible models, we also have to account for how common the base is assumed to be, because the actual rate of change is a function of the frequency of the base. Thus, for any one transformation, we have to account for three factors, not two:

$$A \rightarrow C = \mu a \Pi_C, \text{ given symmetrical change, or}$$

$$A \rightarrow A = -(\mu a \Pi_C + \mu b \Pi_G + \mu c \Pi_T) = -\mu(a \Pi_C + b \Pi_G + c \Pi_T)$$

We can sort all this out into a very general matrix, termed the Q-matrix. It can be described as a matrix that models the rate of change from one base to another during some very small amount of time. For a time-reversible model:

	A	C	G	T
A	$-\mu(a\Pi_C + b\Pi_G + c\Pi_T)$	$\mu a\Pi_C$	$\mu b\Pi_G$	$\mu c\Pi_T$
C	$\mu a\Pi_A$	$-\mu(a\Pi_A + d\Pi_G + e\Pi_T)$	$\mu d\Pi_G$	$\mu e\Pi_T$
G	$\mu b\Pi_A$	$\mu d\Pi_C$	$-\mu(b\Pi_A + d\Pi_C + f\Pi_T)$	$\mu f\Pi_T$
T	$\mu c\Pi_A$	$\mu e\Pi_C$	$\mu f\Pi_G$	$-\mu(c\Pi_A + e\Pi_C + f\Pi_G)$

Note that two of the factors are related as rate parameters while the other is a frequency parameter. For any small amount of time, the actual rate of substitution for A→C is a function of both the relative rate parameter and the instantaneous substitution rate, or μa, coupled with the frequency of the base Π_{base}. As it turns out, the Q-matrix is the product of two matrices (i.e., it can be decomposed into two matrixes), R and Π:

$$
R=
\begin{array}{cccc}
— & \mu a & \mu b & \mu c \\
\mu g & — & \mu d & \mu e \\
\mu h & \mu j & — & \mu f \\
\mu i & \mu k & \mu l & —
\end{array}
\qquad
\Pi=
\begin{array}{cccc}
\Pi_A & 0 & 0 & 0 \\
0 & \Pi_C & 0 & 0 \\
0 & 0 & \Pi_G & 0 \\
0 & 0 & 0 & \Pi_T
\end{array}
$$

Once we have a model, we can use it to calculate changes from one base to another along any particular branch of a tree over time by calculating a transition (or substitution) probability:

$$P(t) = e^{Qt}$$

The transition probability matrix can be evaluated for any particular branch length by decomposing the Q-matrix into it eigenvalues and eigenvectors. Swofford et al. (1996) demonstrate how this is accomplished for some of the simpler models where relatively simple expressions exist for the eigenvalues. For example, the Jukes-Cantor model only states that a particular base can change or not change, so there are only two probabilities to consider where i and j denote bases:

$$P_{ij}(t) = \tfrac{1}{4} + (\tfrac{3}{4})e^{-\mu t} \text{ if } i = j \text{ (no change)}$$
$$P_{ij}(t) = \tfrac{1}{4} - (\tfrac{1}{4})e^{-\mu t} \text{ if } i \neq j \text{ (change)}$$

We considered parameters earlier. Now that we know what a substitution model looks like, we can relate the concept of parameters to phylogenetic likelihood analysis. If a model has a parameter that receives a constant value over the entire model, then that parameter is effectively factored out. For example, if we assume the Jukes-Cantor model, then a = b = c = d = e, etc. Further, Jukes-Cantor assumes equal base frequencies: $\Pi_A = \Pi_B = \Pi_c = \Pi_D$, and this parameter factors out. Thus, we are left only with a single parameter, μ, and a single parameter model. You can see this in the equations above, it all boils down to a function of μ. Kimura (1980) introduces two values for the relative rate parameter, but assumes that base frequencies are equal; so here we have a two-parameter model. We can continue to add parameters, for example, the eight parameters of the general time reversible (GTR: Tavaré, 1986) model.

   Most likelihood analyses use a series of assumptions built into the analysis that simplify calculations.

1. Most models are time-reversible. This permits one to calculate likelihoods by rooting the tree at an arbitrary node and decreases computational effort. The result is an unrooted tree.

2. Characters (nucleotide sites, base positions) are assumed to evolve independently (Kluge's auxiliary principle). This permits likelihoods to be calculated separately for each character and then multiplied in order to calculate a likelihood for the entire tree.

3. Change is assumed to follow a Markov process. We assume that changes along different branches of the tree are independent of each other. In addition, the rate of particular change, e.g., A to T, does not depend on the history of change of the site prior to the acquisition of A.

4. Change is usually assumed to follow a homogeneous Markov process. Homogeneous Markov processes assume that the rate of change from one particular character state to another state (as specified by the evolutionary model) is the same over the entire tree.

Simple models assume equal base frequencies. However, the frequency parameter is often set empirically, by calculating the relative percentages of the four bases in the data matrix. The relative frequency of each base is assumed to be constant over the entire tree. The product of the relative rate parameter and the instantaneous substitution parameter yield a rate parameter that specified, in essence, base turnover rate along branches.

The probability of changing from one base to another is a function of the substitution rate and time. The mean substitution rate is set to one, and the relative rates are scaled so that they sum to the mean rate (Yang, 1994). The rate of evolution is usually allowed to vary over the tree. In this form of inference, each branch on the tree has a parameter that represents its length in terms of the expected number of changes per character. Forcing the rate to be fixed over the tree is adopting the molecular clock hypothesis, and likelihood can be calculated by considering times of divergence in rooted trees (Swofford et al., 1996, and citations).

Calculating the likelihood of a particular tree requires us to consider the likelihoods of the occurrence of each character state at each node given the states of the terminal taxa, the tree topology, and the estimated branch lengths. For the data at hand and a specified tree, likelihoods are calculated taking into account the prior probability of a particular character state (based, for example, on its overall frequency in the data) and the conditional probabilities of the character state remaining the same or changing along branches from the root to terminal taxa (observed sequences). Summing all this leads to a likelihood for the tree as a whole. To find the ML value for a tree, many combinations of parameter values must be evaluated until we find a set of values for all parameters (including branch lengths) that maximizes the likelihood.

The substitution probability matrix can be modified to accommodate rate heterogeneity by adding a rate factor based on some a priori or empirical assessment of variation among sites and rate of change. In fact, if there is significant rate variation among sites and this factor is left out of the model, then likelihood analyses will suffer from some of the faults typically ascribed to parsimony such as "long branch attraction" (Gaut and Lewis, 1995). A simple discrete model is the invariable-sites model (Hasegawa et al., 1985), where some fraction of the sites is considered invariant while the remaining fraction vary at the same rate. The gamma distribution

(Yang, 1993; Steel et al., 1993) provides a continuous model of rate variation among characters, although the model is usually implemented as an approximation that uses discrete rate categories (Yang, 1994). Typically we set the β parameter of the Gamma distribution $1/\alpha$, so that the mean rate across all characters is 1.0 and low values of α denote distributions of variation among sites such that most sites are evolving slowly while a few are evolving at a rapid rate. Higher α-values lead to less heterogeneity in rate, and if α is infinite, then all sites are evolving at exactly the same rate. The value of α can be inferred using ML, as with the other parameters of the model.

Selecting Models

Likelihood methods present the investigator with a plethora of models. The question is: what model should one use? One problem in determining this is that different models contain different numbers of parameters, and as the number of parameters increases, so does the variance associated with the ML values of a given set of trees. We can fall into the trap of tracking noise rather than signal. All parameters except the one the investigator is interested in are termed "nuisance parameters." If the parameter you are seeking is the tree topology, then all other parameters such as branch lengths or rates of change in particular classes of data are "nuisance parameters." However, as the investigator introduces more and more parameters, the possibility arises that the new parameters are contributing little to the result, which is the tree topology. Consider our original example of estimating the mean. We added a parameter of skewness, but our data were hardly skewed and adding the parameter did not significantly contribute to our estimate of the mean of the population. So parameters should be added only when they actually help and the process of adding parameters can reach a point in model complexity where one cannot discriminate between any of the trees due to the large variances associated with each parameter. Thus a model is said to be "over-parameterized." Over parameterization is a well-known statistical problem in such operations as multiple regression and discriminant analysis.

   One may wonder how an investigator might pick a model in the absence of a known phylogeny. Swofford et al. (1996) suggest that one selects a goodness-of-fit statistic and then selects a model that maximizes this statistic without adding additional (perhaps unnecessary) parameters. Many such tests or criteria can be used to inform the investigator of the fact that adding an additional parameter does not significantly improve the fit. The first is the log likelihood ratio test (the G-test of Sokal and Roth, 1981), and the second is the Akaike information criterion (Akaike, 1973; Sober, 2008).

   The log likelihood ratio test is simply the likelihood ratio test of goodness of fit. It tests the significance between two nested hypotheses (that is, one hypothesis is a proper subset of the other, as in our example of mean and standard deviation versus mean, standard deviation, and skewness). In phylogenetics, for example, we can select a particular tree topology and then compare models with different numbers of parameters. In application, one sees if adding an additional parameter to a model produces a significant increase in goodness-of-fit of the data *given the same tree topology*. For example, consider a model of equal rates of evolution across all sites

and another that allows rates to vary. The models are nested and can be tested against a common topology. Does the more complicated model (varying rates) produce a significantly better fit of the data than the model of equal rates across all sites? If so, then one accepts the more complicated model.

You might wonder how the topology is picked. Sullivan et al. (1997) showed that using a topology that is a rough estimate of the actual phylogeny will result in a reasonable model for further analysis. A good rough estimate is a parsimony tree (Yang et al. 1995; Sullivan and Swofford, 2001). However, a random topology or one very dissimilar to the actual phylogeny can produce suboptimal models (Sullivan et al., 1997).

The Akaike criterion introduced a penalty for each added parameter based on the degrees of freedom associated with the parameters (Akaike, 1973). Its principal strength is that it can be used to test across tree topologies, rather than being restricted to testing within nested hypotheses (Prosada and Buckley, 2004; Sullivan and Joyce, 2005).

BAYESIAN ANALYSIS

Adoption of likelihood approaches to phylogenetic systematics lead to consideration of Bayesian approaches. Likelihood asks "what is the probability of the data given the model [p(data|model)]? The alternative question is: what is the probability that the model is correct given the data [p(model|data)]? This has to be evaluated with respect to a set of candidate models. The relationships between these two questions are bridged by Bayes' theorem, a theorem that details the relationship between conditional probabilities. Again, Read (2000) provides a simple and elegant explanation of this relationship.

Consider the probability of A being true, p(A). Now, consider the proposition, p(A|B) as the probability that A is true given that B is true. In other words, what we can say about A is related to what we know about B. There is an overall probability of A being true, and there is a conditional probability that A is true given what we know about B. For example, consider the "probability space" of A and B to be represented by the Venn Diagram in Fig. 7.7.

The area occupied by A is the p(A), and the area occupied by B is the p(B). Note that the two areas overlap: exactly 50 percent of the area of A is included within the area of B and 40 percent of the area of B is included within A. That overlap is the area where both A and B are true. That area of A that does not overlap the area covered by B is the area where A will be true even if B is false. The joint probability can be derived from the multiplication of probabilities:

$$p(A \text{ and } B) = p(A)p(A \text{ given } B)$$

In the Venn diagram, A occupies 20 percent of the probability space, so: p(A) = 0.2. This is the probability of A. Likewise, the probability space occupied by B is 25 percent, so: p(B) = 0.25. This is the probability of B. The region covered by both is p(A and B) = 0.1. Now, if we assume that A is true, then B will also be true half of the time:
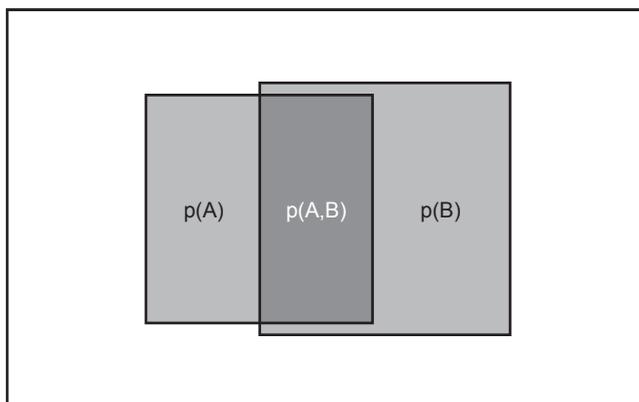
Figure 7.7. The probability of A + B is the joint probabilities of A and B. Redrawn from Read (2000), used with permission.

$$p(B \text{ given } A) = 0.1/0.2 = 0.5 = p(B \text{ and } A)/p(A) = 0.5$$

This is the probability of B being true given A. If we assume that B is true, then 40 percent of the time A will also be true:

$$P(A \text{ given } B) = 0.1/0.25 = 0.4 = p(A \text{ and } B)/p(B)$$

This is the probability of A given B. Now, we can verify the multiplication law of probabilities:

$$p(A \text{ and } B) = p(A)p(B \text{ given } A) = (0.2)(0.5) = 0.1$$
$$p(A \text{ and } B) = p(B)p(A \text{ given } B) = (0.25)(0.4) = 0.1$$

Thus:

$$p(A \text{ and } B) = p(a)p(B \text{ given } A) = p(B)p(A \text{ given } B)$$

and

$$p(B \text{ given } A) = p(B)p(A \text{ given } B)/p(A)$$

Now, consider one of these "B" to represent a model and "A" to represent data. We can cast the formula in likelihood terms:

$$p(\text{model given the data}) = p(\text{model})p(\text{data given the model})/p(\text{data})$$

and in more familiar notation

$$p(\text{model} \mid \text{data}) = p(\text{model})p(\text{data} \mid \text{model})/p(\text{data})$$

The denominator, p(data), is actually complex and is the sum of the probability of the hypothesis multiplied by the probability of the data, given the hypothesis,

$$P(model \mid data) = P(model)p(data \mid model) / \Sigma_{model}\ p(model)p(data \mid model)$$

The denominator performs a desirable function: it normalizes the posterior probabilities so that they add to one. The problem is: the denominator is the sum of all models, and for a phylogenetic hypothesis that is only moderately large there are a great number of models—too large to be calculated. Thus, it is not just the tree topologies that can reach astronomical numbers but also the branch lengths. A complete account of all of the models for a particular tree topology would be all the possible branch lengths for that topology, which are a function of all of the rate parameters. Multiply that by the number of possible trees, and you can see why we need to approximate, rather than calculate, the answer.

Now, let us consider our original very simple example from the Bayesian perspective. Given our data, we can see intuitively that if the ML is reached with a model of mean = 5.0 mm and standard deviation = 1.0, the probability of a model that specified mean 4.9 mm, standard deviation 1.1, given the data, would be higher than the probability of a model with the mean = 6.0 mm and the standard deviation = 2.0. The question is: if we do not know the mean and only have the measurements, how would we reach the solution in a Bayesian manner? Consider the landscape showing our contours, Fig. 7.8. In a Bayesian analysis, the shapes and contours of this landscape are not known; we have to discover them.

If we performed all of the likelihood calculations at, say intervals of 0.1 mm × 0.1 mm, we would have one term in the equation. We also need a prior, the p(model), to complete the equation. This requirement for a prior partly leads to controversy because frequentists might claim that there is no justification to introduce prior expectations into the calculation. Classical phylogeneticists might also object to the statistical nature of this kind of inference. Still, even in classical phylogenetics, concepts akin to priors exist such as designated outgroups used to polarize characters.

If we were sampling the population a second time, the prior for the mean might be centered around 5.0 and the prior for the standard deviation might be a distribution centered around 1.0. If we have never sampled the population, perhaps we don't expect anything in particular. We might consider each fitted model to be equally probable. Since we have dissected the landscape into little squares, we can calculate the probability of each square and the entire lot of these calculations will sum to p =1.0. This creates a hill whose volume is 1 (Fig. 7.8). If we project the volume under each square, the sum of these volumes = 1, so the volume under each square is a measure of p(model|data) and the little square with the most volume happens to be the one at the top of the cone which includes mean = 5 mm, standard deviation = 1. The shape of this hill can tell us a great deal about the model as a whole and cross-sections can tell us a great deal about the model parameters.

The shape of a hill is a function of the credible intervals we might expect for any parameter. If there is only one peak (as in our case, cross-section Fig. 7.8a), then we can conclude that the problem is rather simple. If the surface has multiple peaks, then we can conclude that the fit of different models to the data is complex. If the cone forms a steep peak (cross-section Fig. 7.8b), then the difference in volume
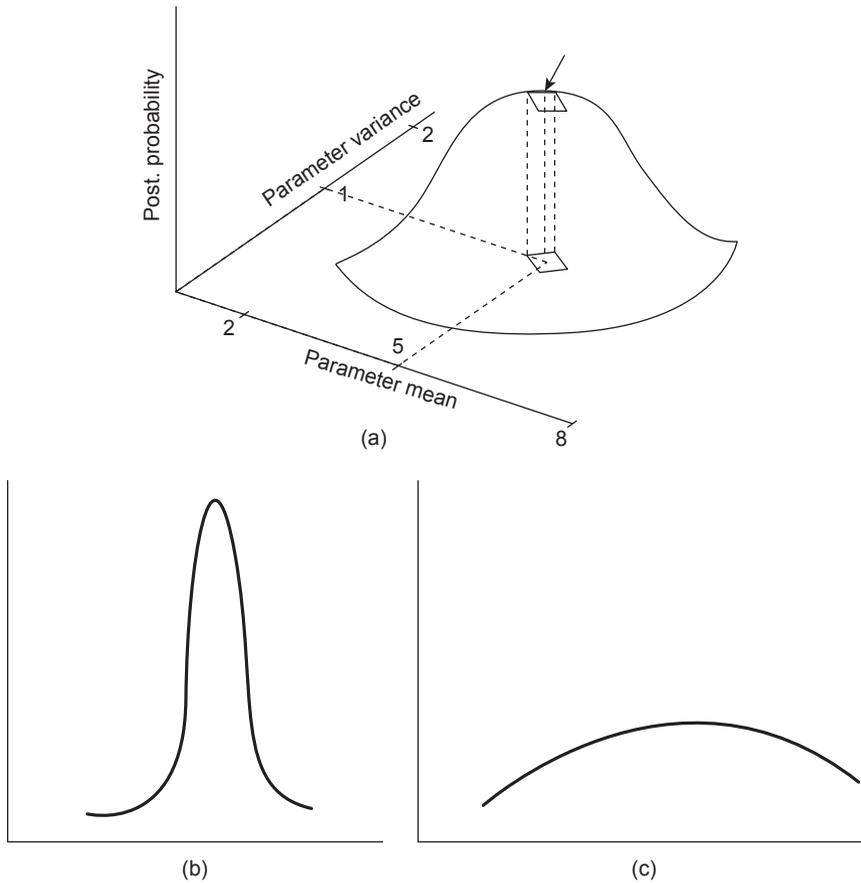
Figure 7.8. (a) Posterior probability space for our example of antenna length from Fig. 7.2. (b–c) Cross-sections of two probability spaces. In (b) there are very few fitted models at the top of the hill that have similar likelihood values. In (c) there are many fitted models at the top of the hill that have similar likelihood values.

among adjacent squares is large and we will have relatively few fitted models on the peak that are statistically equally probable. If the cone is in the form of a low hill (cross-section Fig. 7.8c), then the difference in volume among adjacent squares is very small and we have many models that might be equally probable. This result is very intuitive if we apply frequentist statistical reasoning. We know that confidence limits in frequentist statistics are a function of variation in the population. We may estimate a sample mean, and if the variance is low, we can expect to sample something similar the next time we sample the population for the same number of individuals. However, if variance is very high, then we might expect very different results the next time we sample, and our confidence would be low.

We can also isolate parts of the model, concentrating on only one model parameter and treating the other parts as nuisance parameters. For example, if we were only interested in the mean and its credible interval, we could treat the standard deviation as a nuisance parameter, and we could sum the probability along the axis

that corresponds to the standard deviation. This would give us a one-dimensional curve showing the posterior probability density as a function of the mean. We can examine this curve to learn about what values for the mean are most probable in terms of posterior probability. We can do the same thing for the standard deviation.

In tree inference, we are interested in the tree topology that has the highest posterior probability. Any particular tree topology is associated with many other parameters, and it is theoretically possible to examine the posterior probabilities of each tree topology in a manner similar to examining all posterior probabilities of fitted mean values in our simple example. Some trees will be more probable, given the model, than others. And we can examine the posterior probabilities of other parameters, one at a time, given a particular tree topology, the other parameters, and the model, which is something of interest to evolutionary biologists interested in studying various evolutionary processes.

Now, we mentioned that summing all of the values in the denominator of Bayes' theorem was hard. In fact, it is an intractable problem in phylogenetic inference, due to the numbers of trees, branch lengths, and other parameters mentioned above. So, we cannot visualize the posterior probability surface or even visualize the complexity of the landscape to see if there are multiple optima. So, what to do?

As it turns out, this is not an insurmountable problem in phylogenetic inference because we can approximate using a variety of heuristic methods. The favored method used in Bayesian phylogenetics is to explore the possible space occupied by the models. If enough space can be visited, then perhaps we can find the hill and its peak without visualizing the entire landscape. This is somewhat analogous to exploring parsimony space using such routines as branch swapping. In parsimony we explore a landscape of tree lengths, and in Bayesian analyses we explore a landscape of posterior probabilities. In a parsimony analysis, we cannot see the entire landscape. We simply try to find the highest peak(s), and when we do, we accept the results as our best estimate. This is analogous to Bayesian analysis where we cannot see the entire landscape, but try to find the highest peak. Exploring model space is somewhat similar, except for the fact that we are adopting different criteria (tree length versus posterior probability).

There are several ways to explore the space. For example, Rannala and Yang (1996) used numerical integration of the posterior probabilities over all tree interior nodes, an approach that could be viewed as similar to an exhaustive search in parsimony, and one that suffers from the same defect: it is only useful for problems involving small numbers of taxa. The favored method is Markov Chain Monte Carlo (MCMC) integration.

Monte Carlo methods were first proposed by the mathematician Stanislaw Ulam (1909–1986: Weisstein, 1998). A simple description of how MCMC works is given by Lewis (2001a). In general, MCMC involves using computer-generated random numbers and a set of rules to simulate a walk through the space of trees and parameters. One begins by either randomly picking a model (random tree topology and other associated parameters) or by picking a particular model (one considered a priori probable, usually a particular tree topology and associated parameters). One then randomly picks a second model and compares it to the first. If the proposed model has a higher posterior probability density, then adopt it and pick another random likelihood model to test. But if the proposed model has a lower

posterior, then it can still be picked with some probability (the probability is simply the ratio of the posterior for the proposed state to that of the current state). For example, if the second model is actually better than the first, then we pick the second 100 percent of the time. However, if the posterior probability density of the second tree is 75 percent of the current tree's posterior probability density, then pick the second tree 75 percent of the time by random draw, otherwise we stick with the original tree, pick a new tree model, and make a new comparison.

These methods are described by Metropolis et al. (1953) and Hastings (1970). If MCMC procedures are followed correctly, we will travel over the landscape and (given a long enough simulation) the amount of time the simulation spends in a particular set of models will approximate the posterior probability of that set. If we save the results as a function of the frequency with which we sample various models, the procedure can sample the shape of the hill and allow us to find the approximate location of the peak. This procedure is sampling the *posterior probability density* of the models given the data. It works rather like an n-dimensional histogram, the more probable models are visited more often than the less probable models, building a probability density space where the best model(s) are interpreted as more probable because MCMC visits them more often ("finds them") than the less probable models.

MCMC searches run forever unless stopped by the investigator. In phylogenetics, the end result we seek is a stable topology or set of topologies, so all other parameters (e.g., those determining branch lengths) are treated as nuisance parameters. (Of course, if the objective is not strictly a systematic objective, then other parameters may be of interest.) If we run MCMC long enough, we will be able to accurately estimate the posterior probability of any tree. If the peak is steep and pointed, then the area of the peak is very small. If we could visualize its volume relative to the entire landscape, the volume would be highest. In such a situation, we would expect the peak to be populated by a relatively small set of trees and associated parameters (Fig. 7.8b). If the peak is low and flat, it will be populated by a large number of tree topologies and associated parameters (Fig. 7.8c).

Consider a simple thought experiment. Pretend that you have traveled around and over the surface for a million iterations. That means you have visited 1×10^6 models, with replacement. One model may be visited many times, another only a few times. If all of the states you have visited corresponded to the same tree topology, then 100 percent of the time you would have encountered the same clades. Consider another case in which you visit 15 tree topologies. In this case, it is possible that no single tree dominates the posterior probability surface, but you can still summarize the proportion of steps in which you were in a state that corresponds to each tree. This will allow you to estimate the posterior probability of each tree. You can also calculate the proportion of steps in which you visited a tree that had a particular clade. This proportion is interpreted as the posterior probability that this clade is present in the true tree.

The usual procedure is as follows.

1. Select a model that contains a reasonable set of parameters for change in the same manner as selecting a likelihood model for ML. In the best of all possible worlds, this would be the true model of character evolution, but in the real world, one hopes that the analysis is robust to the use of a simpler model.

2. Run an MCMC analysis. This consists of one or more Markov chains (four is common), each sampling the probability landscape. The point is to sample the posterior probability space as thoroughly as possible and find the area(s) of highest probability density. There may be multiple peaks of different heights. So the typical analysis may consist of multiple MCMC chains and criteria for switching from one chain to another, if it looks more promising.

3. The first MCMC results are likely to be uninteresting because the analysis is just beginning to sample the probability space and is heavily influenced by the starting point for the MCMC (and this starting point is often arbitrary). Frequently the initial portion of the MCMC run is discarded as a "burn-in." Convergence of the chain to accurate approximation of its stationary distribution is difficult to detect. Multiple independent runs are usually required. Discrepancies between different simulations with respect to the estimates of the posterior probability surface indicate that the simulations (or at least some of the runs) were terminated prematurely.

4. If tree topology is the parameter and all other parameters are considered nuisance parameters, the goal is to determine the relative frequency of the appearance of particular clades in the probability space sampled. *The frequency at which you sample any clade provides an estimate of its posterior probability.* It is possible to collect all of the samples, but this would take a great deal of computer space because we would have to save upward of 1 to 3 million tree topologies, one for each visit to the peak by MCMC. More typically, the investigator saves some subsample, typically one tree per thousand sampled. Consensus analysis can then be performed on the collection of trees. The usual method is to take a majority rule consensus tree, which will reveal the percentage of times a particular clade appears in the subsample and this is interpreted as the posterior probability of the monophyly of the clade given the data.

5. To understand the results, it is helpful to consider what might be on the peak. Consider first a very simple problem: a set of data having little homoplasy. If we ran a parsimony analysis, we would quickly find a solution and, if the data are really clean, perhaps only a single tree that is optimal. In this case, we would also expect the same result using likelihood methods, or classic Hennigian argumentation with a priori character polarization. In such a case, we can imagine that the peak is very pointed and at the top are a rather large number of models that do not vary in tree topology, but might be different in the nuisance parameters of branch length, a function of the Q-matrix. (Indeed, there are, potentially, an infinite number of such models because branch length variation is continuous.) In this case, you would expect the algorithm to sample the same clades, over and over again, discovering alternative topologies very infrequently. Now consider a very messy problem: a great number of taxa and many potential homoplasies. A parsimony analysis might result in many parsimonious trees (or many trees that differ only slightly in length), perhaps collapsing into one big polytomy if a consensus analysis was performed. Our confidence in clades that appear only in some subset of the most parsimonious trees is not very high; bootstrap values are low for most clades. In this case, you would expect the peak to be broad and populated with many topologies

plus the complement of diversity in nuisance parameters. You would expect the algorithm to sample any particular clade at a low frequency or perhaps not at all.

There are several computer packages available for performing Bayesian phylogenetics. The most commonly used package seems to be Mr. Bayes (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003). An excellent introduction to Bayesian literature is presented on the Mr. Bayes WWW site: mrbayes.csit.fsu.edu.

INTERPRETING MODELS IN A PHYLOGENETIC CONTEXT

It is important to understand that likelihood methods may be used as tools in both systematic research and in broader evolutionary research. Strictly speaking to systematic research, parameters such as branch length are of concern only when branch lengths interfere with reconstructing common ancestry relationships. This can be a concern when taxon sampling or character evolution results in "long branch attraction" problems (see, for example, Anderson and Swofford, 2004). Otherwise, phylogenetic systematics, as a discipline, takes no account of branch lengths because branch lengths that are not artifacts of "long branch attraction" are not taken into account in determining genealogical relationships or in forming phylogenetic classifications. Trying to account for such things in classifications is a research program for evolutionary taxonomists, not phylogeneticists. If, however, the thrust of the research is in other evolutionary directions, then we may wish to have a tree before the fact and study the evolutionary behavior of characters on that tree. Such research programs are not primarily systematic, but depend on the fruits of systematic analysis.

There are two potential advantages of using explicit models in a likelihood-based framework. First, likelihood-based techniques are parametric techniques, and parametric techniques are usually more powerful than nonparametric techniques. *Power*, in this sense means statistical power, the ability of the statistical test to correctly reject a false null hypothesis. Second, if the analysis uses the correct model of evolution, then ML will have the characteristic of statistical consistency. *Statistical consistency* obtains when the analysis converges on the correct solution as more and more data of the same kind are applied to the problem. Early proponents of likelihood touted statistical consistency as a clear advantage of likelihood over parsimony (e.g., Felsenstein, 1978). Parsimony adherents pointed out that there was no guarantee that ML methods would produce statistically consistent results unless the model was true, and that because no one claimed that their models were true, the claim is unjustified (Farris, 1999). ML adherents replied that ML was robust to deviations in model truth-value, which seems to be reasonable under certain conditions, but not in others, especially when evolutionary rates are heterogeneous across or within sites and when these parameters are not included in the model (e.g., Gaut and Lewis, 1995). Sober (2008) states that scientists use models they know to be "untrue" all of the time and that models should be judged on their power to predict new data such that fitted models can be tested. Sober (2008:351–352) suggests the controversy between parsimony and likelihood is not likely to be solved by simply process models where parsimony and likelihood agree (e.g., Felsenstein, 1981a; Penny et al.,

1994; Tuffey and Steel, 1997) or disagree because parsimony advocates will claim that the models are wrong or unknowable and likelihood advocates will claim that the models are at least good enough to make accurate predictions.

Given this acrimony, we were curious to see exactly what might emerge when we performed both a parsimony analysis and a Bayesian analysis on a large morphological data set and then used the ability of Mesquite (Maddison and Maddison, 2009) to plot the distribution and probabilities of likelihoods of synapomorphies on the tree topology generated from the Bayesian analysis and compare it to the results of the parsimony analysis. We could not find argumentation. The question is: does only parsimony, as an algorithm, fulfill the basic principle of Hennig (1966) that one should not assume convergence in the absence of evidence? An additional question: does a likelihood approach result in monophyletic groups confirmed by synapomorphies? The auxiliary principle and grouping by synapomorphy are thought by us to underlie all phylogenetics. But what if parsimony and classical Hennigian argumentation are not the only way to fulfill the paradigm? Or to put it another way, is there more than one way to be a phylogeneticist in the Hennigian tradition?

With the help of Matthew Davis (then a graduate student at the University of Kansas), we analyzed the Gauthier et al. (1988) matrix of amniote relationships (fossil and living) using both parsimony (PAUP*, TBR, 100RAS, equal weighting) and a Bayesian analysis with the simple Mk morphology model of Lewis (2001b). We did not perform a likelihood analysis as there is no way at present to examine synapomorphies at nodes using likelihood at this stage in the development of algorithms. Ancestral states reconstruction, as implemented in Mesquite (Maddison and Maddison, 2009), was used to study the distributions of characters assigned to ancestral nodes. Because of missing data, we did not study the state reconstructions of the soft anatomical characters, but they were included in the analyses. The single tree topologies of both analyses were identical. Of the 207 hard anatomical characters studied, 198 unambiguous synapomorphies were mapped to the same node in both the parsimony and Bayesian tree topologies. The only difference was that the Bayesian probabilities were on the order of 95 percent to 99 percent while those of parsimony scored CI values of 1.0. Characters that were ambiguous in the parsimony analysis were ambiguous in the Bayesian analysis (9 of 9) and at the same nodes, and there were no character columns that were totally different in their interpretation (although 7 instances of states were different, e.g., probabilities of 75 percent in Bayesian versus CIs of 1.0 in parsimony). Of course, this is only one analysis and one Bayesian model, but these results seems to show that at least for one matrix the Bayesian analysis is attempting to maximize homology and minimize homoplasy, and it is circumscribing monophyletic group with synapomorphies—two of the goals of the Hennigian Paradigm. We conclude that parsimony may not be the only way to achieve classical Hennigian objectives.

CHAPTER SUMMARY

- Likelihood asks the question: what is the probability of observing these data given a specified model (which includes a tree topology)?
- In phylogenetic inference, the model includes a tree topology and some number of other parameters including branch length between nodes.

- Likelihood models take into account the probability that characters will change over time between nodes.
- The simplest model is picked from a variety of models available.
- Bayesian analysis asks the question: what is the probability of the model (including the tree topology) given the data?
- Bayesian analysis also includes both the tree and other parameters such as branch length.
- Bayesian analysis uses the formalities of likelihood calculations and explores posterior probability space using MCMC to find the model/tree topology with the highest posterior probability.
- For phylogenetic research, branch lengths are of concern only in so far as they affect our ability to estimate genealogical relationships, so the important part of the model in both kinds of analyses is the tree topology.

8

PHYLOGENETIC CLASSIFICATION

> Academic taxonomy deals with classes; it merely arranges according to similarities; while natural taxonomy arranges according to kinships determined by generation.
> —Kant, 1775; quoted in Dobzhansky (1962:93)

> [A]ll true classification is genealogical.
>
> —Charles Darwin (1859:420)

Classifications are systems of names organized to show relationships among the entities named. The names derive their meaning from the intent of the persons who are trying to communicate. Biological classifications are used to convey ideas of the relationships among organisms. Technically, biological classifications are not classifications at all because taxa are not classes. Thus, Griffiths (1974) prefers *systematization* and de Queiroz (1988) has suggested that confusion over classification as opposed to systematization inhibited the spread of phylogenetic classifications between the time of Darwin and Hennig. However correct *systematization* might be, the term has not caught on, and we may be stuck with an inappropriate term (*classification*) based on common usage. Still, we have no problem with the term *phylogenetic systematization*, although we will refer to it as *phylogenetic classification*.

Phylogenetic classifications are biological classifications that meet the minimum criteria of being a system of names that imply relationships that are logically consistent with the phylogenetic tree the classification references. Differences between phylogenetic classifications of the same organisms may come from two sources. First, phylogenetic classifications may differ because they adopt different conventions for showing relationships, e.g., a classification that names each branch as compared to

a classification that uses a listing convention, as described below. Or they may differ because one uses Linnean conventions and another uses numerical prefixes or Phylocode conventions. Second, they may differ because the reference phylogeny is different. In the first case, we can think of the classifications as different ways of communicating about the same idea. In the second case, there are disagreements about the underlying relationships. Differences of the second kind are biologically important as they denote differences in the empirical data or the interpretation of the empirical data. Differences of the first kind are a matter of conventions adopted.

Classifications that include what we now recognize as monophyletic groups have been around long before Linnaeus, never mind Darwin. This indicates that the pattern of evolutionary descent in some groups is clear enough that their existence was recognized both by those who believed that natural order was divinely composed and by those who recognized that the order is the result of natural processes (e.g., Louis Agassiz and Charles Darwin, respectively). The distinctive nature of groups such as Vertebrata or Aves makes them stand out as natural individuals regardless of the methods of taxonomists (Patterson, 1977). However, it was left to Hennig (e.g., 1966) to codify the critical distinction between monophyletic and paraphyletic groups and thus clearly distinguish between artificial groups (those poly- and paraphyletic) and natural groups (those monophyletic).

In this chapter, we will discuss the general nature of classification, including some nonbiological examples. We will then discuss ways of making phylogenetic classifications and discuss the annotated Linnean Hierarchy (Wiley, 1979c, 1981a). The merits of this system are discussed and examples given. We will then discuss alternate ways of classifying phylogenetically, including such issues of classifications without rank, numerical prefix schemes, and the PhyloCode. We will end with our rationale for preferring phylogenetic classifications over classifications that include paraphyletic groups.

CLASSIFICATIONS: SOME GENERAL TYPES

The process of classifying is the activity of grouping entities or phenomena and giving names to the resulting groups. The placing of some things into one group to the exclusion of other things implies that the members or parts of the group share some relationships not shared with things outside the group. There are many ways to parse out the classification of classifications. One could recognize a basic dichotomy between hierarchical and nonhierarchical classifications. Or one could divide classifications into natural classifications and artificial classifications. For purposes of discussion, we will distinguish between three types of classifications: (1) those involving natural kinds, (2) those involving historical groups and individuals, and (3) artificial, convenience classifications. This discussion builds on the discussions in Chapter 3 on the nature of supraspecific taxa and in Chapter 5 where we discussed the relationship between characters and groups.

Classification of Natural Kinds

Natural kinds are formed when the entities classified have an indirect a-historical relationship. Entities are members of a kind by virtue of having the property of the

kind and further that property is a property that functions significantly in some theory of the world thought to be valid. The most obvious example is the semihierarchical classification of the Periodic Table. Each kind (Hydrogen, Helium, Lithium, etc.) has properties that are necessary and sufficient for kind membership such that entities can be placed in the kind. In this case, the property is atomic number. Certain kinds can be hierarchically grouped based on consequences of the property. For example, the group of inert kinds (Helium, Neon, etc.) have properties deriving from their atomic number. Entities that are members of these kinds have orbitals that are filled and thus do not participate in chemical reactions under normal, Earthlike, conditions.

Astronomers classify stars based on the size, luminosity, and temperature into such groupings as main sequence, blue giants, and red dwarfs. The pattern on which this classification is derived is the Hertzsprung-Russell (H-R) diagram shown in introductory astronomy classes (Fig. 8.1a). This classification is used for a variety of purposes as a predictive tool. The place of a star on the H-R diagram can be used to predict the age and future ontogeny of the star, within certain limits. Blue giants are all relatively young stars (100 million years old or less) because they are so massive that they burn fuel quickly. Red dwarfs have a life expectancy of many billions of years, because they burn their fuel slowly. The shape of the H-R diagram, when applied locally, has predictive properties. For example, the H-R diagram of stars in an old globular cluster looks much different from Fig. 8.1b due to an excess of giants and a low number of stars on the main sequence. The age of the cluster can be inferred by where along the main sequence the pattern is disrupted.

Classifications of natural kinds are rarely hierarchical and even the Periodic Table is only partly hierarchical, extending only one to a few levels. What all such classifications share in common is that the properties are not shared historically. Instead, they are properties gained by what an evolutionary biologist would term convergence. It is the power of such classifications that they organize and explain convergence in reference to laws and processes thought to be valid.

Historical Classifications (Systematizations)

Historical classifications are based upon inferred historical connections between the entities classified. The properties, as mentioned in Chapter 5, have an indirect historical relationship, but the entities have a direct historical relationship, that is, they form ancestor–descendant relationships. Historical classifications form deep hierarchies because the entities are replicators or they are made up of replicators; thus the entities formed have part–whole relationships. The replication need not be biological. For example, our present continents are "descended" from Pangea; North America is a historical part of Pangea, as is Africa. Phylogenetic classifications of organisms are a type of historical classification. *Homo* and *Rana* are parts of Tetrapoda, and Tetrapoda is part of Vertebrata. *Homo sapiens* is a replicator, whereas Tetrapoda is not.

The properties of entities in historical classifications are time-bound, a quality not found in the properties of natural kinds. The class of properties that systematicists can study are what we term *homologies* (Chapter 5). Each grouping based on one to many synapomorphies is a hypothesis that the entities comprise a monophyletic group. A particular monophyletic group (Aves) is an entity (or more formally,
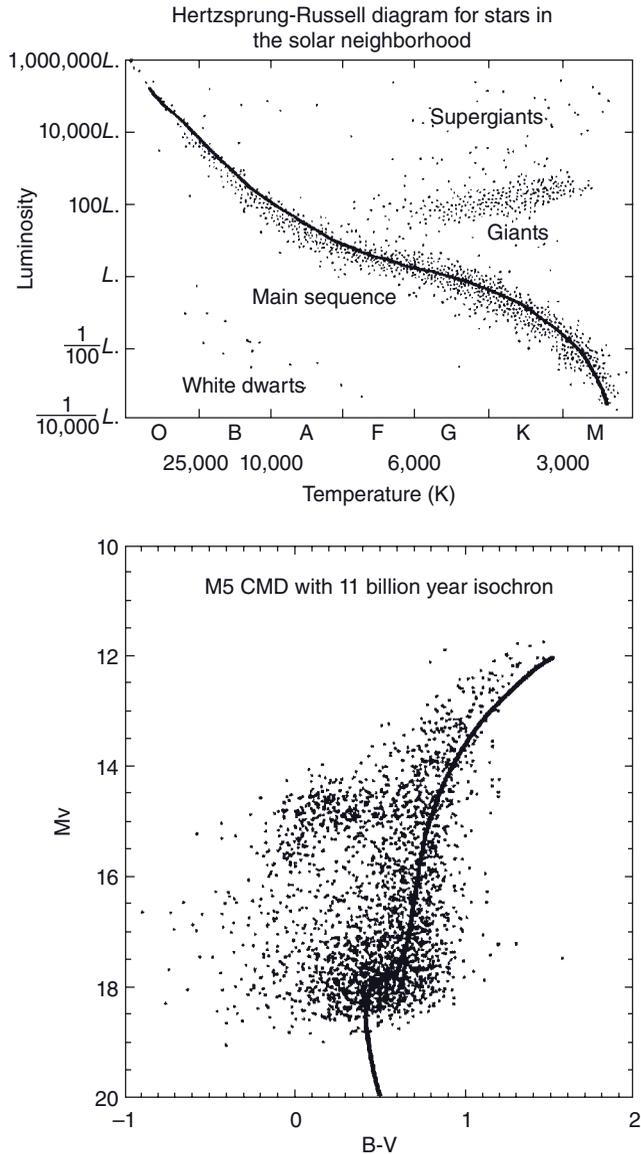
Figure 8.1. Examples of nonhierarchical classification, using Hertzsprung-Russell Diagrams to identify the nature of star clusters. (a) An H-R diagram of stars in the Pleiades open cluster (M45). (b) The H-R diagram of the globular cluster (or dwarf galaxy) 47 Tucanae. The age and classification of different clusters (open and globular) are determined by the pattern of distribution of their constituent stellar population. M45 is a relatively young open cluster with stellar distributions similar to a random selection of galactic stars and thus follows the "Main Sequence." The 47 Tucanae has an old stellar population that is atypical of the galactic population as a whole with many stars off the "main sequence." B-V is the difference between apparent blue (B) and visual (V) magnitude and cooler red stars are to the right of each graph. The Mv is the absolute visual magnitude of stars based on their apparent magnitude and the distance modulus of the cluster. Use of original data from the VizieR catalog service, Centre de Donnée astronomiques de Strasbourg, is gratefully acknowledged.

hypothesized to be an entity). It is important to recognize that although individual monophyletic groups are not natural kinds, the monophyletic group, in general, is a natural kind within evolutionary theory. Moreover, discovery of individual cases of monophyletic groups is confirmation that an important part of evolutionary theory (speciation) is valid.

Convenience Classifications

Convenience classifications are similar to classification of natural kinds in that a hierarchy, if it exists, does not extend to all entities covered by the classification. For example, the kind "felon" might include the kinds "murderer" and "arsonist" but not the kinds "speeder" or "adulterer." Convenience groups have properties that are indirectly a-historical, and their justification does not depend on invoking natural processes. There are many useful convenience classifications, including such classifications as the Dewey Decimal and Library of Congress systems of classifying books.

BIOLOGICAL CLASSIFICATIONS

Biological classifications may fall into any of the three kinds listed above and may be hierarchical or nonhierarchical, and they may group entities into kinds or historical groups (or even convenience groups). An example of a largely nonhierarchical classification of kinds would be energy-flow classifications with kinds such as primary producers that might comprise both photosynthetic and chemosynthetic organisms and with primary and secondary consumers. These kinds relate directly to process theories about movement of energy through ecosystems. We do not expect the entities that are members of each kind to form monophyletic groups that are associated with synapomorphies; both pitcher plants and lions are secondary consumers, and these character properties did not arise via common descent of an ancestor that was a secondary consumer. As mentioned above, evolutionary theory also has its natural kinds. For example, monophyletic group or species, but also Mendelian population. However, systematists are usually interested in fully nested historical classifications, specifically classifications of the natural hierarchy that comprises the tree of life.

Constituents and Grouping in Phylogenetic Classifications

The constituents of phylogenetic classifications are taxa: species and monophyletic taxa. A taxon is a group of organisms, and taxon names are proper names. As there are different views on proper names, we will discuss their nature in a section below. For now, we are worried about taxa, not their names. There are many possible groupings of organisms, but only those grouping of species that result in monophyletic taxa *sensu* Hennig (1966) are recognized as natural in the phylogenetic system. We claim an even deeper significance: monophyletic groups are the only natural taxonomic groups of species in evolutionary biology. The relationship of a constituent to the group is a part–whole relationship. Obvious nonmonophyletic groups of species are dismembered and allocated to monophyletic groups if a phylogeny is available. Many taxa are not associated with a phylogeny at all, and these groups

serve as placeholders in general classifications until such time as they are subject to phylogenetic investigation.

As Hennig (1966) stated, the actual classification of a particular phylogenetic hypothesis is a relatively straightforward procedure accomplished by applying whatever conventions the investigator wishes to adopt. Any classification that is logically consistent with the hypothesized phylogeny is a phylogenetic classification (Wiley, 1981b). The extent to which the classification accurately reflects the topology of the phylogeny is the extent to which the classification informs the community as to the evolutionary/genealogical relationships of the organisms classified. If two phylogeneticists construct different classifications of the same organisms and agree upon the phylogenetic hypothesis, then the difference lies in the conventions adopted. For example, one investigator might construct a completely subordinated classification (McKenna, 1975) while another might name only terminal taxa and use a listing convention to show relationships (Nelson, 1974a). Both classifications are logically consistent and fully informative of the tree, and there are only two rules for classification to be termed *phylogenetic*:

1. Taxa classified without qualification are monophyletic groups or species (Hennig's Criterion; Hennig, 1966).
2. The classification must be logically consistent with the phylogeny, and the conventions adopted must reveal the genealogical relationships among the groups and species classified (Hull's Criterion; Hull, 1964).

The traditional conventions for classifying taxa are embodied in the Linnean system of nomenclature, and we shall use Linnean nomenclature. The advantages and disadvantages of alternative systems are discussed after introducing the Annotated Linnean Hierarchy.

THE LINNEAN HIERARCHY

The Linnean Hierarchy is one of several conventions for classifying phylogenies, and scientists who study each major group of organisms (plants, animals, prokaryotes) have developed rules for naming and the use of names for their organisms. (We will review these rules in Chapter 11.) Species names are formed in two parts, using a genus name and a species epitaph. Taxa of higher rank (genus and above) receive a single name. (More elaborate names are available for taxa recognized below the species level.) The Linnean system expresses the relative position of a taxon within the hierarchy by using a set of tags, categories, that denote relative subordination of taxa relative to other taxa and amends the root of the name with a suffix that is particular to the rank assigned for certain categorical rank levels. Within a single clade, taxa of high rank are hypothesized to have originated earlier than taxa of low rank. Occasionally a taxon of relatively high rank will contain a single species. In this case, the higher taxon is redundant or monotypic (containing the same species as the taxon included within it; Buck and Hull, 1966). Such a taxon functions purely to denote relative age ("age of origin," Hennig, 1966:162) and position of the species or clade in the hierarchy. Two features of the Linnean system must be understood by systematists:

1. Higher rank categories are not comparable between clades. Put simply, a family of frogs does not have the same biological characteristics (e.g., time of origin, degree of distinctness, etc.) as a family of tulips. Rank is relative within clades not absolute between clades. Thus, rank has no particular biological meaning. See Forey et al. (2004) for a particularly good discussion of this point.

2. Named species are potential units of process and as potential units of process they may be compared across clades to study the general characteristics of speciation.

The Linnean system has three major disadvantages:

1. Rank categories and the relative position of each rank to the others must be memorized.
2. Shifting ranks cause changes in the suffix of certain group names that are formed with roots and suffixes.
3. A great number of categorical ranks and name endings would be needed in order to completely name and rank every clade of organisms.

These disadvantages may be a major motivation for seeking different ways of expressing hierarchical position within the Linnean system, and some have been cited as a reason for abandoning the system altogether, as we shall see.

Definition of Linnean Higher Categories

Five higher category ranks are commonly used in the various Codes of Nomenclature: genus, family, order, class, and phylum/division. We shall briefly touch on the fine points of nomenclature in Chapter 11. For now, we wish to draw distinctions between old and new concepts of the categories themselves.

Phylogeneticists take a different attitude toward higher categories than such workers as Mayr (1969:92) who defined the genus as a taxonomic category separated from other genera by a decided "gap." Wiley (1979c, 1981a) noted that phylogeneticists rejected such gaps, and by 1991 Mayr and Ashlock (1991:135) agreed, using a completely comparative and pragmatic definition based on a deeper understanding of the difference between the nature of a taxon and its rank. We take the opportunity to further develop definitions based on Wiley (1979c, 1981a):

1. *Category.* A tag of convenience that denotes relative subordination of a taxon within a particular clade.
2. *Species Category.* A category below genus. Names of species are formed by the name of the taxon ranked at the level of genus plus a species epitaph and name formation follows the appropriate code.
3. *Genus Category.* A category between species and family. Name formation follows the appropriate code.
4. *Family Category.* A category between genus and order. Names of taxa assigned at this level in the hierarchy are uninominal and have endings that are set by the various codes.

5. *Order and Class Categories.* Categories between family and phylum/division. Names formed for taxa at this level in the hierarchy are uninominal, and the codes differ on how such names are treated (e.g., whether they are subject to priority or whether endings are uniform).

6. *Phylum/Division Category.* A category between class and kingdom. Names formed at this level in the hierarchy are uninominal, and the codes differ on how names are treated.

7. *Kingdom Category.* The highest category normally treated by the codes. Again, names formed at this level in the hierarchy differ among the codes.

There are additional categories formally recognized by various codes (e.g., tribe) and subdivisions of categories (e.g., subfamily), and there are also informal categories not recognized by the codes. For example, the informal category "species group" is frequently encountered as an informal category to group species within genera without introducing a formal category such as subgenus.

Conventions for Annotated Linnean Classifications

Regardless of the eventual fate of the Linnean system of nomenclature, international agreements in place at the present govern the names of taxa that use this system. Further, major repositories for genetic information, such as Genbank, use Linnean names, and there is no reason to think that Linnean nomenclature will disappear any time soon. Thus, it is worth dealing with. We will first discuss a series of conventions designed to minimize the use of rank categories. We will then suggest some modern ways to integrate Linnean nomenclature with Web-based technology to further extend its utility. Various authors have found parts of the conventions listed below useful (e.g., Judd et al., 2008).

Convention 1. The Linnean Hierarchy will be used, with certain other conventions, to classify organisms.

Convention 2. Minimum taxonomic decisions will be made whenever possible to construct a classification or to modify an existing classification. This will be accomplished in two ways. First, no empty or redundant categorical ranks and associated taxon names will be used unless they are needed to show the sister group relationships of a small clade or single species relative to its sister. Second, the ranks of well-known clades will be retained whenever possible.

These conventions simply declare that we shall use Linnean nomenclature and that the classification will be minimally redundant and maximally informative (following Farris, 1976). Note that the wording of Convention 2 differs from Wiley (1979c, 1981a) who advocated that redundant names be restricted to "mandatory" categories. No such restriction need be placed on the Annotated System because there are no mandatory categories. (Not even genus and species are mandatory as a fragment of a fossil can be classified to family without having to be classified to genus or species.) An example of the conventions is provided by gar classification (Wiley, 1976).

Division Ginglymodi
    Family Lepisosteidae
        Genus *Atractosteus*
        Genus *Lepisosteus*

The family contains only two genera. There is no need to assign each genus to a separate subfamily as the included taxon and diagnosis of each subfamily are redundant. Likewise, assigning both genera to a single subfamily would render the subfamily redundant relative to the family. However, the classification does contain one redundancy as Lepisosteidae is redundant relative to Ginglymodi. The redundant rank and name does serve a purpose: the sister group of gars (as presently understood) is the taxon division Halecomorphi, a clade that includes one living and many extinct species of fishes. Use of the redundant clade named Ginglymodi at the rank division is purely a device used to permit taxa to be ranked below the rank of division within the sister group of gars. Thus, although it is redundant relative to gars, it is useful, needed, and not redundant relative to bony fish classification writ large.

Convention 3. Asymmetric trees containing terminal taxa may be placed at the same hierarchical rank and listed in order of their branching sequence (Nelson, 1972a, 1974a).

This is the sequencing convention of Nelson. It is used to preserve the phylogenetic information concerning sister group relationship without the need to introduce additional hierarchical levels in the classification. For example, Schuh's (1976) classification of the hemipteran family Miridae exactly reflects the branching order of his phylogenetic hypothesis (Fig. 8.2).

Family Miridae
    Subfamily Isometopinae
    Subfamily Psallopinae
    Subfamily Phyllinae
    Subfamily Cylapinae
    Subfamily Mirinae
    Subfamily Bryocorinae

Without this convention, Schuh (1976) would have required 14 taxon names and appropriate categorical ranks rather than the seven names and two categorical ranks used.

Convention 4. Entirely fossil clades should be noted as such.

This is a revision of the original Convention 3 of Wiley (1979c), which called for the use of the rankless category "plesion" and the sequencing of extinct clades. We discuss the reasons for and the history behind the revision more fully below, but briefly here it derives from the fact that it was thought that fossil taxa were inherently less informative, when it came to phylogenetic matters, than extant taxa. Moreover, others were concerned that newly discovered fossil taxa might overturn or change existing classifications, necessitating the establishment of several new
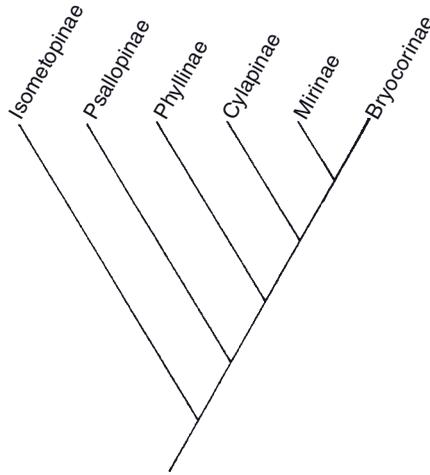
Figure 8.2. Schuh's (1976) hypothesis of relationships among subfamilies of the hemipteran family Miridae (after Schuh, 1976, from Wiley, 1979c).

taxonomic ranks. (This begs the question of whether something similar could happen with the discovery of a new extant organism.) The plesion convention can certainly still be used, but it is potentially problematic because it implies that an organism should be classified differently simply because it went extinct. Given that fossil taxa are just as much a part of the tree of life as extant taxa, and further, they are also natural, historical entities, we would argue that they should not be treated any differently. Although the plesion concept still is occasionally used, it is more common to rank fossil clades part and parcel with extant clades, although they can be set apart with the use of a symbol such as the dagger (†). A reasonable way to exactly reflect the relationships of a fossil clade to its recent relatives would be to use this convention in concert with the listing convention to preserve as many hierarchical ranks as possible.

As mentioned above, the placement of fossil groups vexed some early phylogeneticists. Convention 4 acknowledges how the community has dealt with the problem. It emerged from consideration of three proposals debated between 1966 and around 1980.

Proposal 1. Fossils should be classified separately from clades with recent organisms (Crowson, 1970) and divided along time lines (e.g., Hennig, 1966). This proposal creates a tidy classification of clades with living constituents and takes care of the placement of fossil ancestors, but there are problems. Patterson and Rosen (1977) point out that (1) there is a decrease in information content compared to a combined fossil-recent classification and (2) there would be a needless proliferation of names to accommodate relatively few fossil taxa. In addition, they pointed to some illogical features of such a scheme when the same taxon's range transcends the time boundaries assigned to different classifications. Wiley (1979c) pointed to other problems: (1) horizontal divisions in time are inherently arbitrary and likely to be clade specific, (2) the system produced paraphyletic grades, and (3) the same higher taxon

would occupy one hierarchical rank at one time interval and another rank at a later time period. Such objections also argue against Lovtrup's (1977) axiomizations concerning fossil classification, which were based on accepting Crowson's (1970) proposal.

Proposal 2. Fossil and recent organisms should be classified together and treated the same. This was McKenna's (1975) proposal (within the phylogenetic discussions of the time). It seems to be the proposal that has won out, but it does require that the ranks to which certain taxa are assigned in the hierarchy be radically adjusted from time to time. For example, McKenna (1975) demonstrated the utility of his scheme with a classification of Mammalia, but assigned the rank of class to the clade. In contrast, Nelson (1969) assigned ranks of suborder to Aves and Mammalia, which is more in line with his overall classification of Vertebrata, but this proposal was not well received among ornithologists nor mammalogists.

Proposal 3. Fossil and recent taxa should be classified together, but fossils should be treated differently (Hennig, 1966; Nelson, 1972a, 1974a; Griffiths, 1974; Patterson and Rosen, 1977). There were several proposals for how to accomplish this kind of classification. Hennig (1966) suggested the concept of the stem group (stammgruppe). *Archaeopteryx* and other fossil birds basal to recent ratites would be allocated to the avian stem group. The problem with this solution is that it encourages paraphyletic groups, just the kind of groups we wish to avoid. It is also not very useful for clades where fossil members of the clade are more diverse than living members (Patterson and Rosen, 1977). Nelson (1972a) simply suggested that fossils be tagged with a dagger and inserted within the classification at the appropriate level after categorical ranks are assigned to the living clades. Patterson and Rosen (1977) suggested that the rankless category "plesion" be assigned to fossil clades and that they be sequenced with their living relatives. The plesion concept can certainly be used, but seems out of favor at this time relative to the dagger convention of Nelson (1972a), which has long been used in traditional classifications to denote entirely fossil groups.

Convention 5. Monophyletic groups that form polytomies are given appropriate equivalent rank and placed *sedis mutabilis* at the level of the hierarchy at which their relationships to other taxa are known (Wiley, 1979c).

The *sedis mutabilis* convention is necessary in order to set apart lists of taxa that form ascending dichotomies from list of taxa that form polytomies. For example, the tree in Fig. 8.3 would be classified in the following manner using both the sequencing convention and the *sedis mutabilis* convention.

Family XYZidae
    Genus *X*, *sedis mutabilis*
    Genus *Y*, *sedis mutabilis*
    Genus *Z*, *sedis mutabilis*
        Z aus
        Z bus
        Z cus
        Z dus
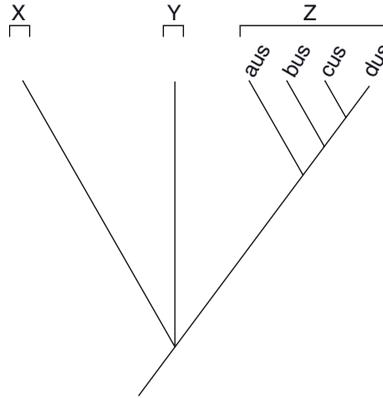
Figure 8.3. Phylogenetic relationships among some members of the hypothetical family XYZ (from Wiley, 1979c).
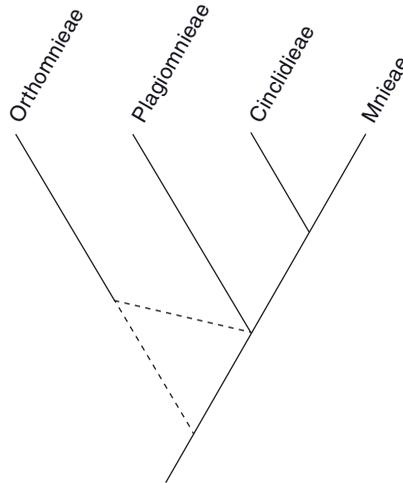
Figure 8.4. Koponen's (1968) hypothesis of relationships of mosses of the family Mniaceae.

Convention 6. Monophyletic taxa of uncertain relationships will be placed in the hierarchy *incertae sedis* at the level and ranks at which their relationships are best understood.

Although some workers have restricted *incertae sedis* only to fossil taxa (e.g., Nelson, 1972a, 1973a; Patterson and Rosen, 1977), this restricted use seems arbitrary and we prefer to follow McKenna (1975). The convention is used in both phylogenetic and traditional classification to denote ambivalence relative to the classification of a taxon of low rank relative to one or more taxa of higher ranks. For example, Koponen (1968) analyzed the bryophyte family Mniaceae and found that three of the four traditional tribes could be parsed phylogenetically into an ascending hierarchy (Fig. 8.4). One tribe, however, could not be placed as it lacked the synapomorphies uniting it with the other tribes. Using this convention, Wiley (1979c)

suggested the following classification, which uses both the sequencing convention and the *incertae sedis* convention.

> Family Mniaceae
> > Mniaceae *incertae sedis*: tribe Orthomnieae
> > > Tribe Plagiomnieae
> > > Tribe Cinclidieae
> > > Tribe Mnieae

Convention 7. A group whose status as monophyletic is unknown or suspect may be included in a phylogenetic classification if its status is clearly indicated by placing the name in shutter quotes to indicate that all included taxa are actually *incertae sedis* at the level of the hierarchy at which the taxon is classified. Such a group will not be accorded a formal rank.

This convention was first used by Patterson and Rosen (1977) to indicate the status of certain Mesozoic fish groups was either unknown (monophyly not demonstrated) or perhaps para- or polyphyletic. Specifically, the "Semionotidae" was considered a collection of diverse taxa that fits somewhere between the gars and the teleost fishes in actinopterygian evolution. Some might be more closely related to bowfins, others to teleosts and others spread out in the phylogeny between gars and teleosts. Part of Patterson and Rosen (1977) illustrates this use.

> Infraclass Neopterygii (higher bony fishes)
> > Division Ginglymodi (garfishes)
> > Division Halecostomi
> > > Halecostomi *incertae sedis*: "Semionotidae"
> > > > Subdivision Halecomorphi (bowfins and relatives)
> > > > Subdivision Teleostei (teleosts)

Ancestors in Phylogenetic Classification   The conventions presented above all deal with the classification of recent and fossil taxa while minimizing the use of rank categories. Exceptions include the placement of hybrid species and the placement of ancestral species (if known). Because part of the solution to hybrid species deals with how to place ancestral species, we shall take up ancestors first.

As mentioned in Chapter 4, ancestral species, if present in the analysis, should form polytomies with two or more descendant clades because ancestors have the synapomorphies of the group but none of the synapomorphies of descendant clades or descendant species. Thus the issue of whether one has an ancestor that needs classification depends on accepting a polytomy as a hard or true polytomy and not simply the result of missing data, a soft polytomy. A hard polytomy might be definitively recognized if repeated character analysis fails to resolve the polytomy. We would also assume that other conditions are met, such as the fact that we have biogeographic and statigraphic information amenable to the hypothesis that the species involved has characteristics we might expect from an ancestral species. Given that there must be many ancestors awaiting discovery (and perhaps many

have been discovered but were not called ancestors), no phylogenetic classification philosophy would be complete if phylogenetic classifications were incapable of correctly classifying ancestral species, regardless of the difficulties of that enterprise. Naturally, the issue of classifying "ancestral groups" does not occur in phylogenetics because such "groups" would necessarily be paraphyletic and thus discarded in favor of monophyletic groups. Whether or not we will ever have enough information to actually identify an ancestral species is another point of debate (Hennig, 1966; Brundin, 1966; Crowson, 1970; Griffiths, 1974); traditionally phylogeneticists have rejected the idea that an actual ancestral species could be identified as a stem species using the tools now at our disposal and this concern was clearly articulated by Hennig (1966:72):

> Naturally, in practice this [ancestor recognition] meets with basically insurmountable difficulties because it is scarcely ever possible to determine with certainty whether one (and in this case which) of the known species of *Archaeopteryx* (to continue our example) is the stem species of all other known species of Aves.

Patterson and Rosen (1977) suggested that ancestors be treated as terminal taxa. Given the peculiar topologies that are predicted to result when an ancestral species is included in a phylogenetic analysis, this proposal would invoke the *sedis mutabilis* convention. Consider, for example, that the ancestral species of all higher bony fishes was discovered. The classification would appear as below (based on Wiley and Johnson, 2010).

> Subclass Neopterygii (higher bony fishes)
> > *Neopterygius primus*, *sedis mutabilis*
> > Infraclass Holostei, *sedis mutabilis*
> > Infraclass Teleostei, *sedis mutabilis*

Such a classification bypasses controversies of whether or not *Neopterygius primus* is the ancestor of all descendants classified as neopterygian fishes. If one was bold enough to actually propose that *N. primus* was the ancestor, then some other convention would have to be applied. Indeed, in a truly general system of classification, we should anticipate that future investigators might be able to reliably identify stem species and take the view that, in general, phylogenetic classification must be able to accommodate all species, not simply descendant species (Wiley, 1979c, 1981a). Hennig (1966:71–72) gives a clue used by Wiley (1979c) for solving the "ancestor classification question."

> From the fact that … the boundaries of a "stem species" coincide with the boundaries of the taxon that includes all of its successor species, it follows that the "stem species" itself belongs in this taxon. But since, so to speak, it is identical will all the species that have arisen from it, the "stem species" occupies a special position in the taxon. If, for example, we knew with certainty the stem species of the birds (and it is only from such a premise that we can start in theoretical considerations), then we would no doubt have to include it in the group "Aves." But it could not be placed in any of the subgroups of the Aves. Rather, we would have to express unmistakably the fact that in the phylogenetic system it is equivalent to the totality of all species in the group.

Convention 8.   A stem species (ancestral lineage/ancestral species) of a supraspe-
cific taxon will be classified in a monotypic genus and placed in the hierarchy in
parentheses at the side of the supraspecific taxon of which its descendants are parts.

Given *N. primus* as the ancestor of all other neopterygian fishes, it would be clas-
sified in the following manner:

Subclass Neopterygii (*Neopterygius primus*)
  Infraclass Holostei, *sedis mutabilis*
  Infraclass Teleostei, *sedis mutabilis*

This convention treats stem species as biologically relevant, yet preserves the rela-
tionship between the phylogeny and the classification. Consider that we discover the
monophyletic nature of Aves through phylogenetic analysis. Although a monophy-
letic group now, Aves arose as a single ancestral species. It is perhaps irrelevant that
if we were alive during the Jurassic and collecting specimens of this species we would
never expect it to give rise to all birds. We would likely have been more concerned
that one or several large and carnivorous similar "dinosaurian" forms were apt to eat
us. It is only looking back through history and fitting Aves into the larger tree of life
that we arrive at the place in the hierarchy that causes us to assign the rank super-
order to Aves. If, sitting in the Jurassic, we could have predicted what would unfold
in the future, we might have placed *Aves ancestorcus* in its own monotypic superor-
der, but without knowledge of the next 150 million years, we very much doubt it.

The convention has additional benefits (Wiley, 1981a). (1) If we can discover
stem species, their incorporation into existing phylogenies will have minimal impact.
(2) Because the classification of ancestral species will have minimal impact, bold
hypotheses can be proposed without having to dramatically change classifications
to accommodate the hypothesis. (3) Only the phylogenetic system of classifying
clades provides a ready backbone of classification that can accommodate the place-
ment of all ancestral species while preserving both the logical correspondence
between the phylogenetic tree and the classification and the biological significance
of the ancestral species themselves.

There is an additional benefit that ties ancestors to the clade containing their
descendants in an empirical manner. Synapomorphies are the evidence we use to
circumscribe monophyletic groups. They are the historical effects of common ances-
try. That is, at least some of the descendants of an ancestral species have the prop-
erty we characterize as a synapomorphy, because they are descended from a common
ancestral species that had the property by the time it speciated to leave descendants.
Each synapomorphy "points to" at least one ancestral species where the synapo-
morphy was originally fixed as an autapomorphy and several may point to the same
ancestor when its history is understood in total. Theoretically, that apomorphy that
diagnoses the ancestral species is the same synapomorphy that groups the descen-
dants in the part–whole relationship. This is not a typological or essentialistic posi-
tion. In saying that all subsequent species are descended from an ancestor that has
a particular synapomorphy, we are not saying that all descendants must have that
synapomorphy. All tetrapods *sensu* Gaffney (1979) are descended from an ancestral
species that had the tetrapod limb, but not all descendants of that ancestor need
have a tetrapod limb (consider extant snakes for example).

Species and Higher Taxa of Hybrid Origin   Species of hybrid origin are relatively rare among animals (White, 1978) but more common among plants (Grant, 1981). Botanical classification recognizes such taxa as nothotaxa and furnishes various rules for naming them (*International Code of Botanical Nomenclature*, 2000). Naming is one task, but phylogenetic classifications must also be capable of showing the relationships of taxa of hybrid origin relative to their parental species (Wiley, 1981a). The solution we adopt here is similar to the solution adopted in annotated classifications for placing stem species.

Convention 9. Taxa of hybrid origin will be classified with one or both parental species and its hybrid nature (apart from any nomenclatural rules applied) will be indicated by placing the names of the parental species, if known, beside the hybrid's name in parentheses. The sequence of the hybrid in a list carries no connotation of branching relative to nonhybrid taxa in a sequenced list of taxa.

Wiley (1981a) used an example of phylogenetic relationships of some members of the composite genus *Anacyclus* L. published by Humphries (1979) to illustrate use of this convention. Of the 12 species of *Anacyclus*, 3 were hypothesized to be of hybrid origin (Fig. 8.5). The classification below uses the sequencing convention and the hybrid convention to illustrate both the origin of hybrid species and the origin of species from species further removed on the tree. Note that two of the sections in the listing convention were not named by Humphries (1979).
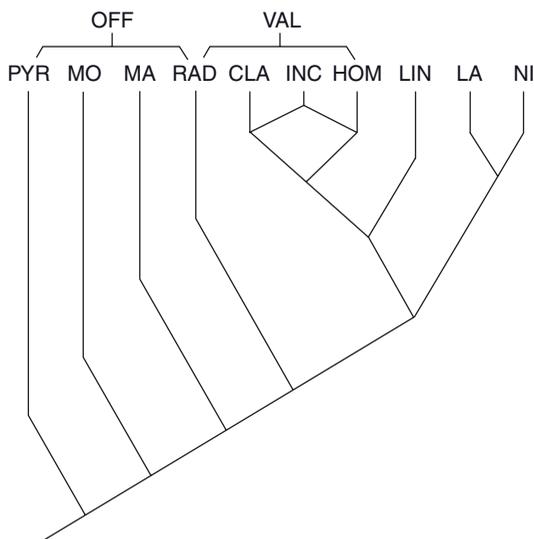
Figure 8.5. Humphries' (1979) hypothesis of relationships of the composite genus *Anacyclus* L. Abbreviations: CLA, *A. clavatus*; HOM, *A. homogamos*; INC, *A.* x *inconstans*; LA, *A. latealatus*; LIN, *A. linearilobus*; MA, *A. maroccanus*; MO, *A. monanthos*; NI, *A. nigellifolius*; OFF, *A. officinarum*; PYR, *A. pyrethrum*; RAD, *A. radiatus*; VAL, *A.* x *valentinus*.

Genus *Anacyclus* L.

    Section *Pyretharia* DC

        Anacyclus pyrethrum (L.) Link

        Anacyclus officinarum Hayne (*A. pyrethrum* x *A. radiatus*)

    Section *Anacyclus* L.

        Anacyclus monanthos (L.) Thell.

        Anacyclus maroccanus (Ball) Ball

        Anacyclus radiatus Loisel

        Anacyclus x *valentinus* L. (*A. radiatus* x *A homogamos*)

    Section (Unnamed 1)

        Anacyclus linearilobus Boiss. & Reuter

        Anacyclus homogamos (Maire) Humphries

        Anacyclus clavatus (Desf.) Pers.

        Anacyclus x *inconstans* Pomel (*A. homogamos* x *A. clavatus*)

    Section (Unamed 2)

        Anacyclus latealatus Hub.-Mor.

        Anacyclus nigellifolius Boiss.

ALTERNATIVE METHODS OF CLASSIFYING IN THE PHYLOGENETICS COMMUNITY

The Linnean system is only one of several ways of classifying organisms. Wiley (1981a) examined two others, numerical prefix schemes and rankless indentation. We shall examine both briefly in this section and then move on to the most recent proposal, the PhyloCode.

Numerical prefix schemes denote hierarchical rank with a prefix that is unique to each taxon and fixes the hierarchical level of each taxon relative to others. Early discussions of this approach are provided by Hull (1966), Hennig (1969, 1981), and Griffiths (1974). Hennig (1969, 1981) used numerical prefixes in conjunction with traditional Linnean names (i.e., the suffixes applied to names with roots and suffixes that conformed to Linnean nomenclature or to traditional practice within entomology) to classify insects in his landmark book *Insect Phylogeny*. An example of the system is shown in Fig. 8.6. Hennig (1969, 1981), began his classification with the prefix 1.0 to denote Entognatha and 2.0 Ectognatha. Each subordinate level added to the prefix designation (1.1, 1.1.1, 1.1.2, etc.). This produced an internally consistent classification and has some decided advantages. For example, fossil species can be inserted anywhere in the system without changing the hierarchical level of recent taxa (Griffiths, 1974). It is also preadapted to work well with computer languages, and the hierarchical levels are self-sustaining in that there is no need to create new rank categories as these are made "on the fly" as a consequence of adding taxa. There are disadvantages, however (Wiley, 1981a). Numerical prefixes are not the language of humans and are foreign to our efforts to communicate. The prefixes are unique, and thus, there will be as many prefixes as there are branches of the tree. It is easy to begin with 1.0 at

2.2.2.2.4.6 Mecopteroidea
2.2.2.2.4.6.1 Amphiesmenophora
2.2.2.2.4.6.1.1 Trichoptera
2.2.2.2.4.6.1.2 Lepidoptera
2.2.2.2.4.6.2 Antilophora
2.2.2.2.4.6.2.1 Mecoptera
2.2.2.2.4.6.2.2 Diptera
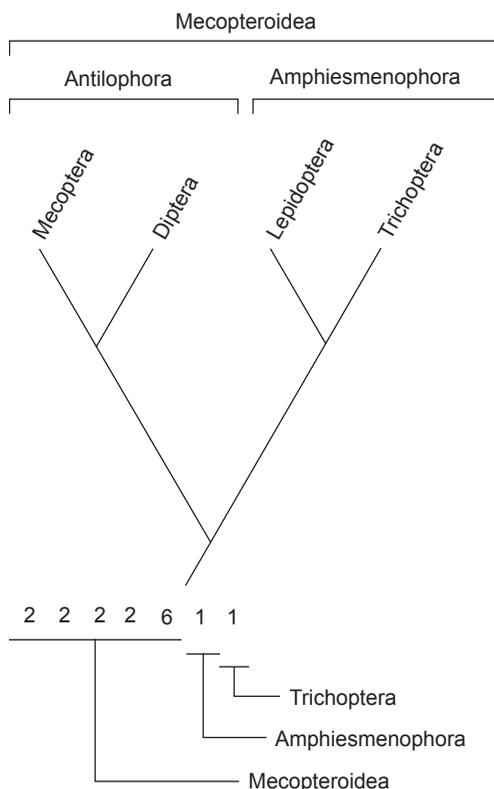
Figure 8.6. A hypothesis of the relationships among mecopteroid insects and the classification of Trichoptera using numerical prefixes (after Griffiths, 1974; from Wiley, 1979c).

Entognatha, but what if one began with prokaryotes? What would be the length of the prefix for Entognatha? Linnean ranks at least have the possibility of being reused (although they have the disadvantage of being interpreted by the unwary or unschooled of being biologically comparable between clades). Although Lovtrup's (1977) proposal that binary coding could be shortened by concatenation would certainly cut down on the length of the prefix, but this does not guarantee that the concatenation would yield unique prefixes (i.e., 1.1.1.1 = 4.0, but so does 2.2).

De Queiroz (1997:132) defends prefixes, stating that they are simple devices for representing hierarchical relationships (see also de Queiroz and Gauthier, 1992).

To de Queiroz, those who have criticized prefixes as "cumbersome and difficult to use in verbal communication" (e.g., Wiley, 1979c, 1981a; Eldredge and Cracraft, 1980; Ax, 1987) have misunderstood their use, as they are really not substitutes for Linnean ranks but "simple devices for representing hierarchical relationships." We make the following observations. First, if Linnean ranks serve to place groups of organisms such that their hierarchical relationships are shown, we fail to see why numerical prefixes, which serve the same purpose, are not "substitutes for Linnean ranks." Second, we leave to the reader to decide if 2.2.2.1.1.0.1 is simple and not cumbersome for humans. Third, we note that while Hennig used numerical prefixes, he also used formal Linnean name endings. In effect, Hennig (1969, 1981) hedged his bets.

Subordination by indentation is another alternative. In such schemes, one may use either ranked taxa or unranked taxa. If the taxa are ranked, the rank does not denote relative position in the hierarchy (Farris, 1976) but the indentation does serve this function. So, it is possible for one family to be included in another family by indenting. This clearly differentiates rank indentation schemes from indenting a traditional Linnean Hierarchy. In the latter, indentation is a way of presenting a visual clue as to the relative position of taxa. Pure indentation schemes dispense with ranks entirely and use only indentation to denote subordination. For example, the classifications below represent a traditional Linnean approach (above) and a pure indentation scheme (below).

    Class Vertebrata
        Subclass Myxini (hagfishes)
        Subclass Petromyzontia (lampreys)
        Subclass Gnathostomata (jawed vertebrates)
            Infraclass Chondrichthyes (sharks, etc.)
            Class Teleostomi (bony fishes and tetrapods)
    Vertebrata
        Myxini
        Petromyzontia
        Gnathostomata
            Chondrichthyes
            Teleostomi

Both classifications exactly reflect the underlying phylogeny generally accepted by many vertebrate systematists. The pure indentation scheme has an advantage in that if the phylogenetic position of hagfishes and lampreys changes relative to jawed vertebrates they can be easily moved. A disadvantage of the Linnean classification is that one has to memorize the hierarchical ranks order. However, rankless classifications have one practical difficulty (Wiley, 1981a): one must be able to line up coordinate taxa to confirm their sister group relationships. This presents no particular difficulty if the classification is relatively small and confined to a single page, but it does present difficulties when the classification is long and sister taxa appear on different pages. One might tend to lose one's place unless some standard was

introduced to measure the amount of indentation (Wiley, 1981a, suggested a marginal blue line for registration). In all fairness, we point out that this might also be a potential problem for the listing convention outlined above.

Our criticisms of both numerical prefixes and pure (or mixed) indentation schemes are practical, not theoretical. The fact is, all methods of classifying that are logically consistent with the underlying phylogenetic hypothesis and which are informative of that hypothesis serve the same basic purpose: to create a series of names of clades and their relationships that can be discussed by those interested in the history of the clades.

THE PHYLOCODE

A relatively recent development in classification is the proposal for an alternative formal system of nomenclature, the PhyloCode (available online: www.ohiou.edu/phylocode/index.html). The aim of the PhyloCode is the same as that of phylogenetic taxonomy in general (including phylogenetic taxonomy using Linnean nomenclature): to produce classifications that are logically consistent and fully informative concerning relationships among organisms. As of January 2011, this alternative code had not been implemented, but apparently it will be with the establishment of the *First Book of Phylogenetically Defined Names: A Companion to the PhyloCode*. This book will be the equivalent of the *Systema Naturae* for those who follow the PhyloCode and will contain names of clades that are approved by a PhyloCode nomenclature committee. Unlike the three major codes, the PhyloCode is not sanctioned by the International Union of Biological Sciences, and its rules of nomenclature will not carry the force of international "sanction" until such time as it is so sanctioned. This, of course, may be a matter of time or politics depending on the reception the PhyloCode receives when its governing body implements its rules. And besides, taxonomists have never let such formalities get in the way if they think the formalities are not useful.

The PhyloCode is designed to ensure the stability of names by defining the names of taxa through the use of specifiers. Specifiers are existing taxa or character homologies referenced to define the name relative to taxa included within the clade. There are three ways of defining the name relative to specifiers. The first is an inclusion statement: "Xinae is the name that refers to the clade stemming from the common ancestor of (the taxa named) *Xus* and *Yus*," where *Xus* and *Yus* are included taxa. The second is an inclusion/exclusion statement: "Xinae is the name of the clade that consists of all species that share a common ancestor with the taxon named *Yus* but not with that named *Zus*." The third is a synapomorphy statement: "Xinae is the name of the clade stemming from the first species to have the character 'hole on top of the head' that is homologous with the hole on top of the head found in the taxon named *Yus*." At least two specified things must be referenced in each case, either a combination of names of taxa or a combination of characters and taxa. There are alternative ways of specifying that are suitable depending on the intention of the author. Specifically, some forms of the definition are more suited for "crown clade" definitions (those designed to circumscribe only extant clades) as compared to "total clade" definitions where the taxon is circumscribed to include a number of more basal fossil species or clades. The PhyloCode differs in some important aspects

from most classification schemes in several ways. These should be understood by those wishing to employ it.

1. The PhyloCode assumes that taxa governed by the PhyloCode have certain biological characteristics; names must refer to clades (Preamble, 2). (Traditional codes have no such assumptions; they are "biology free.")

2. As it purports to govern only the names of clades, it does not serve as a general system for classification because explicit phylogenetic hypotheses are not available for many groups of organisms.

3. The PhyloCode extends the type concept and priority (through the use of specifiers and a registry of names, to be published) to all levels of the hierarchy in zoology, botany, and prokaryote classification.

4. Name endings do not change with changes in hierarchical rank. More to the point: the formation of names is independent of categorical rank.

5. Priority for the use of names is not determined by first use of the name in the literature but by first registration of the name in a list maintained by the governing body of the PhyloCode. Preexisting names can be "converted" by applying the appropriate definitional phrases and specifiers, but priority for the use of that name rests with the date the name was formally registered, not the date they were originally published.

6. Stability of names is achieved by fixing the name to a specific context determined by the specification as recorded. Once used in this context, the name is not available for use in other contexts.

7. The meaning of *clade* is not synonymous with the meaning of *monophyletic group*. A clade can be a monophyletic group of species or a clone or even individual organisms.

8. Clades can be overlapping (restricted to the case of taxa of hybrid origin).

9. There are several ways of recognizing clades (see above) but a single concept of monophyly. This can lead to difficulties depending on what kind of graph is referenced. For example, if one applies the "node-based" concept to a phylogenetic tree rather than to a Hennig tree, the ancestral species of the group would seem to be excluded from the group. This is problematic in a phylogenetic perspective because it is contra the entire phylogenetic enterprise, for instance, see Hennig (1966:71–72). To quote the PhyloCode (Article 2, Note 2.1.4 as of April 2007): "A node-based clade is a clade originating from a particular node on a phylogenetic tree, where the node represents a lineage at the instant of a splitting event." And from Article 9, Note 9.4.1: "A node-based definition may take the form 'the clade stemming from the most recent common ancestor of A and B. …'"

PhyloCode names differ from Linnean names in a number of other respects. Available names are placed in an approved database of clade names maintained by the PhyloCode commission. *Xinae* will forever mean "the name of the clade stemming from the common ancestor of *Xus* and *Yus*." If one wishes to extend the name *Xinae* to include more basal taxa, then one would form a panclade name, Pan-Xinae, and define it appropriately.

The PhyloCode also differs from Linnean codes in its treatment of ranks. Ranks might be used (they are completely optional), but name endings do not change with a change in rank. For example, if we find that Agamidae is a clade that includes Chamaeleonidae, then Chamaeleonidae would be included within Agamidae without changing the suffix of the root name. This treatment of names is one of the sticking points for those who use Linnean nomenclature. In particular, name endings mean something in Linnean nomenclature at certain levels of the hierarchy, where they serve as exclusion devices (a member of Agamidae cannot also be a member of Chamaeleonidae if both are monophyletic). However, name endings are meaning-less in PhyloCode nomenclature at all levels of the hierarchy. It should be noted that name changes in, for example, the Zoological Code, are only affected if the name is referred to a clade of the rank family or below, no name changes are governed for the names of taxa ranked higher than the family group. This does not mean that name endings will not change, but the changes are not governed by the code.

The PhyloCode also differs from the Linnean Codes in its view of the meaning of the taxon names. The Linnean Codes do not purport to give biological meaning to the names of taxa. In this regard, they are Millian, that is, they treat names as mere labels in accordance with the philosophy of proper names espoused by John Stewart Mill (1872). In contrast, PhyloCode names are Russellian, that is, they treat names as synonyms of their definitions in accordance with the philosophy of proper names espoused by Bertrand Russell (1919). For a discussion on these points, see Härlin (1998, 2003a) and Härlin and Sundberg (1998). We pursue this topic more fully below in the section on the meaning of proper names.

PhyloCode Controversies

A vigorous debate has ensued over the PhyloCode since it was first proposed by de Quieroz and Gauthier (1992). Part of this debate has resulted from a seeming mis-understanding, on the part of PhyloCode proponents, about the Linnean Codes. A number of such misunderstandings are significant.

Early advocates of what became the PhyloCode claimed that a new code was needed because Linnean classifications are essentialistic (de Quieroz and Gauthier, 1990), classification has not caught up with the Darwinian revolution, and Linnean taxonomic practices have inhibited the modernization of classification. However, these charges seem to be misplaced. In particular, claims that Linnaeus was an essentialist apparently trace back no farther than Cain (1958), and although repeated by such workers as Mayr (1959, 1963, 1968, 1976, 1982), they have been demon-strated to be false (Winsor, 2006). According to Winsor (2006) it is not clear that Linnaeus studied logic at all, much less Aristotelean logic. In fact, Linnean classi-fications made by Linnaeus and his followers "seemed to involve an active neglect of the classic rules of logical definition" (Whewell, 1847, cited in Winsor, 2003:3). Apparently the "essential characters" of Linnaeus were "key characters," those used in "keys" for identification; essential characters are characters of convenience not essential characters in the sense of Aristotelian logic. As evidence, one only need to consider the writings of Linneaus:

> Anyone who thinks that he can understand botany from the essential character and disregards the natural one is therefore deceiving and deceived; for the essential

character cannot fail to be deceptive in quite a number of cases. The natural character is the foundation of the genera of plants, and no one has ever made a proper judgment about a genus without its help; and it is and always will be the absolute foundation of the understanding of plants (Linneaus, 1751:143, translation in Winsor, 2006).

Further demonstration that Linnaeus was not following Aristotle is the fact that he used genus and species as fixed hierarchical terms; their use in Aristotelian logic is relative (e.g., bird is a genus containing the species swan; bird is a species contained in the genus animal; Winsor, 2006). This "nonessentialistic" concept of the Linnean system was in fact recognized by Darwin (1859:413–414):

> Such expression as that famous one of Linnaeus, and which we meet with in a more or less concealed form, the characters do not make the genus, but that the genus gives the characters, seem to imply that something more is included in our classification, than mere resemblance. I believe that something more is included, and that is propinquity of descent,—the only known causes of the similarity of organic beings,—is the bond, hidden as it is by various degrees of modification, which is partially revealed to us by our classifications.

Apparently, de Queiroz (1997:132) agrees. However, just because Linnean nomenclature is not inherently essentialistic does not mean it is to be preferred over a proposition such as the PhyloCode. It only means that rejecting the Linnean system because is it essentialistic is misplaced.

De Queiroz (1997) observes that changing paradigms from a creationist perspective to an evolutionary perspective did little to change Linnean classification or call into question the efficacy of the Linnean system. However, this change, according to de Queiroz (1997:128), "contradicted the Aristotelian context within which the Linnean Hierarchy was originally developed." Setting aside the question of whether Linnaeus was an Aristotelian (addressed above), it is certainly true that the basis for perceived hierarchy changed, whatever Linnaeus might have thought. But does this matter? The lag between the general acceptance of descent with modification and its explanatory power vis-à-vis groups-within-groups/part–whole hierarchies seems more a matter of two later developments: (1) the concept that genealogy alone should be the primary basis of relationships expressed in hierarchies and (2) the realization that many groups previously thought to be monophyletic were actually paraphyletic. These are conceptual issues that underlie the Hennigian Revolution and were not, so far as we can determine, inhibited (or promoted) by the use of the Linnean system of nomenclature. Thus, while it might be true that descent with modification played a rather superficial role in classifications between 1859 and today (a matter for historians of science to examine and which we doubt), it has not been established that the formalities of Linnean nomenclature *caused* this inhibition.

De Queiroz (1997) suggests that the realization that species are lineages effectively "redefined the Linnean category Species" and decoupled it from the rest of the Linnean Hierarchy. We certainly agree that species are lineages (Wiley, 1978, 1981a; Lieberman, 1992; Wiley and Mayden, 2000a; Wiley, 2002, 2007). Further, that species-as-taxa are different from clades-as-taxa is not doubted, nor do we doubt that real

species and real clades have objective reality apart from Linnean nomenclature. But we fail to understand why this would lead to rejection of Linnean nomenclature. The various international codes disavow placing biological meaning on names; biological meaning is left to the biologist to interpret (bringing us back to the distinction between Millian and Russellian meanings of words discussed previously). Systematists are free to consider a particular species such as *Fundulus nottii* as a lineage or a phenetic cluster. (Those who wish for species to function as comparative tools in evolutionary biology will gravitate to one concept, and those who see species names as only taxonomic devices may not.) As for higher taxa, we suspect that many are not clades, but there is little we can do about this until a phylogenetic analysis is performed on such a group's members and presumed relatives. In fact, the codes for the Linnean category "species" simply outline the rules for naming entities thought to be species and govern the use of such names when conflicts occur. It is up to the biological community to examine whether these names apply to lineages (or whatever your favorite species concept might be). It is true that there is nothing in the Linnean Codes that prohibits the naming of para- or even polyphyletic taxa. But it is also true that there is nothing in the Linnean system per se that prohibits a purely phylogenetic system of classification (Barkley et al., 2004a, b).

Two additional misconceptions need to be addressed. First, claims that the Linnean system is typological because names are defined by characters (de Queiroz and Gauthier, 1992) are false. In Linnean nomenclatural systems the *taxa* are "defined" or diagnosed by characters, not names of taxa. Names are formed using certain rules at certain levels. The form of the name may be dictated by the form of the name of a type, but if anything actually defines the name, it is the type (where typification applies), not the characters of the type. Second, in spite of repeated claims that there are five mandatory taxonomic categories (e.g., Laurin, 2005), such mandatory categories are nonexistent.

If we set aside their misconceptions surrounding Linnean Codes, PhyloCode proponents do have some valid points. For example, there are simply not enough categorical levels, especially within the genus group, to serve the needs of some practicing phylogeneticists (e.g., Hillis, 2006). In fact, this lack of hierarchical ranks is exactly what precipitated the many proposals incorporated into the Annotated Linnean Hierarchy and is behind other alternative systems discussed above. Theoretically, there is no reason to restrict the number of hierarchical ranks (Farris, 1976). However, coming up with unique suffixes is problematic, and thus, the inclusion/exclusion function of name endings might only serve for as many ranks as can be provided reasonable and grammatical endings to be added to the root. (Note that this is a problem not articulated by PhyloCode proponents.)

Second, it is true that changing a phylogeny may precipitate many name changes, with the same taxon changing name endings as it is pushed up or down the phylogeny. Of course, phylogeneticists who adhere to Linnean nomenclature might welcome such changes as signals of a paradigm shift in the classification of their groups (Härlin, 2003a). So, changing the meaning of names by changing the content of clades or recognizing that a name did not belong to a clade at all is not necessarily bad.

If we set aside adherence to pure tradition of Linnean nomenclature, there are other, additional criticisms of the PhyloCode. At one level, PhyloCode circumscription of taxon names is unproblematic. Indeed, the same strategy could be used

within the Linnean system, although one of the specifiers would have to be the type where typification applies. However, there are problems. These problems exist on two levels. The first problem is empirical: the stability of names relative to specifiers is given precedence to the stability of the total content of the clades named (point made by Härlin, 2003b). The second problem is ontological. To achieve stability of names relative to specifiers, taxa must be treated as kinds and not as individuals (Härlin, 1998, 1999, 2003b). This seems paradoxical, especially given that one of the original reasons for developing the PhyloCode was to combat essentialistic thinking, but we will show why this is the case.

Stability of Names Relative to Clade Content

The PhyloCode purports to define names through the use of specifiers. As pointed out by Härlin (1999) and Forey (2001), future research might have unexpected nomenclatural results relative to the contents of clades, if the relationships among the specifiers changes. The node relative to the specifiers might not have a stable place in the larger phylogeny, and thus, the ancestor involved might actually end up having different relationships and a different group of descendants than originally intended. One is left wondering if this is the same ancestor as that intended by the person who formed the proper name originally (it is not, as a matter of contingency; Härlin, 2003b). Traditional classifications have the same problem, shifting ideas of relationship results in changing content and meaning of clades. The difference is that PhyloCode names must follow the specifiers even with content change, while traditional (Linnean) names do not. Linnean names only have to follow the rules of priority if they fall within the scope of the particular code and community consensus if they do not.

Part of the problem with the reception of the Phylocode by its critics is what we might term the "part-whole" instability problem (e.g., Härlin, 1999; Nixon and Carpenter, 2000; Forey, 2001, 2002; Carpenter, 2003). While striving for the stability of names, it seems to create instability in the content of clades relative to names. Since no one must pay attention to the historical seniority of names as an arbitrator, and since Phylocode extends priority to all levels, the "neo-seniority" of Phylocode names by the registration process may cause unintended problems. For example, the phylogeny and names of three clades of apical bony fishes as of 2009 is shown in Fig. 8.7a. We form and register a node-based name, Halecomorphi, with specifiers *Amia calva* (AC, the living bowfin) and *Elops saurus* (ES, a teleost). Later we form the node-based name Neopterygii with specifiers *Lepisosteus osseus* (LS, a gar) and *E. saurus*. However, Grande (2010) demonstrates that the bowfin is actually more closely related to the gar than to the teleost (Fig. 8.7b). The name used since the 1920s to designate the group that includes gars, bowfins, and teleosts must now be replaced by the name adopted in the 1970s meant to *exclude* gars.

Stem-based names also create problems with unintended consequences. Consider Halecostomi as that clade defined as the clade that sprang from the common ancestor of *A. calva* but not from the common ancestor of *L. osseus* and *A. calva* (Fig. 8.7c) and the definition that Neopterygii is that clade that sprang from the common ancestor of *L. osseus* but not from the common ancestor of *L. osseus* and the paddlefish, *Polyodon spatula* (Fig. 8.7c). Grande's (2010) phylogeny would dictate that
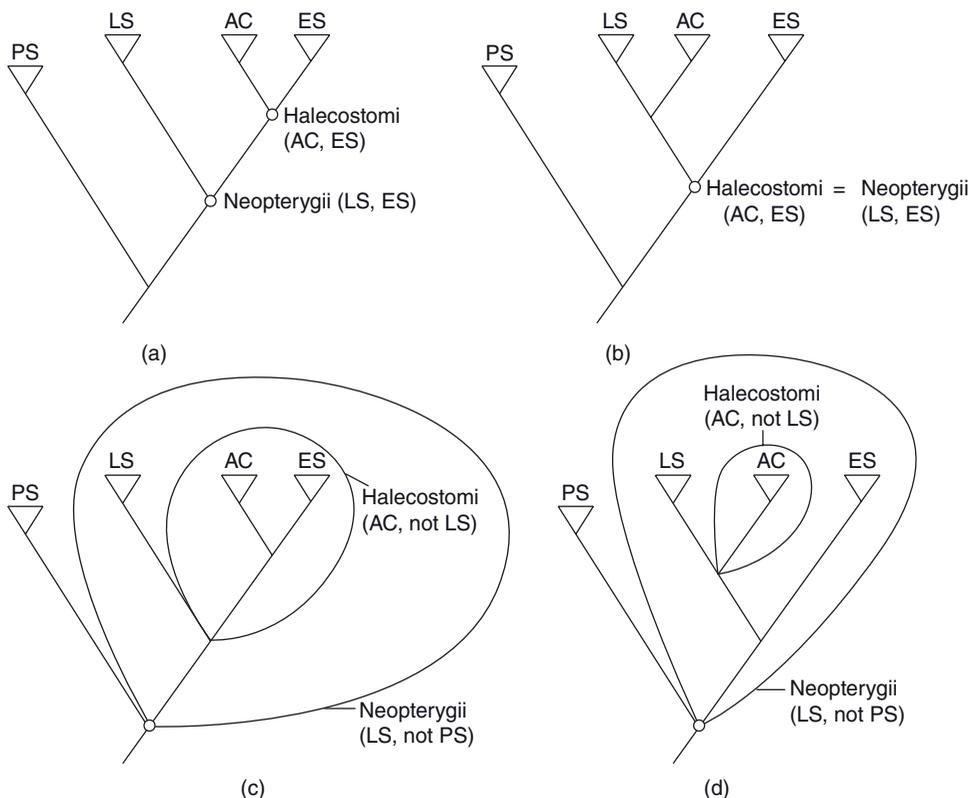
Figure 8.7. Name and content changes following the PhyloCode. (a–b) Alternative hypotheses of the relationships of basal actinopterygian fishes. If the 2009 community consensus hypothesis (a) is replaced by the alternative (b), then the names Halecostomi and Neopterygii become synonyms and Neopterygii is discarded as the younger name. (c–d) Node-based names fare no better. The tree in (c) is the traditional tree with groups circumscribed by ovals. If we accept tree (b), then Halecostomi is restricted only to *Amia calva* and the synapomorphy of this group must be changed from having an interoperculum to, for example, having double vertebral centra.

Halecostomi includes *A. calva* but excludes *E. saurus*, and this was not the intent of previous workers such as Rosen and Patterson (1977). Of course, we might be fortunate to have registered Neopterygii before Halecostomi, but since there are no rules about the historical priority of names relative to Phylocode registration, there is no guard against such a scenario.

Now consider what might happen to the characters associated with the names. The presence of an interoperculum has been traditionally associated with the name Halecostomi as a synapomorphy uniting bowfins with teleosts. Grande (2010) showed that some fossil gars, basal on the gar phylogeny, have interoperculars. In traditional nomenclature the presence of interopercular would simply be included in the diagnosis of Neopterygii (Halecostomi being discarded), and this would cause no problems with a node-based Phylocode Halecostomi (=Neopterygii) so long as

the community was willing to accept Halecostomi as a replacement for Neopterygii. If, however, we opted for stem-based names, Halecostomi would need a replacement character diagnosis since the interoperculum no longer diagnoses Halecostomi but the larger group Neopterygii. Finally, there is the problem of homoplasy; characters that are used as specifiers for clades that are later found to be homoplasious will cause needed name changes (Forey, 2001).

We have no doubt that the Phylocode can work as its adherents intend. In the end, after all, any system of nomenclature will work if we adhere to the rules of the system and this system contains no internal conflicts. And we do not question the good intent of Phylocode adherents to wishing to make nomenclature totally phylogenetic. However, we wonder, along with Forey (2001, 2002), whether the benefits are worth the costs. Is the stability of names rather than the community meaning of those names what we are really striving to achieve?

We have primarily been focused on perceived epistemological problems of Phylocode names. There is an additional ontological problem and we now deal with this problem because it is a larger issue of more general interest to our understanding of names in taxonomy.

PROPER NAMES OF TAXA

Given that taxa are individuals (Ghiselin, 1966; Hennig, 1966), the names of taxa are proper names. There are two basic philosophies of the nature of proper names that are related (broadly) to this discussion. These two philosophies cause controversy among phylogeneticists about the nature of proper names in taxonomy. The controversy is not easy to comprehend, because it underlies different philosophies of meaning and language. John Stewart Mill held that "a proper name is but an unmeaning mark which we connect in our minds with the idea of an object, in order that whenever the mark meets our eyes or occurs to our thoughts, we may think of that individual object" (Mill, 1872:22). Proper names are attached to an object (or entity) and are not dependent on any properties of the object. As Wettstein (1999:124) concludes: "A proper name, once attached, becomes a socially available device for making the relevant name bearer a subject of discourse."

In contrast, Russell (1919) held that proper names were substitutes for a description or set of properties. Russell's approach is associated with Frege in that both held that proper names are abbreviated descriptions. This is sometimes called the Russellian or descriptivist philosophy of simple proper names.

Mill's approach (the Millian philosophy) was considered by Kripke (1980), who argued that simple proper names (the names given to individuals) were not determined by a descriptive condition but by a causal chain that links name to references. Härlin (1998) called attention to these different approaches and asserted that only the Millian approach is consistent with the view that taxa are ontological individuals, as asserted by Ghiselin (e.g., 1966, 1995, 1997, 2007).

Both views have their philosophical problems, and two examples will show this. The Millian account of names has problems with statements that refer to nonexistent entities like "Santa Claus lives at the North Pole" or "Reptilia is paraphyletic." Millians counter that while neither Santa Claus nor Reptilia are real, thoughts are nevertheless communicated, and thus the proper names of unreal entities serve a

useful purpose in that communication is achieved. Kripke's "causal chain that links name to reference" can be invoked to ensure that communication is meaningful. "Reptilia is paraphyletic" can be meaningfully linked: "Reptilia *sensu* Romer (1966) is paraphyletic." Indeed, by linking name to reference, we can always know the meaning of a Millian proper name, so long as we know the reference. Patterson (1977) meant Teleostei to be that clade of bony fishes found in the phylogeny above the branch leading to *Amia calva* (a more inclusive clade). Arratia (1999) meant Teleostei to be that clade found in the phylogeny above *Proleptolepis* (a less inclusive clade). We completely understand how the same proper name is used in different ways if we understand the reference, or chain of inference. If one particular use of the proper name is dominant or its use is unproblematic, we dispense with the referent. For example, among neonotologists, uttering "Teleostei" conjures up the same concept of a particular clade of living fishes, but among paleonotologists who are worried about whether a particular fossil is or is not a teleost, a finer distinction might be needed to tie down exactly what we are talking about. Is Pachycormiformes a teleost *sensu* Patterson and *sensu* Arratia? (Answer: it is to Patterson but not to Arratia).

The descriptivist account has problems of rigidity and descriptive adequacy; while purporting to provide a description, the description when applied to evolving systems is frequently inadequate. For example, while the clade Tetrapoda *sensu* Gaffney (1979) is diagnosed with the synapomorphy of the tetrapod limb, not all tetrapods have this limb. (Indeed, no tetrapod has this limb during all phases of its life cycle.) Strict application of the definition would leave out snakes and other limbless tetrapods. Descriptivist accounts of natural kind names are quite another matter. Given that candidates for natural kinds have necessary and sufficient definitions, one could argue that helium, for example, is the kind name for those atoms with the property of having two protons. However, kind names are not proper names and philosophers of both camps generally agree that kinds do not have proper names.

Running through this controversy is the problem of what proper names should be referring to in the first place. The usual referent is a species or a clade. When Arratia or Patterson refers to Teleostei, they are referring to an entire clade (a whole), not just to some exemplars (only parts of the whole). These are contingent propositions that reference particular phylogenies, not necessary truths that refer to all possible phylogenies (Ghiselin, 1995; Härlin and Sundberg, 1998). They are not associated with necessary properties, but only contingent properties—those properties that are true given that the phylogeny in question is true. As such, they can hardly be mistaken for descriptivist proper names. This is so even within the Linnean system that uses type species to form names. The type species of *Homo* (*H. sapiens*) does not define the genus *Homo* in a descriptivist manner. Instead, it limits the content to the proper names of that clade, ranked as a genus, to those other species that share a common ancestor with *Homo sapiens* but not with, for example, *Pan tryglodytes* (because *Pan* is the name of another clade ranked as a genus). *Homo* is a Millian name, and if in doubt, a referent can be cited to tie down the meaning in a particular context.

Kripke (1980) coined the term *rigid designator* and claimed that proper names can only be used rigidly. This is a Millian concept, given that no definite description gives meaning to a proper name. But Kripke argues that one might "fix the reference" of a proper name. One way of fixing the reference is to point at the individual.

This amounts to fixing the reference by ostention. In phylogenetic classifications, the reference for clade names are to all the parts of the clade. The parts of the clade are "pointed at" through the inclusion of subclades within the named clade by classifying the parts. One always "knows" what a clade name means as a contingent proposition because the people who use the name signaled their intention for the meaning of the name by pointing directly at or alluding to the subclades within the clade so named. If we follow Kripke, this name would be rigidly designated and would apply to the clade in all possible worlds where the clade exists as a clade. (Of course, in worlds where the clade did not exist as a clade, then the name would not apply to anything.)

Another way to ostensively fix the reference for a name is to make a declarative statement of the sort advocated by advocates of the Phylocode. They convey the intention by pointing at two parts of the clade and making a declarative statement. For example: "Mammalia might be defined as the clade stemming from the most recent common ancestor of horses and echnidas" (de Queiroz, 1995:224). Fixing the reference through such ostensive "definitions" is said to lead to taxonomic stability. Unfortunately, this is not true, as the names are proper names. Consider Smith and Jones. Let's imagine the phylogeny of Smith has hippos and kangaroos joining the phylogenetic tree between echnidas and horses, but the phylogeny of Jones has kangaroos branching before echnidas. Mammalia *sensu* Jones is not the same clade as Mammalia *sensu* Smith because the whole of Mammalia *sensu* Jones is not the same whole as that of Smith.

The simple solution to this problem is to adopt the concept that taxa are not individuals, but rather, some sort of kind or class (Härlin, 1998). This seems the solution advocated by de Queiroz (1995). This would lead to exactly the same sort of stability as the stability of other kinds. Helium is always the kind of atom that has two protons. Members of the kind vary from day to day, depending on, among other things, the number of hydrogen fusions and uranium decays in the universe. Mammalia is always the kind that includes the common ancestor of echnidas and humans, and whether hippos or horses or kangaroos happen to be members of the kind is quite irrelevant.

But there is a problem with this solution. Mammalia, as a kind, can hardly be a natural kind. Mammalia is not predicted by any general theory of natural processes. Clades, in general, are predicted to exist due to speciation, but particular clades are not predicted; they are a matter of historical contingency. If Mammalia is a kind, it must be a nominal kind, not a natural kind. But if taxa are nominal kinds, then why are names of monophyletic taxa any better than names of polyphyletic taxa?

We conclude that the benefit of treating the names of taxa as kind names with necessary and sufficient properties in order to achieve name "stability" is far outweighed by the cost, both philosophical and phylogenetic. Stability in classification, as such, is not a particular goal of the phylogenetic system. Rather, consilience of the phylogeny with the classification is what we seek. And with that consilience the names will take care of themselves.

THE FUTURE OF LINNEAN NOMENCLATURE

The future is a matter of historical contingency and thus impossible to predict even if we know the constraints of history. However, there are several possibilities

for the future of biological nomenclature. With the rise of computers it is conceivable that classification per se might not be needed at all, simply some rules of priority. For example, we could hyperlink every name to a tree graph and place the names in any order we wish. Select a name from the list and up pops a tree detailing its closest relatives. We suspect that nomenclature could evolve into even more hybrid systems than now provided by systems like the Annotated Linnean Hierarchy. For example, Linnean ranks could be used to denote some number of levels of hierarchy with appropriate rules (priority, naming, etc.) and within these levels another system could be used (e.g., indentation, numerical prefixes, or simply lists hyperlinked to trees). How phylogenetic classification will evolve is up to the community and is unlikely to be well served solely by committees, however well meaning.

ALTERNATIVE "SCHOOLS" AND LOGICAL CONSISTENCY

In Chapter 4 we used the tool of logical consistency to examine the nature of paraphyletic groups relative to monophyletic groups. We demonstrated that there were two kinds of taxon groupings relative to phylogeny: monophyletic and nonmonophyletic. This was based on the observation that paraphyletic groups, like polyphyletic groups, were not logically consistent relative to a phylogeny that contains the groups. Wiley (1981a) used a variety of arguments to counter the claims by evolutionary taxonomists such as Mayr (1969, 1974), Ashlock (1971, 1972), Simpson (1961, 1975), and Bock (1974) that their classifications containing paraphyletic groups were superior to those containing only monophyletic groups. Wiley (1981a) refuted the notion that the school of evolutionary taxonomy was superior to the school of phylogenetic systematics. However, Hull (1964) had provided a more succinct argument earlier. He showed that the claims by Simpson (1961) that classification should be logically consistent with the underlying phylogeny were true, but Simpson's (and other evolutionary taxonomists') recognition of paraphyletic groups renders classifications containing such groups logically inconsistent with the underlying phylogeny. Again, this undermines the entire program of evolutionary taxonomy. This fact, and the fact that phenetics has largely disappeared, negates the need for a separate chapter contrasting the phylogenetic system with alternatives, because there are no other.

CHAPTER SUMMARY

- Phylogenetic classifications are systematizations.
- Phylogenetic classifications are logically consistent with the phylogeny that they purport to summarize.
- There are a variety of ways of classifying that are both natural and useful but have different knowledge goals.
- Classifications of natural kinds tend to be nonhierarchical or only partly hierarchical.

- Phylogenetic classifications are part–whole hypotheses and fully hierarchical.
- Linnean ranks do not rank comparable groups between clades.
- Classifications using the Linnean system are but one of several ways of achieving the goals phylogenetic classification
- An annotated Linnean system is summarized.
- The PhyloCode is critiqued, and the philosophy of proper names is discussed.

9

HISTORICAL BIOGEOGRAPHY

Historical biogeography is the study of the geographic distributions of organisms over relatively long time spans, time spans of thousands to millions of years. Our goals in this chapter are to describe the nature of this research area and elucidate how phylogenetic methods provide productive ways to study it. Historical biogeography has played a significant role in our understanding of evolution, and the study of biogeographic patterns played a major role in convincing scientists like Wallace and Darwin that evolution occurred. Today biogeography remains a discipline fundamentally relevant to evolutionary biology; this continued relevance can be partly attributed to research techniques that allow scientists to study biogeographic patterns and processes in greater detail.

As we turn from reconstructing the pattern of evolutionary descent to unraveling the processes that may have driven this pattern, evolutionary biologists and phylogeneticists tend to posit several distinct mechanisms. Among the most important mechanisms are those relating to abiotic factors that are external to organisms, including climate and geological change. These are distinguished from other important mechanisms involving internal or biotic factors such as competition and gene flow. A question that has often struck us (and other phylogeneticists) as intuitively interesting is, what are the relative contributions of biotic and abiotic factors in shaping the evolution of life on the planet? We consider parsing out biotic and abiotic factors as well as historical and proximal causes for animal and plant distribution to be important because they allow us to isolate and study causal factors in their proper context. In this chapter we shall argue that data from phylogenetics are essential for addressing these questions.

Ultimately part of the reason why phylogenetics matters for biogeographers is straightforward: we use phylogenetics to reconstruct evolutionary patterns; these patterns can in turn be used to test evolutionary processes (Eldredge and Cracraft, 1980; Wiley, 1981a) including the influence that abiotic forces exert on the evolution of life. Croizat (1964) expressed the pithy dictum that the Earth and life have co-evolved. If this dictum is true, there is a fundamental connection between biogeography and evolution, and phylogenetic methodology can be used to test the extent of this co-evolution and the nature of the underlying pattern. There are strong analogies between these fields (e.g., Croizat et al., 1974; Brooks et al., 1981; Nelson and Platnick, 1981; Wiley, 1981a, 1988a, b; Brooks, 1985; Brooks and McLennan, 1991; Lieberman, 2000a; Morrone, 2008), and in this chapter we will explore their similarities and differences in some detail. The similarities suggest a close alliance, but the differences point out that these fields should not be treated as identical. Analogy does not imply identity, and therefore a one to one map between the aims, precepts, and practices of phylogenetics and biogeography seems ill advised. The analogy breaks down because biogeographic areas are not ontologically equivalent to phylogenetic taxa. We will explore the connections between phylogenetic methods and biogeographic methods, concentrating on those that can be used to study biogeographic patterns in a phylogenetic context.

We will also focus on aspects of the history of biogeography and suggest that many of the concerns of both early biogeographers and evolutionists are still relevant today. For example, it is now universally recognized that most speciation occurs in a geographic context. Darwin endorsed this view in his notebooks but paid only token attention to it in the *Origin*, much to the detriment of our understanding of speciation. As another example, the relative roles of vicariance and dispersal were active debates in the nineteenth century, just as they are in the twenty-first century. We will discuss how each of these processes can act in a congruent manner and how each must be considered by phylogenetic biogeographers.

Another topic much debated by early evolutionary biologists was what patterns in the fossil record can tell us about evolution. We will consider the relevance of the fossil record for our understanding of biogeography, and in particular focus on how extinction and the incompleteness of the fossil record effects our ability to reconstruct biogeographic patterns. We will argue for a view that incorporates all biodiversity, living and extinct, to build a more detailed picture of biogeographic patterns. Finally, we will consider the current biodiversity crisis in a biogeographic context. We will argue that there are similarities between the current biodiversity crisis and past biodiversity crises (so-called times of mass extinction) as biogeographic phenomena.

THE DISTINCTION BETWEEN ECOLOGICAL AND PHYLOGENETIC BIOGEOGRAPHY AND THE IMPORTANCE OF CONGRUENCE

Biogeography is usually parsed into two major research programs: ecological biogeography and phylogenetic (sometimes called historical) biogeography (Brooks and McLennan, 1991; Brown and Lomolino, 1998; Lieberman, 2000a, 2003a; Morrone, 2008; Lomolino et al., 2010). The distinction between these research programs can be clearly seen if we consider the entities and processes in an explicitly hierarchical

context. We suggest that the distinction between these programs corresponds to the distinction between the ecological (or economic) and genealogical hierarchies (*sensu* Eldredge and Salthe, 1984; Eldredge, 1985; Lieberman, 2000a).

Ecological biogeography is a research program that focuses on entities and processes in the economic or ecological hierarchy (Brooks and McLennan, 1991; Lieberman, 2000a, 2003a). Ecological biogeographers are most interested in studying patterns in ecological entities like populations or communities while testing for the role of processes like population dynamics, competition, niche partitioning, and dispersal relative to the biotic context of communities and the abiotic environment. Examples of ecological biogeographic analyses at large scales would be the study of latitudinal diversity gradients (e.g., Stevens, 1992), species area relationships (Rosenzweig, 1995), or body size distributions in various taxa (Brown and Maurer, 1989).

Phylogenetic biogeography is a research program that focuses on entities and processes in the genealogical hierarchy like species and clades and testing how processes such as geological and climatic change influence speciation, extinction, and geodispersal (congruent range expansion of biotas). Examples of phylogenetic biogeographic analyses would be the study of how different clades of fish diversified in Central America (Rosen, 1978), the effects of Andean uplift on co-evolving vertebrates and their parasites (Brooks et al., 1981), the influence of geographic and geologic processes on modes of speciation in North American freshwater fishes (Wiley and Mayden, 1985), or the origins of the flora of the Indian subcontinent (Conti et al., 2002). The focus of this chapter will be on phylogenetic biogeography, but it is worthwhile to explore the distinction between these research programs and consider past views on this topic; this helps us to understand the manner in which those interested in the phylogenetics program approach the study of biogeography.

Some have tried to distinguish ecological and phylogenetic biogeography along the lines of whether or not dispersal is an important biogeographic process. It is supposed to be important for ecological biogeographers but not for phylogenetic biogeographers (Nelson, 1983; Patterson, 1983). We shall present evidence later in the chapter, however, that episodes of geodispersal are critically relevant to phylogenetic biogeographers and can be considered within a phylogenetic context. Indeed, we will further argue that if geodispersal is ignored by phylogenetic biogeographers the patterns they reconstruct may be incomplete and potentially confusing or meaningless. Because range expansion is important both to ecological and phylogenetic biogeographers, the difference between these research areas does not correspond to a difference between whether or not dispersal is an important process. Instead, the key element of phylogenetic biogeography is the search for congruent patterns of evolutionary descent across geographic space and between clades. Indeed, congruence is the means by which phylogenetic biogeographers test how Earth history changes influence these descent patterns (phylogenetic trees). A congruent pattern displayed among several clades that occur in the same region is evidence that Earth history played an important role in driving evolution. It would be evidence that the individual ecologies of organisms played a more muted role, at the grand scale, in determining the patterns of evolutionary divergence. By contrast, if different clades in the same region show different or incongruent patterns, it is evidence that factors specific to each clade's distinct ecology and biology played the more important role in driving evolution. Thus, an important distinction between

phylogenetic and ecological biogeography is that in the former the goal is to find congruent patterns in the evolutionary histories of clades of species while in the latter the goal is to find (in the best case) congruent patterns of convergence of individual species responses to biotic and abiotic factors. (These patterns, of course, are very interesting and demand explanation; they are just not the phylogenetic biogeographer's research program.)

Some have distinguished between ecological and phylogenetic biogeography on the basis of time. For example, it has been suggested that shorter time scales are the purview of the ecological program while longer time scales (deep history) are the purview of phylogenetic biogeography (e.g., Brooks and McLennan, 1991). This distinction between the two research areas is more useful than that of Nelson (1983) and Patterson (1983), but its validity needs to be considered in greater detail. For example, it is clear that species (as well as clades) can persist for long periods of geological time (an original thesis of Eldredge and Gould's [1972] punctuated equilibria model). Therefore, Brooks and McLennan (1991) were correct when they argued that phylogenetic biogeographic studies, at least of species and clades, must focus on patterns and processes that operate over deep time. However, some phylogenetic biogeographic studies will consider shorter time scales. Whether ecological biogeographic studies solely focus on short time scales or also focus on longer time scales depends on how long ecological entities persist (Lieberman, 2000a). One of the fundamental debates in ecology and paleoecology centers on the extent to which communities or regional ecosystems persist. If they are ephemeral, as some have argued (e.g., Davis, 1983; Graham, 1986; Foster et al., 1990, etc.), then clearly it is impossible to study ecological biogeographic patterns in such entities over the long term. But if communities or regional ecosystems are stable over the long term, as others have argued (e.g., Brett and Baird, 1995), then the study of biogeographic patterns in these entities should also focus on patterns and processes that operate in deep time. Brooks and McLennan (1991) may very well be correct that events in deep time are exclusively within the purview of phylogenetic biogeographic studies; but if not then a whole new realm of deep time ecological biogeographic studies is opened up.

Outlining phylogenetic and ecological biogeography along the lines of the study of patterns and processes in entities of the genealogical and economic hierarchies means that there will typically be a prominent distinction between these two areas. This is because most biological entities appear in only one hierarchy. Species and clades are in the genealogical hierarchy, but communities and regional ecosystems reside in the economic hierarchy (Eldredge, 1985, 1986; Lieberman, 2000a, 2003a). However, there are commonalities. Organisms belong to both hierarchies, and populations may as well, although there seems to be a distinction between reproductive populations or demes and ecologically interacting populations or avatars (Eldredge, 1986, 1989). Ecological biogeographic studies of individual organisms are usually undertaken in order to understand range dynamics including dispersal (c.f., Nathan et al., 2003, for review of techniques of studying long-range dispersal) and dispersion (discussed more fully below). Phylogenetic biogeographic studies done on the scale of several to many individual populations within species are frequently undertaken by phylogeographers, and these bridge ecological and phylogenetic biogeographic studies (e.g., Alexander et al., 2006). Biogeographic studies will naturally diverge into ecological and phylogenetic foci as the scope of the project increases from

populations to species to clades. However, all biogeographers need to consider processes like range expansion and all are potentially interested in processes and patterns that unfold over long (geological) time scales.

HIERARCHIES OF CLIMATE AND GEOLOGICAL CHANGE AND THEIR RELATIONSHIP TO PHYLOGENETIC BIOGEOGRAPHIC PATTERNS AND PROCESSES

Hierarchies of process are such that each hierarchical level is associated with a set of emergent properties and thus has its own distinct patterns and processes that cannot necessarily be extrapolated to explain the patterns and processes in entities at higher or lower levels. This characteristic of process hierarchies is significant to our biogeographic studies because there are a host of time scales over which various types of climatic and geological changes transpire. These climatic and geological changes can powerfully influence the geographic distributions of organisms and their patterns of evolution and extinction. For instance, these processes will cause entities to move (shift biogeographic ranges as they track habitat), become isolated (facilitating differentiation and thus speciation), and perhaps even vanish (go extinct).

Many aspects of climate change are related to astronomical cycles, and these types of climate change cycles run the gamut from days, seasons, years, and decades on up to the order of tens of thousands to hundreds of thousands of years (see the detailed discussions from a biogeographic perspective in Huntley and Webb, 1989, and Bennett, 1997). There is even evidence that climate cycles related to astronomical cycles might operate on the order of millions of years (Van Dam et al., 2006; Lieberman and Melott, 2007). Geological processes, distinguished from climate, and including plate tectonic changes, operate on time scales of hundreds of thousands to millions of years (see detailed discussion in Lieberman, 2000a). Thus, there is also a hierarchy of time scales over which climatic and geological changes occur.

The larger the entity studied in the process hierarchy and the older the entity, the greater the role of the longer term processes in shaping biogeographic patterns. For example, we might predict that short-term Milankovitch cycles play the greatest role in shaping biogeographic patterns within species. Tectonic processes often, but not always, operate too slowly to influence these. Climatic cycles operating on even shorter time scales might produce many of the biogeographic patterns within individual populations, although they are likely to operate too quickly or frequently to have other than ephemeral effects at the species or clade level. By contrast, we would also predict that the long-term Milankovitch cycles and also plate tectonics will play the greatest role in influencing patterns of biogeographic differentiation among species within a clade, and typically happen too slowly to effect changes within species or populations. We will discuss several examples of biogeographic studies that focus on patterns within and among species more fully below. Recognize for now, however, that the biological realm is organized into distinct, hierarchically arrayed entities; further, Earth history has presented life with an intergrading set of oscillating and changing conditions from the smallest to the largest scales that have influenced patterns of geographic distribution and evolution. As phylogenetic

biogeographers, we seek to tease apart the various patterns and processes that occur at the different hierarchical levels; indeed "all the sciences, and not just the sciences but all the efforts of intellectual kinds, are an endeavor to see the connections of the hierarchies" (Feynman, 1965:125).

THE IMPORTANCE OF VICARIANCE IN THE CONTEXT OF EVOLUTIONARY THEORY

If we consider microevolution to be evolution within species and macroevolution to be speciation and extinction, then the importance of vicariance to evolution in general and phylogenetic biogeography is immediately apparent. Allopatric speciation, while not ubiquitous, is the mode of speciation most commonly encountered in nature and the mode most often associated with tree hierarchies. This is not to claim that speciation only happens allopatrically, or even that in some groups other modes of speciation are not more common. But it is the mode of speciation in which we expect to find congruent patterns of distribution and the fact that allopatric speciation is so common (e.g., Coyne and Orr, 2004) leads us to methods that can discriminate between congruent and incongruent patterns of distribution that form the basis for subsequent evolutionary studies.

THE IMPORTANCE OF "DISPERSAL" IN PHYLOGENETIC BIOGEOGRAPHY

The problem with solely subscribing to either the dispersalist or the vicariance perspective is that neither is entirely complete without the other. For a species to be primitively widespread and later undergo vicariance, it must have somehow become widespread in the first place: presumably dispersal was required to attain the broader distribution. Similarly, when a species disperses out of a narrower area, one might ask why that distribution was originally narrow in the first place (Lieberman, 2000a). Also, even allopatric speciation via peripheral isolates requires some aspects of biotic dispersal, but it also requires pre-existing geographic barriers that would have been created by climatic or geological processes that could cause vicariance in other taxa.

Biogeographers have largely talked past one another on this subject, endorsing solely a vicariance or a dispersalist perspective. Part of the difficulty of reconciling these two viewpoints is the different meanings of dispersal when the term has been invoked. To the ecologist and population biologist, *dispersal* simply means the movement of individuals over a landscape; it is part of the dynamics of the interactions of populations with no necessary biogeographic consequences at all because it may occur entirely within the present range of a species. Excellent presentations describing dispersal through time are available in Huntley and Webb (1989) and Brown and Lomolino (1998), and outstanding paleontological examples are provided by Coope (1979).

To some historical biogeographers, dispersal tends to mean movement of populations of a species into newly occupied territory, usually over a pre-existing geographic barrier, coupled with evolutionary divergence (speciation, adaptive radiation,

etc.). This is the definition that has been used in a strict, cladistic context, e.g., Humphries and Parenti (1986). It is also the type of dispersal endorsed as among the most important biogeographic process by Darwin (1859, 1872).

The key aspect of this version of biogeographic dispersal is that it was defined to never produce congruence between Earth history and phylogenetic history; instead, it involved a single instance within a single species of range expansion caused by chance or unique ecological factors that would rarely be replicated in other lineages (Croizat et al., 1974; Platnick and Nelson, 1978; Rosen, 1978, 1979; Brooks et al., 1981; Nelson and Platnick, 1981; Wiley, 1981a, 1988a, b; Brooks, 1985; Wiley and Mayden, 1985; Kluge, 1988; Brooks and McLennan, 1991; Lieberman, 2000a; Morrone, 2008). Cladistic biogeographers were right to recognize that the problem with historical biogeographic studies that invoked "traditional" biogeographic dispersal was that it was a mechanism that could be invoked repeatedly and without the need for testability. Using traditional dispersal, biogeographic reasoning and analysis would never be able to account for biogeographic congruence except by invoking chance and time.

That several cladistic biogeographers denied the relevance of dispersal for biogeographic studies, e.g., Platnick and Nelson (1978) and Nelson and Platnick (1981), may seem paradoxical given the universal recognition that ecological dispersal is well known and documented. Of course, these authors understood this and attempted to use the term *dispersion* to describe ecological dispersal (Platnick, 1976). As real as the distinction between more ecologically relevant dispersal and more biogeographically relevant dispersal, the term *dispersion* already had a well-known meaning in ecology (the pattern of spatial distribution of individuals and populations, not movement of individuals). Phenomena such as seasonal migrations of song birds, terns, and whales seem better described as dispersal within the total range of a species, and lie within the scope of ecological dispersal and general life history patterns. Of course, other terms have also been co-opted, and we do not wish to make too much out of using terms well known in one discipline with changed meaning in another. *Immigration* used in MacArthur and Wilson's (1967) theory of island biogeography might be either kind of dispersal, depending on the circumstance, and *immigration* has a specific population genetic meaning, not just of dispersal and dispersion but of successful gene flow.

Geodispersal: Not Dispersal

At first blush *ecological dispersal* might not seem directly relevant to cladistic biogeography because as defined by Platnick (1976) it is not associated with cladogenesis, and hence will not be a macroevolutionary phenomenon. Further, it does not entail congruence. However, consider the case where a geographic barrier falls due to sea-level rise or climate change, or a new island appears. This may precipitate coordinated range expansion in several taxa congruently, such that dispersal occurs across different lineages, perhaps involving entire biotas: congruent biotic dispersal, if you wish. Then, imagine at a subsequent point in time a new geographic barrier forms due to sea level fall, climate change, or even continental collision; such a barrier could be in the same place as the original geographic barrier or in a different place; if the barrier is persistent, the result will be congruent vicariance in the many lineages that expanded their range through such dispersal. Or consider ocean island

chains: new islands become colonized as they appear and their biotic relationships reflect the order of their appearance.

Continental collisions facilitate vicariance and this type of dispersal: classic examples are the Great American Interchange between the mammal faunas of North and South America and the collision of India and Asia. India is a particularly interesting case. On its way toward Asia, it took a swipe at the Arabian Peninsula, and the floras of the region appear to have undergone episodes of this type of dispersal (Conti et al., 2002). (Note that in the case of India and Asia the continental collision that first enabled such dispersal eventually led to the uplift of the Himalayas, a major topographic barrier that has led to subsequent geographic isolation and vicariance.) This kind of dispersal produces congruence, and thus to distinguish it from more traditional dispersal concepts in biogeography Lieberman and Eldredge (1996) named it geodispersal. The name was chosen to recognize the fact that it is often geological or climatic processes that cause barriers to form and subsequently rise. One might wonder how geodispersal differs from more traditional ideas of dispersal. The key is to understand that under concepts of geodispersal, concepts like *center of origin* lose their meaning. Geodispersal involves the movement of ancestral species to increase their range, followed by vicariance. The descendants occupy the entire *center of origin*, rendering the term meaningless.

Geodispersal via the rise and fall of barriers is seen in the fossil record, where some of the paradigm examples involve trilobites, extinct fossil arthropods that were abundant and diverse in the Paleozoic. Lieberman and Eldredge (1996) found several examples of coordinated range expansion or geodispersal in Devonian trilobites that lived roughly 380 million years ago; moreover, the geodispersal appeared to oscillate with episodes of vicariance. Lieberman (2003b) subsequently described other examples of geodispersal from Cambrian trilobites. Each of these examples is discussed more fully below in the section where we describe methods of biogeographic analysis.

Recently, Halas et al. (2004) argued that the term *taxon pulse*, from Erwin (1979), was equivalent to geodispersal, and they preferred its usage when referring to congruent episodes of range expansion. The term *taxon pulse* is certainly a potential alternative, although we do not prefer it because Erwin's (1979, 1981) formulation of that term relied heavily on adaptive processes. In particular, he suggested that a taxon pulse was driven by a taxon's adaptive shift from one habitat to another (Erwin, 1981:175), which could be mediated by ecological mutualisms and coevolutionary dynamics, although climate and geology might play a role as well. Because *geodispersal* makes no assumptions about the adaptive nature of range expansion, which is often untestable, it is a less theory-laden term. Further, because taxon pulses are posited to be driven primarily by the distinct ecological characteristics of particular taxa, and congruence among relatively unrelated taxa is not expected, they are not directly analogous to geodispersal. However, it is clear that Halas et al. (2004) are correct that Erwin (1979, 1981) identified a process that shares commonality with geodispersal and is an intellectual antecedent. Indeed, although Lieberman and Eldredge (1996) were the first to use the term *geodispersal*, the concept actually has a long history, extending back nearly to the origins of the field of biogeography, as Lieberman and Eldredge (1996) and Lieberman (1997, 2000a, 2003c, 2006) acknowledged. At the end of the day, the debate about terminology is ultimately fruitful because it causes a renewed focus on the broad range of

biogeographically relevant phenomena: taxon pulses and geodispersal seem to describe real phenomena; the term to be used depends on the extent to which authors ascribe short time scale causes (dispersal, adaptation) or long-time scale causes (geological or climatic change) as the primary forces behind the biogeographic patterns observed.

Some important examples of what Lieberman and Eldredge (1996) termed geodispersal come from other paleontological studies. For instance, McKenna (1975, 1983) documented numerous examples of geodispersal by mammals during the Cenozoic, between Europe and North America and between Asia and North America, with episodes of geodispersal oscillating with episodes of vicariance facilitated by geological and climatic changes. Hallam (1977, 1981a, 1981b, 1983, 1994), who published several pioneering paleobiogeographic studies, documented numerous examples of geodispersal in fossil invertebrates. Vrba's (1980, 1985, 1992) Turnover Pulse hypothesis was also developed to explain patterns of oscillating geodispersal and vicariance in fossil African mammals, with climatic cooling causing different lineages of tropical mammals to become restricted to narrow, forested refugia. Such conditions led to extinction but also to population differentiation and speciation. Later, when climatic conditions ameliorated, their preferred habitats expanded and the lineages of tropical mammals would geodisperse outward. These changes would occur on timeframes of tens to hundreds of thousands of years, associated with Milankovitch climate cycles.

More recently, Sereno (1997, 1999) and Beard (1998, 2002) have documented other examples of geodispersal in fossil vertebrates. The examples in Beard (1998, 2002) involved Cenozoic mammals moving between Asia and North America, and the geodispersal in many mammal lineages at this time appears to have been caused by global warming. By contrast, the dinosaurs studied by Sereno (1997, 1999) showed episodes of vicariance in the early Jurassic, followed by geodispersal in the early Cretaceous, followed by vicariance in the late Cretaceous (Lieberman, 2003c) that appear to have been more mediated by plate tectonic changes. Other studies of fossil organisms that have emphasized how both geodispersal and vicariance have produced congruent biogeographic responses can be found in the work of Rull (2004), Rode and Lieberman (2005), and Hembree (2006). Finally, recently Folinsbee and Brooks (2007) presented evidence for a fascinating set of biogeographic and evolutionary dynamics involving geodispersal and vicariance in our own clade, the hominoids, and other clades of African mammals: the hyaenids and proboscideans. They hypothesized that each of these clades displayed instances of vicariant differentiation within Africa followed by geodispersal out of Africa into other regions including Asia and Europe. Further, vicariant differentiation then occurred within each of these regions, followed by geodispersal back into other areas including Africa (Fig. 9.1).

One prominent theme of recent paleobiogeographic studies is that at different time periods throughout Earth history the prevalent biogeographic mode has oscillated between vicariance and geodispersal. This makes sense given that certain types of geological changes, for example, widespread continental rifting, will affect a host of organisms in a similar fashion, though of course there are times when patterns in marine organisms may be the opposite of those in terrestrial organisms. This has important macroevolutionary implications because it means that due to abiotic conditions, at certain time periods rates of speciation might be unusually high,
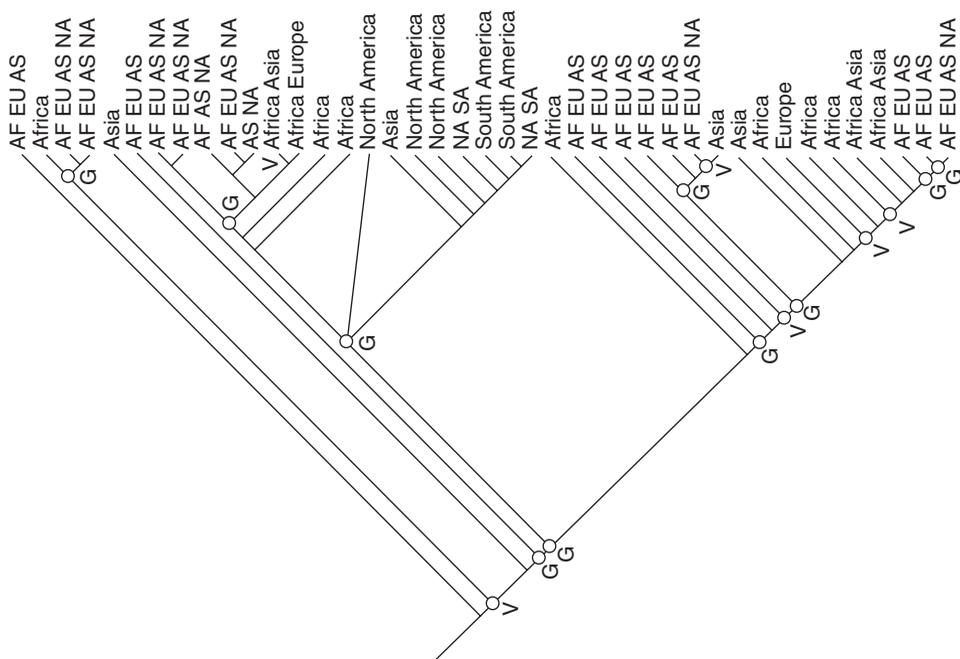
Figure 9.1. Area cladogram from Folinsbee and Brooks (2007) summarizing biogeographic patterns in three clades of mammals, hyaenids, proboscideans, and hominoids; each of these had been treated in phylogenetic analyses that incorporated both extant and fossil representatives. "*V*" refers to nodes where there were congruent episodes of vicariance replicated across several clades, and "*G*" refers to nodes where there were congruent episodes of geodispersal replicated across several clades. Notice that throughout the evolutionary histories of these clades there are repeated episodes of vicariance, followed by geodispersal, followed by subsequent vicariance. Used with permission of Dan Brooks, Wiley-Blackwell, and the *Journal of Biogeography*.

because of abundant vicariance; other time periods, by contrast, may show more muted rates of speciation and evolution. This topic is discussed more fully below.

Examples of geodispersal in early cladistic and phylogenetic biogeographic studies. Early on, many phylogenetic biogeographers recognized the potential significance of geodispersal (although they did not use that term) to biogeography. For example, Brundin (1988), Cracraft (1988), Noonan (1988), and Wiley (1988a, b) all argued that geodispersal and vicariance likely oscillated as barriers fell and then later rose; subsequently Bremer (1992), Ronquist (1994, 1998b), and Hovenkamp (1997) endorsed similar views (Lieberman, 2000a, 2003c). Most recently, Brooks and McLennan (2002), Conti et al. (2002), Halas et al. (2004), Brooks and Ferrao (2005), Brooks and Folinsbee (2005), Wojcicki and Brooks (2005), Folinsbee and Brooks (2007), and Morrone (2008) provided strong endorsements of the importance of geodispersal. Brooks and McLennan (2002), Halas et al. (2004), Brooks and Folinsbee (2005), and Morrone (2008) also supported the use of analytical methods to document geodispersal, and we describe such analytical methods more fully below.

There are also excellent examples of geodispersal in the writings of cladistic biogeographers. This is perhaps ironic given that many of these authors argued

stridently against dispersal as a relevant biogeographic phenomenon. But Platnick and Nelson (1978), Rosen (1978), and Nelson and Platnick (1981) all recognized that geodispersal was necessary to produce widespread biotas that could be subsequently divided by vicariance. Nelson and Platnick (1981) went so far as to suggest that dispersal is vicariance in disguise, a statement that Brundin (1988) mocked, but also used in support of his contention that congruent range expansion, i.e., geodispersal, affected biotas. Given that some cladistic biogeographers recognized that all biotas must have been affected by one early episode of geodispersal, it may not require much to take the additional step of recognizing that biotas may be affected by many episodes of congruent range expansion or geodispersal, followed by vicariance.

Conclusions. We introduce the long intellectual pedigree of geodispersal partly as evidence of its importance. Because geodispersal has powerfully influenced the evolutionary and biogeographic history of many, perhaps all, biotas, it is a process that must be taken into account by biogeographers. In particular, biogeographic methods need to be able to search for and analyze episodes of geodispersal. Those biogeographic methods that cannot capture or study episodes of geodispersal will produce results that are certainly incomplete, and likely also flawed and inaccurate. For this reason, in our discussion of biogeographic methods, we focus on an analytical method that can recover both episodes of vicariance and episodes of geodispersal. We also find the long history of discussion on the topic of geodispersal useful because it provides a possible means of resolving the long-standing debate among biogeographers about the relative importance of vicariance or dispersal (Lieberman, 2000a). In reality, the resolution of this debate comes from focusing on those processes that will produce biogeographic congruence: similarities in patterns of evolution across geographic space. Clearly "traditional" dispersal will not produce congruence (by definition). By contrast, vicariance and geodispersal will produce congruence. The focus of phylogenetic biogeography should be centered on the search for congruence, whether this congruence is manifested as vicariance or geodispersal.

Historical Perspective on Geodispersal and the Cyclical Nature of Oscillations between Vicariance and Geodispersal

The earliest scientific work that indisputably identified a process akin to geodispersal was Lyell (1832). Lyell argued that geological barriers were ephemeral over the vast span of Earth history; climatic changes or geological changes could cause barriers to disappear enabling large-scale migrations and the movement of biotas. Further, as a uniformitarian, he emphasized the cyclical nature of geological and climatic changes. The resultant biogeographic patterns influenced by these changes would be oscillatory and would correspond to what we today call geodispersal and vicariance. Given Lyell's influence on Darwin, it is not surprising that the Darwinian notebooks (Barrett et al., 1987) contain a few descriptive passages of phenomena resembling geodispersal (see Browne, 1983, and Lieberman, 2000a, for examples).

Wallace (1860), and even Wallace (1876), described biogeographic patterns involving waves of emigration caused by falling geographic barriers that resemble geodispersal (see Bowler, 1996, and Lieberman, 2000a). For example, he held that there had been biotic migrations from the northern to the southern continents that occurred when these continents became temporarily united (Wallace, 1876:155). The

Great American Interchange (Webb, 1978), was a geodispersal event that was well known to Darwin and Wallace. Huxley (1870), Beddard (1895), and Lydekker (1896) also argued that at various times during the history of life migrations between major regions occurred that were facilitated by the removal of geographic barriers. However, when examined in detail, each of the patterns these authors described differed from geodispersal as defined by Lieberman and Eldredge (1996) because they relied on competition as one of the primary forces driving the movement of organisms. Thus, their ideas are more akin to Erwin's (1979, 1981) taxon pulse. In contrast, Wortman (1903) and Matthew (1915, 1939) described patterns equivalent to geodispersal in the sense of Lieberman and Eldredge (1996) (see Lieberman, 2000a).

AREAS AND BIOTAS

Area and biota can be problematic concepts. Some of the problems arise when authors try to apply the same noun (and imply the same concept) at different scales and to different phenomena. We shall suggest that this is because there are actually two kinds of areas of biogeographic interest. Sometimes these different areas may be equivalent, but sometimes they are not. We will begin by considering the concepts of area and biota in very general terms. We will then attempt to sort out two biogeographically relevant concepts: geologic area and area of endemism.

Areas of the Earth may be purely nominal (Canada, a political unit) or natural. Here, however, we are primarily interested in natural areas of the Earth. In its most general from, we can say that a geologic area is an area individuated by geological processes and that the area forms a part–whole relationship with the Earth and shares relationships with other parts of the Earth. Relative to the Earth as a whole, geologic areas might be thought of as roughly analogous to parts of individual organisms. Because the Earth is dynamic, constantly undergoing a process analogous to ontogeny, such parts are ephemeral, although they may seem permanent over quite long time spans. We treat such areas as individuals (*sensu* Ghiselin, 1974; Hull, 1976, 1978, 1980; Wiley, 1978, 1979b, 1981a; Eldredge, 1985), and thus each has a history with a unique birth and death point, and some spatiotemporal localization throughout its history.

A biota is simply the sum of all individual organisms of all species living in an area, be it nominal or geologic. Biologists may study areas and associated biotas at many levels. Conservation biologists and park rangers working in a wildlife park may take great interest in inventorying and maintaining the biota of a park that is purely a nominal area. Such an area would be nominal because it is devoid of biotic endemism and is only circumscribed by the boundaries of the park; further, such an area might be continually exchanging taxa with other equally nominal adjoining areas. Ecologists working on a soil community may work with nominal areas of a square meter and may be primarily interested in diversity as it relates to community structure at this scale.

Evolutionary biologists may have no particular interest in geologic areas per se, but they can still use phylogenetic biogeographic techniques to search for patterns of speciation between clades whose members share one or more common species boundaries (e.g., Wiley and Mayden, 1985) in areas of endemism or biotas.

Phylogenetic biogeographers are primarily interested in correlating Earth history and biotic history, so they are interested in areas of endemism and the relationship of these areas to Earth history. The concept of *area of endemism* is not mature. We shall discuss some of the problems with the concept in a later section. For now, it is sufficient to characterize an area of endemism as an area of the Earth that can be circumscribed by the common geographic ranges of some number of species that do not have worldwide distributions. Areas of endemism may be distinct from geologic areas, but they can be treated as individuals in biogeographic analyses in the same manner as geologic areas. Indeed, they are, during their existence, properties of the geologic area on which they reside. For example, Africa is, for the present, a geologic area. It contains many areas of endemism (e.g., the Cape Flora) that are properties of Africa. These in turn may be geologic areas that have histories somewhat different from other parts of Africa; for instance, they may have once been separate smaller cratons that amalgamated to form a larger African craton in the distant past. If so, their biotic histories could have had, at least initially, nothing to do with Africa per se, but then ultimately they would become associated with different parts of a larger Africa.

The relationship between geologic areas and areas of endemism is obviously complex. If the area of endemism is constrained or circumscribed by a geologic area, then we can assume that there is a cause and effect relationship between the geologic history of the area and the biotic history of those organisms contained within it. In particular, there is likely some geologic feature that acts to constrain the area of endemism. This may effect many clades of organisms, and the greater the number of clades, the more distinctive the area of endemism. Not all areas of endemism are constrained or circumscribed by a geologic area, however. Some may be circumscribed by climatic conditions, and as climate changes so too will the boundaries and limits of the area of endemism change. Indeed, while two contiguous geologic areas might separate two endemic floras, if the barrier between them falls, the floras might mix. The geologic areas remain, but the areas of endemism disappear. If the barrier reappears, then new areas of endemism are established, but the parts of the Earth *relative to* the Earth may remain the same. We will treat each of these phenomena below.

Case 1. Geologic Areas and Areas of Endemism Covary. Consider two areas of endemism whose boundaries match those of geologic areas. If biogeographic congruence is found in the sampled endemic clades, we conclude that the history of the geologic areas played a direct causal role in the history of the biotas. Further, we can use the biogeographic patterns in these endemic clades to infer the history of the geologic areas they occur in. This provides an independent (relative to other geologic tests) biotic test of how the areas are related geologically. Indeed, in the case of the initial efforts to get plate tectonics (or continential drift) accepted, it was the study of such areas of endemism that helped cinch the case for the idea and led to the rejection of previously held hypotheses based on errant geologic assumptions (such as the notion that sunken land bridges joined Africa and South America).

As an example, consider a (marine) biota with endemic constituent species found in an ocean basin. The basin is then split in two by a fall in sea level, and ultimately some constituents of the biota speciate; imagine one species in the new basin *A* and

its sister species in the new basin *B*. The relationships of the species reflects the history of the areas they occur in. This situation resembles what is usually studied in the cladistic biogeographic research program (e.g., Humphries and Parenti, 1999). One way of proceeding in this type of research program would be to treat areas as if they were taxa in a Brooks Parsimony Analysis (Brooks, 1981; Wiley, 1988a, b; Brooks and McLennan, 1991, 2002; Morrone, 2008) where terminals are the areas under analysis and the nodes are ancestral areas that gave rise to the terminal areas by some vicariance event (rifting of continents, lowering of sea level isolating basins, etc.). In such vicariance analyses, the areas themselves are geologic areas supposed to have a common and unique divergent history, analogous to common, hierarchical phylogenetic ancestry, during the time period covered by the analysis.

In such classical cladistic biogeographic studies, which focus on vicariance, dispersal and extinction are usually treated as noise in the system. But this need not be the case. Related geologic areas may have a complex history of fragmentation and agglomeration, and the history of their biotas may reflect this history. Vicariance may reflect the response of organisms to the appearance of barriers, but geodispersal may reflect the response of organisms to the functional disappearance of the same barriers. For example, two continuous ocean basins may be separate bodies of water during periods of low sea level but a single basin during periods of high sea level. This may be reflected in complex relationships among their biotas that signal periods of time where isolation leads to speciation, continuity leads to faunal mixing, and subsequent isolation leads to more speciation. Unraveling the phylogenetic relationships among clades that responded to these events leads to insights into the histories of the areas involved. We shall see how such analyses are conducted in later sections.

Case 2. The Geologic Areas and Areas of Endemism (Biotas) Do Not Covary. The separation of once contiguous areas is not always directly correlated with the observation that two areas of endemism are related geologically. For example, Xiang and Soltis (2001) investigated disjunct biotas of several angiosperm lineages in the Northern Hemisphere. Chinese endemics have relatives endemic to Eastern North America. These two areas are thought to be remnants of a widespread Oligocene and Miocene tropical boreal forest that was once continuous across what are now vicariant areas of endemism. Obviously, the floras have a relationship, but the geologic areas in which they now occur do not have a unique geologic relationship *relative to* other Earth areas. We cannot infer something unique about China and Eastern North America as geologic areas as compared to, for example, Eastern North America and Western North America or China and Indo-China based on this biotic history. We can infer that these two biotas were once part of a continuous tropical boreal forest and seek explanations for that vicariance; climate change is an obvious candidate.

Case 3. Vicariance Without Areas of Endemism or Geologic Areas. Vicariance is simply the division of a gene pool by an extrinsic event. Some, such as Hovenkamp (1997), have argued that it is not necessary to focus on areas of endemism but instead simply attempt to match vicariance events with what appear to be common histories among groups. This is similar to the approach used by Wiley and Mayden (1985). It requires neither identification of geologic areas or biotas per se, but some level of congruence between the evolution of two to several clades that might be in

response to a vicariance event (of whatever nature). However, presumably a poste-riori such analyses could be used to argue for the existence of areas of endemism and possibly geologic areas, although that was not Hovenkamp's (1997) stated preference.

"Area" as It Relates to Phylogenetic Biogeographic Analysis

The distinction between geologic areas and areas of endemism has consequences for biogeographic analysis. If the goal of the study is to establish the relationships between geologic areas, then the assumption must be that the areas have a history, relative to the Earth that is independent of any organisms living in (or that had lived in) the area. The form of the analysis will be to treat the terminals as geologic areas and the nodes as ancestral geologic areas. Phylogenies will either be congruent with that history or incongruent. If enough congruence between the phylogenies of organisms that inhabit the geologic areas is found, this may lead to novel ideas about the history of the areas. Incongruence might signal that some (or all) of the taxa do not speak to the history of the areas in question, perhaps because dispersal or extinction obscures an otherwise clear pattern, or simply that the incongruent clades have a different biogeographic history.

If the goal of the study is to investigate the relationships among areas of endemism, then the terminals will be areas of endemism (nominal geologic areas) circumscribed by taxa endemic to those areas and the nodes will be ancestral biotas of endemics that might or might not be associated with an area. The ancestral biotas may not be additive areas of terminals, and they may occupy areas where no member of the endemic groups are found today. For example, the ancestral biota of the boreal forest hypothesized by Xiang and Soltis (2001) is not simply "China + Eastern North America." Rather, it is a large part of the Holarctic.

Fortunately, as it turns out, phylogenetic biogeographic techniques can be used in either case. It is simply a matter of the interpretation we place on the vertices of the tree graph. In an analysis of biotas where geologic areas do not play a causal role, hypothesized ancestral biotas appear at the vertices. In an analysis of biotas where geologic areas do play a causal role, ancestral geologic areas (the classic areas of vicariance biogeography) can appear at the vertices. Of course, an analysis might, not necessarily intentionally, consider a mix of each of these types of areas. However, in the first case (biotas) it is entirely possible that vicariance is caused, for example, by climate change. In the second case, vicariance is possibly caused by separation of the areas of the Earth via geologic processes. It would seem that the most direct connection between phylogenetic biogeography and areas is when these geologic areas play a causal role in generating the phylogenetic patterns. To understand that role and how phylogenetic biogeography can be carried out using areas, we must have a clear understanding of the epistemological and ontological status of such areas.

Consider the case of the continent North America; this is a coherent geological bloc well back into the pre-Cambrian and shows abundant examples of patterns akin to phylogenetic divergence. Around 750 million years ago Antarctica, southern China, and Australia split off from North America's then western margin; perhaps 100–150 million years or so later parts of Scandinavia and South America split off of its eastern margin, whilst Siberia split off from its present northern margin

(Fig. 9.2). Each of these geological events represents classic examples of vicariance, and the North American continental bloc was then largely independent from other continental blocs for hundreds of millions of years. Later though, toward the end of the Paleozoic, North America's eastern margin collided with Europe and Africa, forming the supercontinent Pangea (Fig. 9.3); also, at least since the middle part of the Paleozoic it had begun to amalgamate with other smaller continents or terranes that stuck to both North America's eastern and western margins, i.e., North America shows multiple "tokogenetic" relationships with other continents. (Interestingly, and as an aside, there is no evidence that these small terranes brought living biotas with them [Scott, 1997], although it has been shown that they did bring fossilized biotas with them [Ross and Ross, 1985].) Sometime during the Mesozoic, Pangea began to split apart and a new cycle of vicariance began. Ultimately, Greenland, which had been joined to North America since the pre-Cambrian, split away around 40 million years ago. Is the North America of the Cenozoic in any sense equivalent to pre-Pangean North America? Geologically perhaps yes, as the same continental crust and basement, absent the addition of a few younger layers, are in place, and its areal extent is comparable, setting aside the addition of the western most and eastern most margins of North America, along with the loss of Greenland. Biogeographically, the answer probably would be no, but one could use an epistemological criterion based on biology to see if any Cenozoic North American taxa are biogeographically derived from Paleozoic North American taxa such that there is a skein of evolutionary and geographic connectivity between the two.

Consider a related example involving South America. Throughout the Cenozoic, South America had a long and independent geological and evolutionary history until it collided with North America roughly 3 million years ago, triggering widespread geodispersal during the Great American Interchange. As of yet, however, the geodispersal has not been so rampant as to fully efface the fact that South America long had an independent history; South America still is an area of endemism (and not all that endemism is the byproduct of post–Great American Interchange divergence). South America as a geologic area is not extinct, and geodispersal has not effaced the biotic areas associated with it.

An interesting question becomes, when does extinction of an area occur? In the case of biotic areas, these become extinct when they no longer contain any unique taxa. In the case of geologic areas, they become extinct when they merge with other areas, split into different areas, or are subducted back into the Earth's interior. Of course, the merging of geologic areas might be signaled by geodispersal of the now united biotas, the splitting might be signaled by speciation and the subduction might cause extinction. But geologic areas do their thing on other geologically active planets and moons, and they do so without organisms at all. In summation, we prefer the view that geologic areas are individuals. They may diverge in a fashion that mirrors phylogenetic divergence, they may combine in a fashion that mirrors tokogeny, and they may disappear through climatic change or tectonic processes. Such is the fate of areas, biotic or geologic, on a dynamic world. The processes that individuate areas operate on very long time scales, and there may be long intervals of time when it is difficult to see the precise individual geologic areas that exist. Still, difficulties with individuation even exist in the case of biological organisms, consider colonial organisms, yet this does not make the definition of individual organisms intractable.

(a)

(b)

580 Ma

Figure 9.2. Paleogeographic reconstructions showing the approximate position of what were then the Earth's major continent blocs roughly (a) 750 and (b) 580 million years ago. These were part of a supercontinent that included Laurentia, primeval North America. The rifting that split up this supercontinent proceeded first on present day Laurentia's western margin, roughly 750 million years ago, and 150–200 million years later on its present day eastern margin. Major continental blocs are abbreviated in (b): Lau, Laurentia, North America, plus Greenland; Ama, Amazonia; Bal, Baltica; Ind, India; Aus, Australia; Sib, Siberia; Ant, Antarctica; Ara, Arabia; Arm, Armorica; Ava, Avalonia. Major oceans are labeled in bold. Parts of present-day South America and Africa, which were also once distinct continental blocs, are also abbreviated, including Rio, Sao, in the case of South America, and, in the case of Africa: Waf, West Africa; Con, Congo; Kal, Kalahari. Images courtesy of J. Meert, University of Florida. See color insert.
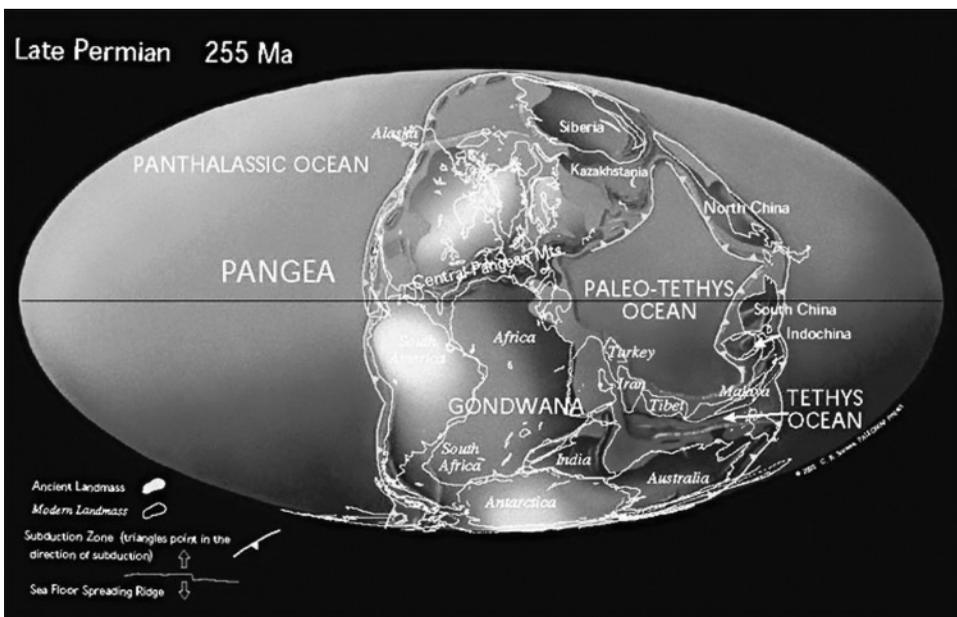
←——————————————————————————————————————

Figure 9.3. The supercontinent Pangaea, in existence from roughly the end of the Paleozoic Era to the middle part of the Mesozoic Era, roughly 250–160 million years ago. Image courtesy of C. Scotese, University of Texas at Arlington, Paleomap Project. See color insert.

The Boundaries of Biotic Areas and Comparing the Geographic Ranges of Taxa   Defining the ontological status of geologic areas of biogeographic interest is no easy matter, and defining the boundaries of a biotic area is not much easier. Epistemologically the presence of endemic biological taxa can be used to identify biotic areas, but what sets the limits of such an area? Individual species in an area have individualistic geographic range boundaries, and the ranges of different species in a region may coincide closely though usually not precisely. Given the typical lack of exact correspondence between species ranges, how can the precise boundaries of a biotic area be defined? The truth is that it is probably necessary to make some statement about whether geographic ranges of various species and taxa in a region are broadly homologous (Lieberman, 2000a). We have discussed the various issues

related to homologizing characters in Chapter 5; in that context, homologous characters are those that share commonality because they are derived from a common ancestor. By analogy, geographic ranges are presumed to be homologous if they were influenced by the same geological and climatic processes. Homoplasious ranges would then be those ranges that may appear homologous but are "achieved" independently. Because geographic range maps are surfaces in three dimensions, it is not necessarily easy to come up with a quantitative algorithm that meaningfully compares them, though perhaps various outline-based techniques from morphometrics could be used, starting with two dimensions for simplicity. The difficulty in comparing geographic ranges of organisms is one reason to also incorporate information about the geographic and physiographic boundaries of the geologic/geographic region when delineating the geometry of biotic areas. Even with regions as pronounced as continents, their boundaries vary slightly through time: consider daily with tides and over millennia because of rising and falling sea-level driven by the waxing and waning of the ice sheets. Still, given the large scale of the areas being delineated, these differences are minor. Such detailed delineation of areas using geographic barriers is also possible at smaller scales. For instance, Rosen (1978), Wiley and Mayden (1985), and Mayden (1988b) conducted phylogenetic biogeographic analyses of freshwater fish from Central and North America; in these systems they were able to define biogeographic regions for these freshwater organisms that corresponded to the physiographic and topographic boundaries of river drainage systems. There are other times, however, when even physiographic and topographic boundaries may be fairly fuzzy: consider the case of current systems in large ocean basins.

Conclusions Currently, detailed, quantitative definitions of biotic areas that are workable are lacking, and this remains one of the underexplored areas of biogeography with real potential for significant developments. Certainly though, criteria previously used to define areas, albeit a bit nebulous, are workable. What is important to recognize, though, is that the number of areas used in a biogeographic analysis, and how these areas were defined, will have an important effect on the results of a biogeographic analysis. For example, subsuming two smaller areas into a larger area may lead to different results than if those areas were treated separately. In a sense this is not surprising. Think of how much taxon diagnoses matter for phylogenetic studies. Imagine if the definition of a taxon was changed, and some other taxa previously treated as distinct were subsumed into that taxon, resulting in different character codings for the newly defined taxon. This new taxon's position in a phylogenetic analysis could differ from its position in a previous analysis when it had been defined in another way. The same caveats will ultimately be true of a biogeographic analysis and are important to bear in mind given the lack of precise, consistent criteria for defining areas.

ANALYTICAL METHODS IN PHYLOGENETIC BIOGEOGRAPHY

Phylogenetic analysis is based on the use of characters to make hypotheses about how various taxa are related. As we have described, characters are evaluated to determine the best supported pattern of evolutionary relationship. Because rarely if ever is there complete congruence among all character data, a means must be chosen to evaluate among competing characters. The principle of parsimony is one

way of evaluating competing character data, and likelihood is another means (albeit in a different framework). Phylogenetic biogeographic analysis aims to determine area/biotic relationships (not taxic relationships) and relies on having phylogenies of clades available, and also information about where the taxa in these clades are distributed (Croizat et al., 1974; Nelson, 1976; Platnick and Nelson, 1978; Brooks, 1981, 1985, 1990; Brooks et al., 1981; Nelson and Platnick, 1981; Wiley, 1981a, 1988a, b; Cracraft, 1988; Brooks and McLennan, 1991, 2002; Morrone and Carpenter, 1994; Morrone and Crisci, 1995; Lieberman and Eldredge, 1996; Lieberman, 2000a; Morrone, 2008). We support the view that the character evidence used in phylogenetic biogeographic analysis is the geographic distributions of taxa, as well as information about how these geographic distributions have changed during cladogenesis. Please note that "area" in this context can be a geologic area [leading to part–whole relationships among such areas] or a biotic area, which may or may not be a geologic area. One may not even know which applies until after the analysis.

Sources of Signal in Biogeographic Analyses. The basic idea in a biogeographic analysis is that similar geographic distributions coupled with similar phylogenetic histories is reason to suspect that two or more clades were influenced by similar geologic or climatic factors. As we shall see, this is not simply a matter of vicariance and allopatric speciation; such congruence can also obtain when entire faunas and floras move, spreading their ranges. So, signal is not simply a matter of identifying common vicariance patterns but of understanding congruence and noise.

Sources of Noise in Biogeographic Analyses. There are several processes that might conspire to make different biogeographic patterns not fully congruent and thus represent noise that can affect any biogeographic study (Rosen, 1978; Platnick and Nelson, 1978; Wiley, 1981a, 1988a, b; Wiley and Mayden, 1985; Brooks, 1985, 1988; Lieberman, 2000a; Turner et al., 2009). Some of these sources of noise relate to the fact that individual clades can have their own distinctive biologies and will not all respond in precisely the same way to the various Earth history changes they have experienced. For example, one source of biogeographic noise is traditional dispersal, which is related to unique aspects of a clade, or species within that clade, that allows it to disperse incongruently over a geographic barrier while other taxa cannot. Sympatric speciation will also lead to biogeographic incongruence because it involves speciation in the absence of the formation of geographic barriers: instead, speciation is driven by various ecological and competitive interactions.

There are two other sources of noise that represent a subtly different category from those described above. This is because, depending on the biogeographic method used, they may or may not actually introduce a significant degree of noise into a biogeographic study. One example of this type of noise arises when geographic barriers form within a pre-existing region, but they produce geographic isolation and cause speciation in only some of the taxa occupying that region; such a situation might arise because certain types of geographic barriers are more likely to isolate one type of organism than another, due to different dispersal capabilities. This is sometimes referred to as noise arising from failure to speciate. Extinction is another source of noise (Lieberman, 2002b; Turner et al., 2009), especially for studies that focus solely on extant taxa. Extinction becomes a problem because certain taxa (the extinct ones) from certain geographic regions might not have been sampled, implying that evidence for the "true" patterns of area relationship might be absent. Given that 99.99 percent of all species that have ever lived are extinct, this is a nontrivial issue (Lieberman, 2002b). An equivalent problem emerges in biogeographic studies

focusing on fossil organisms. In such cases the problem is that the fossil record is incomplete, and not every species that has ever lived has been preserved as a fossil (or has been found, even if it has been preserved) (Lieberman, 2002b; Turner et al., 2009). The effects that extinction and paleontological incompleteness can have on our ability to study biogeographic patterns are discussed more fully below.

HISTORICAL BIOGEOGRAPHY USING MODIFIED BROOKS PARSIMONY ANALYSIS

Modified Brooks Parsimony Analysis (MBPA, termed Lieberman-modified BPA in Maguire and Stigall, 2008) is an extension of Brooks Parsimony Analysis (BPA) originally proposed by Brooks (1981). BPA (Brooks, 1981) is a biogeographic method that takes information from area cladograms and converts that information into a data matrix. The information includes the geographic distribution of individual taxa and the inferred geographic distribution of the ancestral nodes of the tree. A taxon's, or its ancestor's, presence in more than one area is treated as evidence that these areas once formed a continuous biota (and perhaps, but not necessarily, that the areas have a unique common history as geologic areas).

BPA was initially developed by recognizing that there is analogy between biogeography and phylogenetic analysis (Brooks et al., 1981; Brooks, 1981, 1985, 1990; Wiley, 1981a, 1988a; Wiley et al., 1991). In phylogenetic analysis, we seek to analyze the relationships of taxa using characters and a criterion. In BPA we seek to analyze the relationships of biotas or geologic areas using the relationships of taxa and a criterion. Phylogenetic biogeographic data consist of the organisms that occur in those areas and also how these organisms are related to one another on a phylogenetic tree hypothesis. Because not all of these data are likely to predict precisely the same set of area relationships, a criterion is needed to choose among the competing biogeographic data. For example, some parts of an area cladogram may indicate one set of areas shares a more recent relationship; another part of an area cladogram, or a different area cladogram, may indicate other areas share a more recent relationship.

In BPA, a parsimony criterion is used to decide among competing biogeographic hypotheses. After applying the algorithm, the resultant most parsimonious tree(s) is the tree showing the best supported pattern of biogeographic relationship among the areas. In effect, it is a tree of area relationships. The closer two areas are on a tree the more recently their component biotas shared a common history. Just as with phylogenetic analysis, an outgroup is needed to polarize the character data and see which characters represent the plesiomorphic state and which characters represent the apomorphic state. BPA employs an all zero outgroup or ancestral biogeographic region, which presumes that all taxa were primitively absent from the areas of interest.

The original version of BPA had been applied quite successfully. However, some authors (e.g., Morrone and Carpenter, 1994; Morrone and Crisci, 1995) criticized it because it sometimes yielded problematic results. For instance, in a phylogenetic study morphological characters support different parts of a resulting cladogram. In a biogeographic study using BPA, these characters are the taxa themselves and also their ancestral nodes. Sometimes, in traditional BPA, an ancestor and its descendants did not map on at the same node of the tree, and it was argued
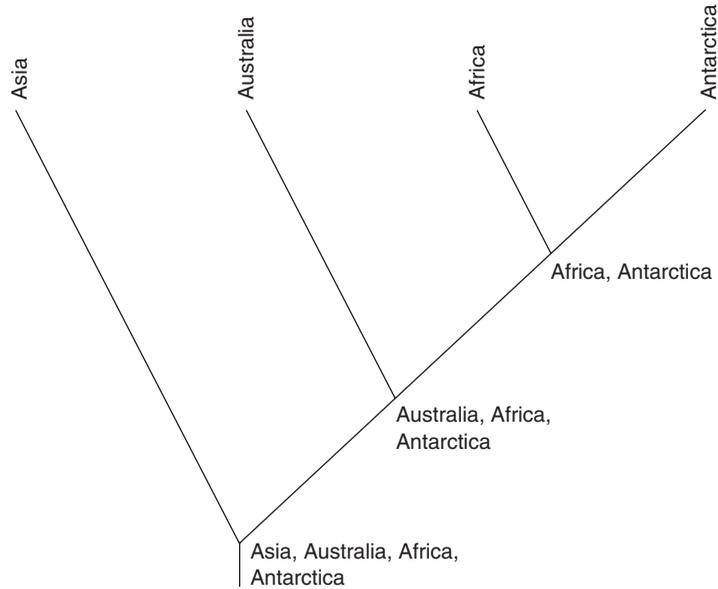
Figure 9.4. A hypothetical area cladogram that shows, in conjunction with Table 9.1, how to apply standard BPA. Redrawn from Lieberman (2000a).

TABLE 9.1. An example, based on Figure 9.4, illustrating how to apply standard BPA. The rows are the areas and the columns are the biogeographic data where columns 1, 3, and 5 are the three nodes moving up the tree and columns 2, 4, 6, and 7 are the terminal taxa moving from left to right.

	1	2	3	4	5	6	7
Outgroup	0	0	0	0	0	0	0
Asia	1	1	0	0	0	0	0
Australia	1	0	1	1	0	0	0
Africa	1	0	1	0	1	1	0
Antarctica	1	0	1	0	1	0	1

that this did not make much sense evolutionarily. It turns out that this was an artifact that arose because the original version of BPA did not use a parsimony algorithm to determine the state of the ancestral nodes of the tree (Lieberman, 2000a). Instead, the ancestral nodes of the tree were estimated by continually summing the distribution of all of the descendants; an approach called inclusive OR-ing (Brooks and McLennan, 1991; Wiley et al., 1991), see Fig. 9.4 and Table 9.1. For this reason, Lieberman and Eldredge (1996) modified traditional BPA and used a parsimony algorithm to infer the ancestral states of the tree. The choice of how to optimize characters was critical.

A variety of methods of ancestral character state reconstruction exist, running the gamut from parsimony-based approaches to maximum likelihood (ML) approaches (e.g., Farris, 1970; Fitch, 1971; Harvey and Pagel, 1991; Brooks and McLennan, 1991; Maddison and Maddison, 1992; Schultz et al., 1996; Pagel, 1999; Ree et al., 2005). Among the parsimony-based approaches used to reconstruct

biogeographic patterns within an individual clade, a very useful method is that first explored in detail by Mickevich (1981). In her method, each area represents one state of a multistate character. These characters can then be optimized to nodes using parsimony. For this particular type of biogeographic application, Ronquist (1994, 1995), Lieberman and Eldredge (1996), and Lieberman (1997, 2000a) argued that Fitch's (1971) unordered parsimony may be the best available approach to use, at least a priori. This is because it makes minimal assumptions about how taxa are moving between areas (W. P. Maddison, 1991). Given that one is usually interested in exploring how a set of taxa evolve across geographic space, it may not be prudent initially to bias the types of change that occur. However, if there is enough known about the biogeographic history of a region, it may be feasible to constrain patterns of biogeographic change, through character step matrices of the type discussed by Ree and Donoghue (1998). For instance, consider the case where a researcher was considering the biogeographic relationships of taxa in the Canadian Arctic, the Great Plains of the United States, and Central America. It might be reasonable to assume that to move between the Canadian Arctic and Central America an ancestor had to first pass through the Great Plains.

A second problem with the original version of BPA (and some more modern versions such as secondary BPA) was that it was only focused on recovering congruent episodes of vicariance and thus could not identify congruent episodes of geodispersal. Given that geodispersal has also left a significant imprint on the biogeographic history of biotas, a method that cannot account for geodispersal as signal is incomplete. Therefore, Lieberman and Eldredge (1996) further modified BPA so that it could be used to study congruent episodes both of vicariance and geodispersal (see also Lieberman, 1997, 2000a, 2003a, b, c). This was done by creating, for each analysis, two data matrices: one designed to retrieve evidence for congruent episodes of vicariance; the other designed to retrieve evidence for congruent episodes of geodispersal. This also addresses one of the other criticisms raised against BPA by Sober (1988) who argued that biogeography and phylogenetic analysis were not truly analogous. In particular Sober focused on the notion of dispersal and area amalgamation and argued that there really is nothing prohibiting area amalgamation or dispersal from occurring, in contradistinction to evolution where long distinct evolutionary lineages cannot re-anastomose (setting the case of horizontal gene transfer aside). By allowing for the possibility of congruent episodes of range expansion (geodispersal) to occur, MBPA more fully develops the potential to explore both the analogies, and the distinctions, between phylogenetic and biogeographic analysis. (This issue was also considered in our discussion above about the difference between geologic and biologic areas.)

In the next series of sections we present a formal account of MBPA as it applied to biogeography.

Overview of MBPA

MBPA consists of a series of steps that results in two trees that summarize the inferred histories of the biotas analyzed. One tree emphasizes shared vicariance events, if any; while the other emphasizes shared dispersal events, if any. Congruence among clades signaled by dichotomous relationships among biotas signals common vicariance or geodispersal contained in the relevant matrix. In Fig. 9.5 we summarize the flow of MBPA.
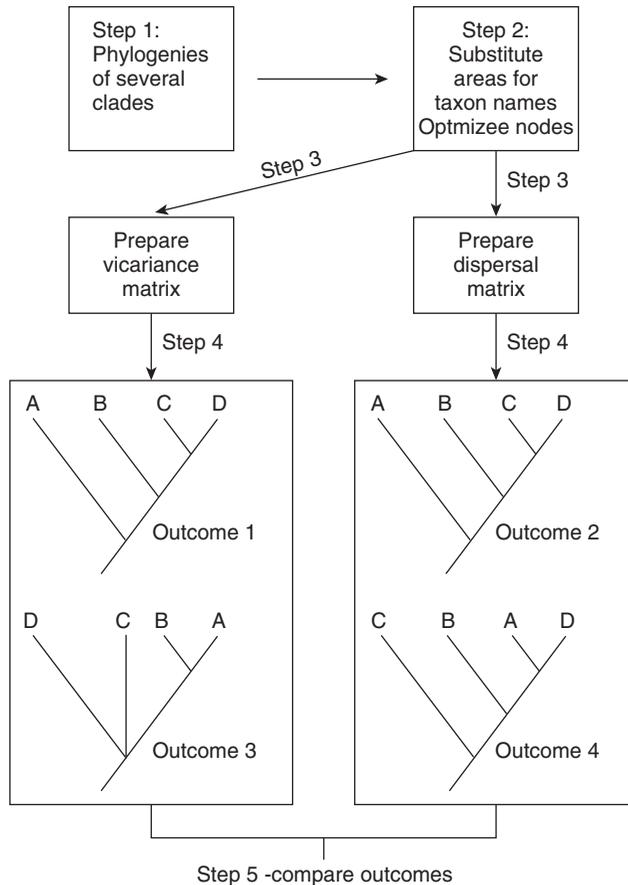
Figure 9.5. A flow chart showing how to apply mBPA.

Step 1 consists of a series of phylogenetic analyses of the candidate groups, each of whom are distributed among hypothetical biotic areas A, B, C, and D.

Step 2 consists of substituting the area of occurrence for each taxon and a Fitch optimization to obtain the inferred ancestral ranges.

Step 3 is the preparation of two matrices, one emphasizing vicariance, and the other dispersal.

Step 4 is the analysis of each matrix to obtain two hypotheses: a vicariance area tree and a geodispersal area tree.

Step 5 is an evaluation of the match of the two trees.

Figure 9.5 shows two possible results of the analysis. One result is congruence of the two area trees (outcomes 1 and 2). This would imply a dynamic of geodispersal followed by vicariance in a cyclic fashion as barriers fall and rise. The second shows a lack of congruence at two levels. In the vicariance matrix, there is only limited congruence among clades in the study; the relationships between clades can be explained by biotic vicariance only in areas A and B, but not in C and D. Further,
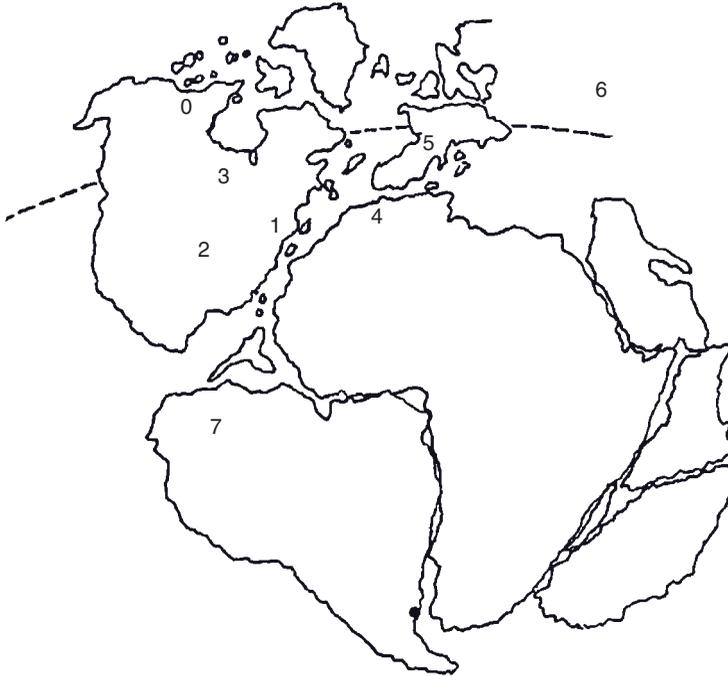
Figure 9.6. Areas of endemism for Devonian trilobites, from Lieberman and Eldredge (1996), showing the major continental blocs with the dashed line representing the paleo-equator. 0, the Canadian Arctic; 1, the Appalachian Basin of Eastern North America (ENA); 2, the Illinois basin of ENA; 3, the Michigan Basin of ENA; 4, North Africa; 5, Armorica; 6, Kazakhstan; 7, northern South America. From *Paleobiology*, used with permission of the Paleontological Society.

there is no congruence between the dispersal tree and the vicariance tree, implying that the patterns of vicariance are different from the patterns of geodispersal.

We will use an example of Devonian trilobites studied by Lieberman and Eldredge (1996) to step through MBPA. During the Devonian there was extensive tectonic collision during the early stages of the assembly of the supercontinent Pangea. Also, there were several major episodes of sea-level rise and fall related to climate change; each of these episodes of Earth history change seems to have led to repeated episodes of congruent vicariance and geodispersal. The patterns recovered in Devonian trilobites appear to also be reflected across a diverse array of other organisms including brachiopods, bivalves, and crustaceans (Rode and Lieberman, 2005). Our second example, presented only in the form of the resulting trees, concerns groups of Cambrian trilobites. Patterns in the Cambrian appear to be quite different from those in the Devonian, showing largely noncyclical patterns. Lieberman (2003b) and Meert and Lieberman (2004) applied MBPA to a series of phylogenies of Early Cambrian trilobites. They found well-resolved patterns of vicariance, but the consensus patterns of geodispersal were less well resolved. These biogeographic patterns match the general Earth history regime of the Cambrian, which was a time of continental fragmentation (Meert and Lieberman, 2004). The areas analyzed are shown in Fig. 9.6.
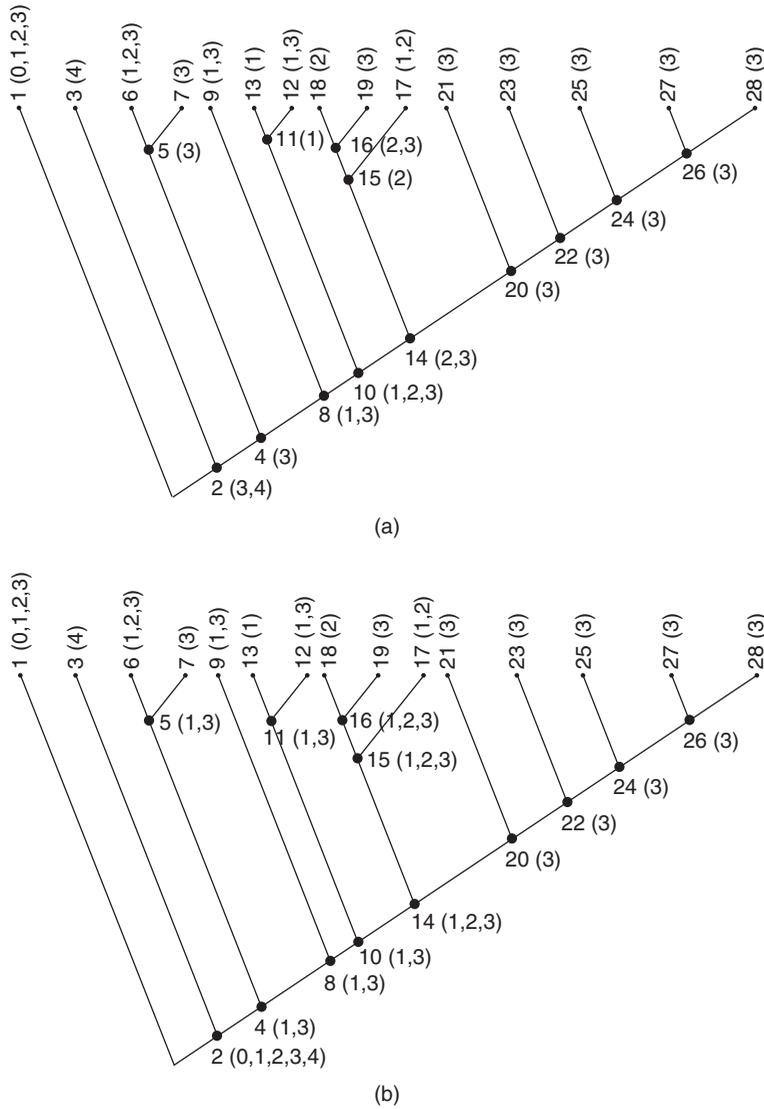
Figure 9.7. An example modified from Lieberman and Eldredge (1996) showing how the (a) downward pass and (b) upward pass optimization in the modified version of Fitch parsimony character states works, using the phylogeny for the Devonian trilobite genus *Basidechenella*. Modified and redrawn from Lieberman (1994).

Steps 1 and 2: Fitch Optimization of Area States on a Phylogeny   Lieberman (2000a) discusses the value of using a modified version of Fitch optimization for biogeographic studies. The usual implementation of Fitch optimization leaves polymorphisms at the tips but optimizes ancestral nodes with a single state. However, there is no reason to restrict the distributions of ancestors to single areas (Ronquist, 1997). Fitch (1971) recognized that this might result in unrealistic ancestral states and suggested the modification used here. Figure 9.7a and b show the steps of Fitch Optimization of areas onto the phylogeny of one of the clades of Devonian

trilobites, the genus *Basidechenella*, taken from Lieberman (2000a). The labeling is a bit confusing because we use numbers for both areas and taxa so that the trees and matrices correspond with Lieberman (2000a). A map of the areas with their numbers is shown in Fig. 9.6. As with all Fitch optimization, the final optimization consists of a two-pass algorithm. Figure 9.8 presents the trees for each of the four remaining clades analyzed by Lieberman (2000a) with the final optimization of area distributions of each node.

The down-pass is the standard Fitch method. It follows one of two rules depending on the state set of the children of a node. We use Fig. 9.7a for examples below.

(1) If two children of an ancestral node share one or more area states, then assign the ancestor the intersection of their state sets. For example, taxa 6 and 7 share area 3, thus:

$$5(\text{state set}) = 6(1,2,3) \cap 7(3) = 5(3).$$

(2) If two children of an ancestral node have different state sets, then assign the ancestral node the union of these state sets. For example, taxa 18 and 19 have different area state sets, thus:

$$16(\text{state set}) = 18(2) \cup 19(3) = 16(2,3).$$

The up-pass is a bit more complicated given that we will allow ancestral nodes to be polymorphic. We begin at the root of the tree, using taxon 1 as the root and proceed to visit each node, starting with node 2 (Fig. 9.7b).

(3) If the node does not have all of the area states of its ancestor, then go to step (5). If it does have all of the area states of its own ancestor, then go to step (4).

(4) If the node has more area states than its ancestor, delete those states from its state set, otherwise do nothing. Proceed to the next node up the tree.

(5) If the down-pass optimization of the node was the result of step (1), an intersection of area states, then go to step (6). If it was the result of step (2), a union, then add any states to the state set that are found in that node's ancestor. Proceed to the next node, and begin at step (3).

(6) Add to the state set of the node any state that meets both conditions: (a) it is present in the ancestor of the node and (b) it is present in at least one of the two children of that node. Proceed to the next node, and begin at step (3).

Some examples from Fig. 9.7 may be helpful. We will follow the up-pass optimization by considering selected nodes, and we will be switching between Fig. 9.7a and b in an iterative fashion as we move up the tree.

Node 2. In Fig. 9.7a the state set of the root 1 (0, 1, 2, 3) is larger than that of 2 (3, 4) in Fig. 9.7b. We proceed to step (5). We observe that 2 (3, 4) is the result of the union of 3 (4) and 4 (3). We add those states found in the ancestor of 2 to the state set of 2, resulting in 2 (0, 1, 2, 3, 4) in Fig. 9.7b, and proceed to the next node.

Node 4. In Fig. 9.7a the state set of 4 (3) is smaller than that of its ancestor 2 (0, 1, 2, 3, 4) in Fig. 9.7b, so we go to step (5). We note that the down-pass resulted in 4 (3), which was the intersection of 8 (1, 3) and 5 (3), so we proceed to step (6).
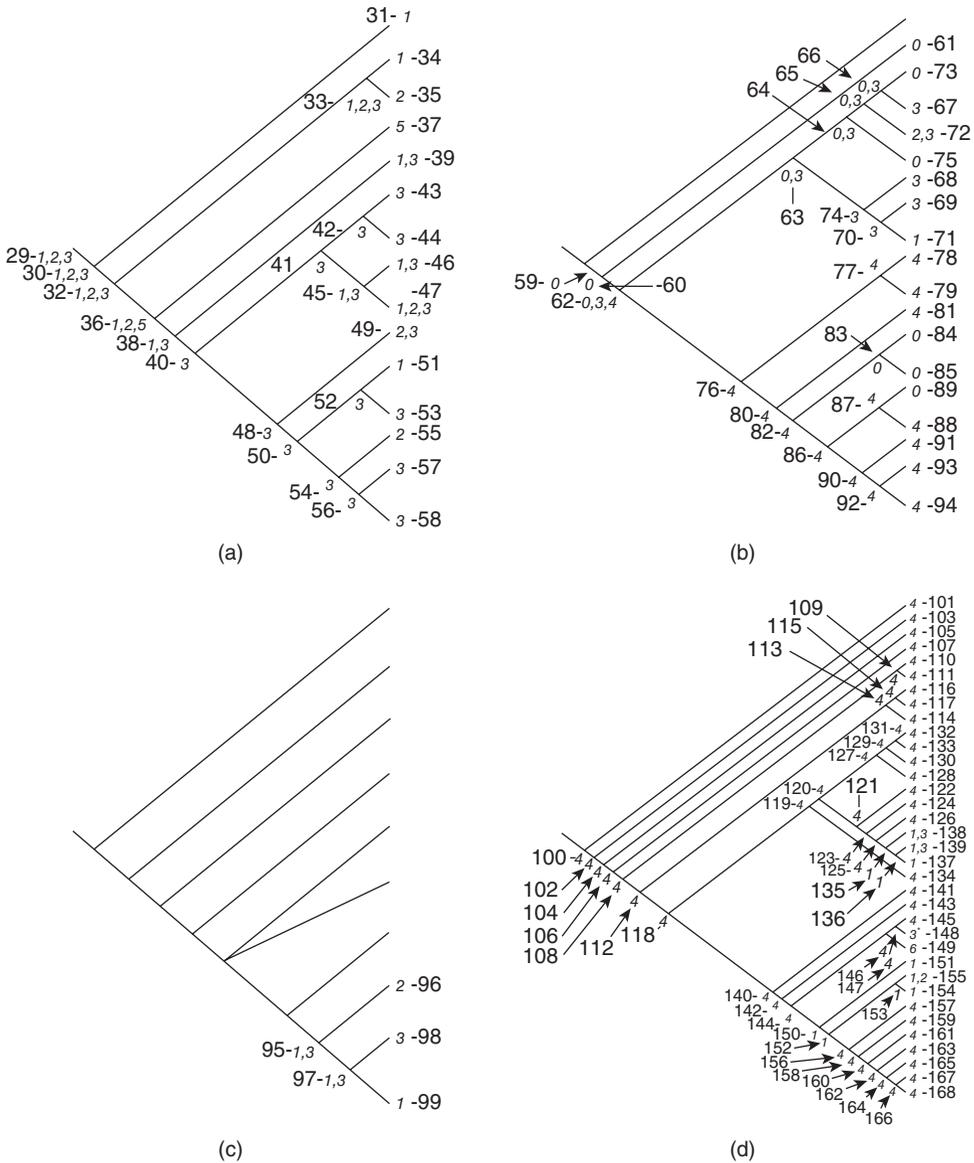
Figure 9.8. An example based on an analysis of phylogenetic biogeographic patterns of Devonian trilobites modified and redrawn from Lieberman and Eldredge (1996) that shows how to code the vicariance and geodispersal matrices in a modified BPA as described in the text. From *Paleobiology*, used with permission of the Paleontological Society.

Area 1 meets the condition of being in the ancestor of 4, 2 (0, 1, 2, 3, 4) and in one of two descendants, in this case descendant 8 (1, 3). Thus we add area 1 to the state set of node 4 (1, 3) in Fig. 9.7b, and proceed to the next node.

Node 8. In Fig. 9.7a node 8 (1, 3) has all of the areas of its ancestor 4 (1, 3) in Fig. 9.7b, thus we proceed to step (4). Because the state sets are identical, we retain node 8 (1, 3) in Fig. 9.7b, and proceed to the next node.

Node 10. In Fig. 9.7a node 10 (1, 2, 3) has all of the states of its ancestor 8 (1, 3) in Fig. 9.7b, so we proceed to step (4). We delete state 2 because it is not found in the ancestor of node 10, that is, node 8, see Fig. 9.7b, and proceed to the next node.

Although we have provided a narrative designed to show how polymorphic state sets are derived using Fitch Optimization, most computer packages will provide these optimizations directly. A complete account of the optimization of the remaining four clades is shown in Fig. 9.8a–d.

Area Distributions   The next steps in MBPA are to take the results of the Fitch optimization and translate the observed or inferred distributional patterns into two matrices: vicariance and dispersal. Coding of both matrices is built around five possible distributional patterns, shown in Fig. 9.9. We briefly discuss each, expressing the relationship between ancestral ranges and descendant ranges. These will be used to create rules for scoring each matrix.

> Pattern 1. Ancestor (A) and descendant (D) have the same range (Fig. 9.9a). The state set A = D.
>
> Pattern 2. Ancestral range is larger than descendant range (Fig. 9.9b). The state set A is a superset of D: A $(\supseteq)$ D.
>
> Pattern 3. Ancestral range is smaller than descendant range (Fig. 9.9c). The state set of A is a subset of D: A $(\subseteq)$ D.
>
> Pattern 4. The ranges of the ancestor and descendant overlap, but each is also found in one or more areas (Fig. 9.9d): $A \cap D \neq 0$, $A \cup D \neq A$, and $D \cup A \neq D$.
>
> Pattern 5. The ranges of the ancestor and descendant do not overlap (Fig. 9.9e): $A \cap D = \emptyset$.

Step 3.1: The Vicariance Matrix   The vicariance matrix emphasizes inferred vicariance events by weighting range contractions inferred to have occurred between an ancestor and a descendant. This range contraction signals a vicariance event and the conclusion that entire parts of the biota participated in such a vicariance event would require congruent range contractions of several clades included in the study. Coding the vicariance matrix follows two rules based on the distributional patterns.

Rule 1. Score "1" for all areas derived from the union of the ancestral and descendant state sets and "0" for the complement of that union.

$$T1 = D \cup A$$

$$T0 = \sim(D \cup A)$$

Rule 2. Score "2" for all areas shared by the ancestor and descendant. Score "1" for all areas of the ancestor not found in the descendant. Score "0" for all areas not found in the ancestor or descendant.

$$T2 = D \cap A$$

$$T1 = A \cap (\sim D)$$

$$T0 = \sim(A \cup D)$$
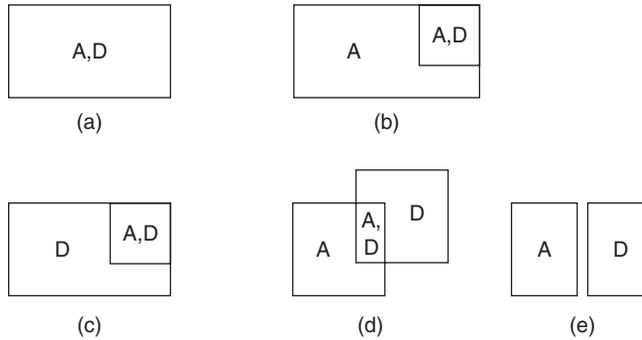
Figure 9.9. The five possible area relationships that can exist between two taxa. (a) Ancestor (A) and descendant (D) have the same range. (b) Ancestral range is larger than descendant range. (c) Ancestral range is smaller than descendant range. (d) The ranges of ancestor and descendant overlap, but each is also found in one or more different regions. (e) The ranges of ancestor and descendant are different.

Relative to the distributional patterns discussed above, we provide some examples. The full vicariance matrix is available for download at: http//:paleo.ku.edu/geo/lieberman.html; select the "Phylogenetics" book icon.

Pattern 1. Invoke Rule 1, see Fig. 9.7b. Ancestral node 4 and descendant node 8 provide an example. Their state sets are identical. Thus we code 8 with the vector (0101000).

Pattern 2. Invoke Rule 2, see Fig. 9.7b. Ancestral node 2 and descendant node 4 provide an example. The ancestor, node 2, has a larger range than its descendant, node 4. For descendant node 4, we score areas 1 and 3 with state "2," areas 0, 1, 2, and 4 with state "1" and areas 5 and 6 with state "0," creating the state vector 4 (12121000).

Pattern 3. Invoke Rule 1, see Fig. 9.7b. Node 6 and its ancestor node 5 provide an example. The descendant, node 6 (1, 2, 3), has a larger range than its ancestor, node 5 (1, 3). We score areas 1, 2, and 3 as "1" and all other areas as "0," creating the state vector 6 (0111000).

Pattern 4. Invoke Rule 2, see Fig. 9.8a. Node 38 (1, 3) and its ancestor node 36 (1, 2, 5) provide an example. For node 38 we score "2" for area 1; "1" for areas 2, 3, and 5; and "0" for areas 0, 4, and 6, creating the vector 38 (0211010). Note: the original Lieberman coding scored area 3 with the score of "2," but we have changed our interpretation now as the inference is that node 38 dispersed into area 3. This dispersal pattern will be picked up in the dispersal matrix.

Pattern 5. Invoke Rule 1, see Fig. 9.8b. Node 71 (1) and its ancestor node 70 (3) are an example. The union of 71 (1) and 70 (3) results in a score for 71 of (0101000). This differs from the original coding of Lieberman (2000a) who scored such patterns as autapomorphies.

Step 3.2: The Dispersal Matrix   The dispersal matrix emphasizes inferred dispersal events by weighting the expansion of a descendant range relative to the range of

its ancestor. The conclusion that range expansion represents a geodispersal event would require congruent range expansions among several clades included in the study. Coding the dispersal matrix follows a single rule.

Rule 3. Dispersal matrix only. Score "2" for any area of the descendant not occupied by the ancestor, score "1" for any area shared by the ancestor and descendant, and score "0" for all other areas not occupied by either the ancestor or descendant.

$$T2 = D - (A)$$

$$T1 = A$$

$$T0 = \sim(A \cup D)$$

Dispersal may be encountered in patterns 3, 4, and 5. We present examples of each below. The full dispersal matrix is available for download at: http//:paleo.ku.edu/geo/lieberman.html; select the "Phylogenetics" book icon.

Pattern 3. Invoke Rule 3, see Fig. 9.8a. Node 45 (1, 3) and its ancestor node 41 (3) provide an example. Score area 1 as "2" and area 3 as "1" while all other areas receive a score of "0." The coding for 45 is (0, 2, 0, 1, 0, 0, 0).

Pattern 4. Invoke Rule 3, see Fig. 9.8a. Node 36 (1, 2, 5) and its ancestor node 32 (1, 2, 3) are an example. Score area 5 as "2"; score areas 1, 2, and 3 as "1"; and score all other areas as "0." The coding for 36 is (0, 1, 1, 1, 0, 2, 0).

Pattern 5. Invoke Rule 3, Fig. 9.8b. Nodes 71 (1) and its ancestor node 70 (3) provide a contrasting example to the way vicariance is scored. We score area 1 as "2," area 3 as "1," and all other areas as "0." The coding for node 71 is (0, 2, 0, 1, 0, 0, 0).

Steps 4 and 5: MBPA Analyses and Comparison   Analysis of each matrix is performed. We have chosen a parsimony approach for reasons discussed, but one might elect to perform a likelihood analysis. Below we discuss some of the results of the analysis of these clades of trilobites (Fig. 9.10) as well as another example of the analysis of some clades of Cambrian trilobites (Fig. 9.11).

The results from the analysis of each matrix is presented as a most parsimonious tree(s) representing the patterns of vicariance best supported by the available biogeographic data and the patterns of geodispersal best supported by the available biogeographic data. The vicariance tree provides information about the relative times at which regions became isolated due to the formation of barriers that separated regions and isolated their respective biotas, ultimately leading to evolutionary differentiation. For example, the vicariant area cladogram shown in Fig. 9.11 suggests that the barriers separating southwestern Laurentia and Siberia formed more recently than the barriers separating southwestern Laurentia and Baltica. Thus, southwestern Laurentian and Siberian biotas share a more recent evolutionary history. By contrast, the geodispersal tree provides information about the relative times at which regions became joined as barriers between regions fell and respective biotas merged, ultimately leading to biotic mixing. Thus, the geodispersal area cladogram shown in Fig. 9.11 suggests that the barriers separating eastern Laurentia and
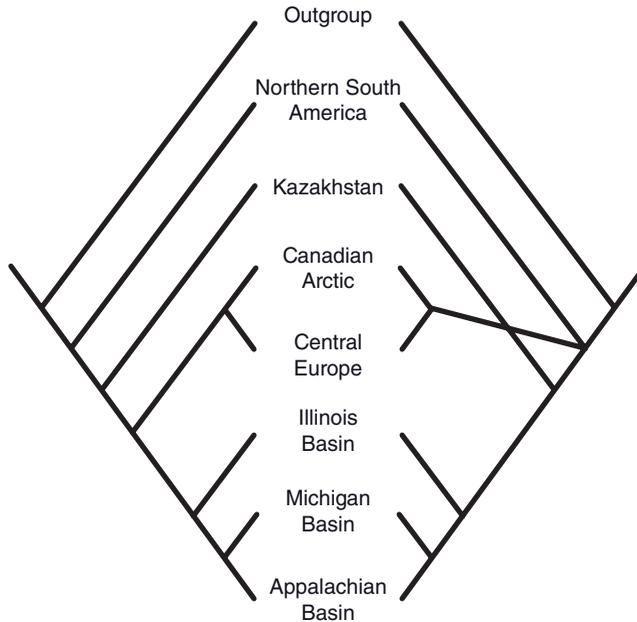
Figure 9.10. Results from biogeographic analysis of Devonian trilobites (~390–370 million years old) using the modified version of BPA described in the text, from Lieberman and Eldredge (1996). Areas considered constitute major sites of endemism in the Devonian and are shown in Fig. 9.6. On the left are the most parsimonious patterns of vicariance and on the right the most parsimonious patterns of geodispersal. Notice both are well resolved and imply very similar biogeographic patterns. From *Paleobiology*, used with permission of the Paleontological Society.

northwestern Laurentia fell more recently than the barriers separating these parts of Laurentia from Baltica: eastern and northwestern Laurentian biotas were more recently homogenized than Laurentian and Baltic biotas.

After the vicariance and geodispersal trees are generated by MBPA, they can also be compared with one another. This provides additional information about the processes that may have been most responsible for producing the biogeographic patterns (Lieberman and Eldredge, 1996; Lieberman, 2000a, 2003a, c, 2005). For instance, if the two trees are very similar (Fig. 9.10), it suggests that the same processes that produced vicariance may also have produced geodispersal. In the case of marine taxa, this might involve repeated episodes of sea-level rise and fall that joined and later sundered populations. Sea-level rise and fall could also influence biogeographic patterns in terrestrial taxa, albeit in an opposing manner, for example, by flooding and then uncovering spits of land that might isolate and then join different populations. By contrast, if the most parsimonious vicariance and geodispersal trees are different (Fig. 9.11) (note: we use the term *different*, not *incongruent*), it suggests that processes not cyclical, at least on a time scale commensurate with speciation, played the primary role in shaping biogeographic and evolutionary patterns (Lieberman and Eldredge, 1996; Lieberman, 2000, 2003a, c, 2005). Such events include single tectonic events like a collision between continents, or a chance
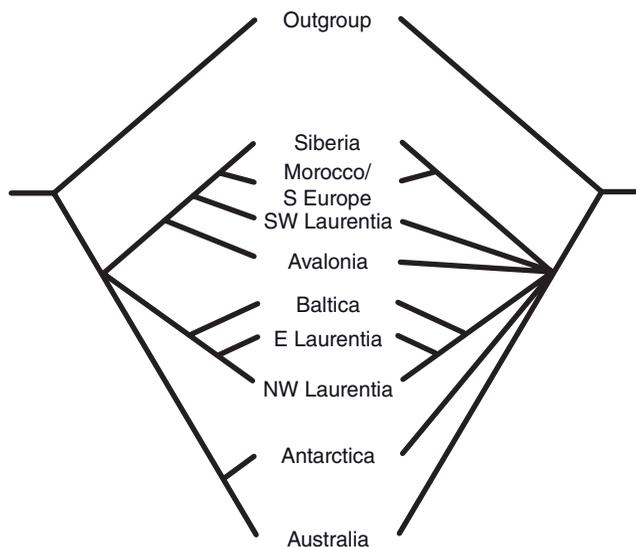
Figure 9.11. Results from biogeographic analysis of Early Cambrian trilobites (~525–510 million years old) using the modified version of BPA described in the text, from Lieberman (2003b) and Meert and Lieberman (2004). Areas considered constitute major sites of endemism in the Early Cambrian. On the left are the most parsimonious patterns of vicariance and on the right a strict consensus of the most parsimonious patterns of geodispersal; the former show more resolution and make more concrete predictions about biogeographic patterns. From Lieberman (2003c).

long-distance dispersal event, or a change in a drainage pattern or rise of a mountain range.

The Devonian pattern, showing cyclic vicariance/geodispersal, is also found in several other groups of organisms (Rode and Lieberman, 2005), and this was a time of pervasive sea-level rise and fall and continental collision. Lieberman's (2003b) and Meert and Lieberman's (2004) analyses of Early Cambrian trilobites produced a quite different result (Fig. 9.11). These biogeographic patterns match the general Earth history regime of the Cambrian, which was a time of continental fragmentation (Meert and Lieberman, 2004).

Another important difference between Cambrian and Devonian trilobites are their rates of speciation. Rates of speciation in trilobites were relatively high during the Cambrian, associated with the initial proliferation of the clade during the Cambrian radiation; by contrast, Devonian trilobites show more muted rates of speciation (Lieberman, 2003a, b, c). These differences make sense given that during the Cambrian there were many opportunities for geographic isolation and vicariance associated with the breakup of the supercontinent Pannotia. By contrast, during the Devonian there were abundant opportunities for geodispersal. More geodispersal means fewer opportunities for vicariance and thus reduced rates of speciation (Lieberman, 2003b, c; Meert and Lieberman, 2004; Rode and Lieberman, 2004, 2005).

Hembree (2006) provided another example of the use of MBPA. Focusing on the amphisbaenians, a group of burrowing limbless lizards, and incorporating information from both fossil and extant members of the group, he retrieved well-resolved patterns of geodispersal, which is quite interesting because they suggest that even burrowing vertebrates may at times undergo congruent episodes of geodispersal.

ALTERNATIVE BIOGEOGRAPHIC METHODS

Brooks parsimony analysis (BPA). The method is described in detail in Brooks et al. (1981), Brooks (1985, 1988, 1990), Wiley (1988a, b), Brooks and McLennan (1991, 2002), Wiley et al. (1991), Lieberman and Eldredge (1996), Lieberman (1997, 2000a, 2003), and Morrone (2008). As mentioned above, BPA is a biogeographic method that takes information from area cladograms and converts that information into a data matrix. The information includes the geographic distribution of individual taxa and the inferred geographic distribution of the ancestral nodes of the tree. Brooks et al. (1981), in a classic and landmark study, used BPA to consider evolutionary and biogeographic patterns in the freshwater stingrays of South America. They concentrated on the biogeography of worms parasitic on the stingrays. Brooks et al. (1981) converted several phylogenies of these parasitic worms into area cladograms and then subjected them to BPA and used this to consider the geological evolution of South America and the co-evolution of the continent's biota (see also Brooks and McLennan, 1991); they found a pattern of vicariance stretching across much of the continent.

Another interesting application of BPA was Mayden's (1988b) analysis of biogeographic patterns in North American freshwater fishes; this is an endemic and very diverse fauna containing perhaps hundreds of species. Mayden (1988b) applied BPA to phylogenies of seven clades of fishes that inhabit major modern river drainage systems in eastern North America. Mayden (1988b) identified an interesting pattern in his biogeographic cladogram: the grouping of rivers differed from the topology of extant river drainages and instead showed greatest similarity to the configuration of pre–Pleistocene river drainages. Brooks and McLennan (1991, 2002) and the references therein document in detail many other studies that have successfully used BPA to study phylogenetic biogeographic patterns.

Phylogenetic Analysis for Comparing Trees (PACT). Wojcicki and Brooks (2005) developed PACT as a new method of biogeographic analysis (see also Brooks and Folinsbee, 2005). This method does not require generation of a data matrix from area cladograms, *contra* BPA, but instead compares the various area cladograms and, in this respect, is somewhat more akin to component analysis (discussed below). However, it differs fundamentally from component analysis because it treats the original data as real (follows assumption 0, discussed more fully below, *sensu* Wiley, 1987, and Zandee and Roos, 1987) and thus does not use assumptions 1 and 2 to modify the input area cladograms. Finally, PACT does not presuppose that all diversification follows the vicariance model and instead allows diversification to be driven by range expansion (traditional dispersal and geodispersal). The method is based on the fact that any tree can be expressed as a set of taxon names and

parentheses that can be represented as a Venn diagram. Using an initial input tree, PACT builds a template tree from this and compares where it agrees and disagrees with other area cladograms; there are various rules used to combine different area cladograms. Thus far, this method has not been applied to many examples, but it appears to have substantial potential. Notably, in common with MBPA, it focuses on identifying overall biogeographic congruence without presupposing that patterns must have been generated solely by vicariance. It will produce results different from BPA or MBPA because individual areas can appear multiple times on the area cladogram that results from synthesizing the various individual area cladograms.

Component Analysis. Platnick and Nelson (1978) and Nelson and Platnick (1981) developed an analytical biogeographic method that was similar to a method presented by Rosen (1978, 1979). Area cladograms for different groups will often differ (hence the need for analytical biogeographic methods) and also do not fully match a simple pattern of vicariance. For example, some area cladograms will not have all areas represented and some area cladograms will have the same area represented multiple times, etc. What component analysis does is compare the different parts or "components" (hence the name) of the tree and attempts to explain or account for the differences between various parts of the area cladograms in order to come up with a consensus or general area cladogram. Component analysis does this methodologically by creating new area cladograms that are formed by adding certain branches to, or breaking certain branches of, the original area cladograms to make them match and to make every part of the tree match a strict pattern of vicariance. In certain cases, many new branches need to be added to each area cladogram. Each branch is held to represent a real taxon in a real area, but that taxon would not have been recovered either because of error by the biologist(s) studying the group or because of its extinction but nonpreservation (or nondiscovery) as a fossil. Where and how many areas are added (by actual addition or by breaking apart combined areas) depends on implementing certain assumptions, called assumption 1 and assumption 2 (Nelson and Platnick, 1981; Morrone and Carpenter, 1994; Enghoff, 1995; Morrone and Crisci, 1995; Lieberman, 2000a; Morrone, 2008).

Some aspects of component analysis have been criticized by Wiley (1988a, b), Brooks and McLennan (1991), and Lieberman (2000a), who question its applicability as a general biogeographic method. In particular, assumptions are made about the correctness of some but not all aspects of the data represented in the area cladograms: only those data disagreeing with strict vicariance explanations are modified. There is no justification, however, for assuming that only these aspects of the area cladogram are incorrect; the other data compatible with vicariance may also be incorrect. Thus, component analysis is at odds with one of the fundamental principles of phylogenetic systematics dating back to Hennig (1966), who argued that any character shared between taxa should initially be treated as indicating common descent. In a sense, this principle implies that data should at first blush be treated as real (Wiley, 1988a, b; Lieberman, 2000a). This notion is sometimes referred to as assumption 0, a basic assumption that the data are valid in the first place (Wiley, 1987, 1988a, b; Zandee and Roos, 1987). Component analysis does not rely solely on assumption 0 and instead introduces biogeographic data that are not part of the original data, through assumptions 1 and 2, to render the results

compatible with a strict vicariance pattern. The method not only significantly modifies biogeographic patterns without factual data (factual data being the data used to construct the phylogenetic relationships in the first place), but it also can never recover a pattern at odds with strict vicariance. For instance, geodispersal is a biogeographic process that can never be retrieved by component analysis. Widespread taxa that have failed to speciate in the face of more recently formed geographic barriers also create undue noise in component analysis. Thus, the inability to recognize geodispersal, undue noise due to failure to speciate, and the unjustified modification of biogeographic data, seem to represent insurmountable problems for component analysis.

A final problem with component analysis is the fact that it is a consensus technique, and such techniques cannot produce a more parsimonious result (relative to the original trees) and often produce results that are nonparsimonious (Miyamoto, 1985; Barrett et al., 1991; Kluge, 1988; Wiley, 1988a, b). Moreover, by comparing different area cladograms with one another and trying to get their various components to square with one another, the method implicitly treats the biogeographic signal from each cladogram as equivalent, and that may not be the case (Lieberman, 2000a).

Because of these differences, we cannot endorse the use of component analysis as a general biogeographic method. By the same token, however, we recognize that component analysis shares many features in common with methods such as BPA and MBPA. In particular, all of these methods involve searching for congruence among various phylogenies. They principally differ in the way they assess this congruence and also the extent to which they use information derived from evolution. Van Veller et al. (1999, 2000, 2001, 2002) and Van Veller and Brooks (2001) usefully distinguished between phylogenetic and cladistic biogeographic methods; the former include those based on some form of BPA and have a more evolutionary focus, using information from the distribution of ancestral taxa and considering how these distributions change as taxa evolve. By contrast, cladistic biogeographic methods, including component analysis, do not take such information into account. Van Veller et al. (1999, 2000, 2001, 2002) and Van Veller and Brooks (2001) referred to component analysis as an a priori method because certain assumptions are made about the input data, and these data are subsequently modified based on these assumptions. By contrast, they referred to methods related to BPA as a posteriori because they do not allow the input phylogenetic and geographic data to be modified; instead, they may explain any differences between the overall area cladogram and individual area cladograms after the fact. Not all methods that have been classified as "phylogenetic biogeographic" employ some form of BPA (Ebach et al., 2003; Morrone, 2005). For instance, Hennig (1966) and Brundin (1966, 1988) conducted biogeographic analyses that used phylogenies but their studies did not employ matrices to compare the results from different area cladograms.

Dispersal Vicariance Analysis (DIVA). Ronquist (1996, 1997) developed DIVA (see also Zink et al., 2000; Whitcher and Wen, 2001; Bremer, 2002; Drovetski, 2003; Chatterjee, 2006; Gómez and Lobo, 2006), which seeks to explain a clade's evolution by invoking vicariance, plus some amount of dispersal and extinction. The latter are treated as noise to be minimized. We return to the example involving North America and China mentioned earlier in the chapter. DIVA was used by Xiang and Soltis

(2001) to investigate several angiosperm lineages in the Northern Hemisphere that are currently disjunct. No reasonable person would argue that these areas share a unique vicariance history within the last 550 million years. However, it is entirely reasonable to suspect, after studying their plant biotas, that these might be vicariant (or partly vicariant) remnants of a widespread Oligocene and Miocene tropical boreal forest that covered most of the Holarctic. (We suggested earlier that in the case of something like the Oligocene-Miocene tropical boreal forest the units of analysis appearing on such a cladogram are biotas, not areas.) Xiang and Soltis (2001) found elements of both vicariance and dispersal (within the widespread Northern Hemisphere plants) using DIVA (Ronquist, 1997) and analyzing groups one at a time. However, there is no reason to think that analyzing component clades one at a time is a requirement. Ronquist (1997) has suggested that cost–benefit analysis can be applied to area cladograms to assess the fit of multiple clade analysis to a single solution. However useful Ronquist's DIVA analysis is, it still treats dispersal and extinction as noise, and when it comes to certain types of dispersal this can be a problem. It is true that some types of dispersal create noise for studies seeking to unravel the history of geological areas and biotas, in particular, individual episodes of uncoordinated dispersal we have already described (the presence of African Cattle Egrets in the Americas is such an example). However, the dispersal of entire biotas (geodispersal) is congruent biogeographic signal that would be treated as noise by DIVA (but not by MBPA).

Event-Based Models. Ronquist (1998b, 2002) elucidated another method in addition to DIVA that can be used to consider biogeographic patterns; this method is also described in detail in Sanmartin and Ronquist (2004). They argued that it might not be appropriate to apply DIVA in cases where the amount of vicariance significantly exceeds dispersal. Instead, they suggested a parsimony-based tree fitting approach, in conjunction with permutation tests, is most apt. This method allows four events to occur: vicariance, duplication or speciation within an area, dispersal, and extinction. It is implemented using the program TreeFitter (Ronquist, 2002). Each event is associated with a cost, and then the organism phylogeny, along with the geographic distributions, are fit to an area cladogram (Sanmartin and Ronquist, 2004); in a sense, the method is based on mapping a phylogenetic tree to an underlying area cladogram. The area cladogram with the minimum cost, that is most parsimonious, is favored. Setting the costs is, of course, crucial when applying this method (Sanmartin and Ronquist, 2004); thus far, what has been done is to use hypothetical data sets to find the event costs that seemed to perform well under a wide range of conditions. In particular, based on such hypothetical data sets Sanmartin and Ronquist (2004) argued that the cost for vicariance and duplication should be 0.01, for extinction it should be 1.0, and for dispersal it should be 2.0; whether this means that vicariance is 10 times more likely than extinction, and 20 times more likely than dispersal is difficult to evaluate. This biogeographic method also experiences difficulty dealing with phylogenies that have widespread taxa (Sanmartin and Ronquist, 2004). When these are present, it must be assumed that either there was recent dispersal, vicariance failed to occur, or there was an anything-goes strategy. In a sense, the way widespread taxa are treated by this method is very similar to how they are treated in component analysis with assumptions 1 and 2. However, this method does have an advantage over component analysis because it allows for the possibility of dispersal. If there is evidence for dispersal between the same areas in

several clades, then there is the potential to identify geodispersal (e.g., Sanmartin and Ronquist, 2004).

Parsimony Analysis of Endemicity (PAE). PAE was developed by Rosen (1988; see also Morrone, 1994, 2005, and Morrone and Crisci, 1995) and has become one of the most frequently applied biogeographic methods. For example, it has been used in a range of biogeographic analyses on a great variety of fossil and extant taxa (e.g., Myers, 1991; Morrone, 1998; Glasby and Alvarez, 1999; Waggoner, 1999; Bisconti et al., 2001; De Grave, 2001; Morrone and Márquez, 2001; Morrone and Escalante, 2002; Aguilar-Aguilar et al., 2003; Lieberman, 2004; Huidobro et al., 2006; Quijano-Abril et al., 2006; Vargas et al., 2007). The method is not technically phylogenetic because it does not incorporate information from phylogenies; instead, the focus is on comparing different regions to see if they share different taxa (species or genera typically). A data matrix is generated with the rows comprising the different regions being studied; the matrix would be akin to a data matrix created for a phylogenetic analysis or a BPA. The taxa become the different characters of the data matrix: the presence of a taxon in a region is scored as a "1" in the data matrix; the absence of a taxon is scored as a "0." An all "0" outgroup is used to polarize the character data; then the matrix is analyzed using a parsimony algorithm, with the results depicted as a general area cladogram. (What the method shares in common with phylogenetics is of course the use of a parsimony algorithm and the generation of a character–taxon [in this case area] matrix.)

The advantage of the method is that it does not require input phylogenies; generating phylogenies, of course, requires a significant time commitment. All that is required is distributional data, which is much easier to obtain. Indeed, this is why the method has been so frequently applied, yet by the same token this is why the method should not be classified as a phylogenetic biogeographic technique in the broad sense. Further, the absence of phylogenetic information is a problem: ultimately the method cannot be used to distinguish whether shared patterns of distribution have resulted from vicariance or geodispersal. Because of these and other difficulties, Brooks and Van Veller (2003) have criticized the general use of PAE; although Nihei (2006) more broadly endorsed the method, he suggested that the method needed to be modified in various ways. It is likely, because of its relative ease of use, that PAE will continue to be employed for some time, and it may be useful as a general first pass biogeographic method. Further, the method clearly has advantages over phenetic biogeographic analyses because it does not conflate plesiomorphic as opposed to apomorphic distributions of taxa, nor does it assign any weight to areas that have a large number of unique (autapomorphous) taxa.

HOW EXTINCTION AFFECTS OUR ABILITY TO STUDY BIOGEOGRAPHIC PATTERNS IN THE EXTANT BIOTA

It is estimated that more than 99.9 percent of all the species that have ever lived on this planet are extinct. The net effect is that the extant biota is a dramatically pruned sample of total diversity. Given that biogeography focuses on the coevolution of the Earth and its biota, the effects of extinction are worth examining in the context of how they affect our ability to retrieve accurate biogeographic patterns. Gauthier et al. (1988), Donoghue et al. (1989), and Wheeler (1992) have demonstrated that

extinction affects phylogenetic studies of extant organisms because it determines which taxa can be studied by neontologists. Incorporating additional taxa into phylogenetic studies, especially fossil taxa with unique character states, can increase phylogenetic accuracy. Given that biogeographic results depend on available phylogenies, this is clearly one manner in which extinction can affect our ability to retrieve accurate biogeographic patterns; it further clearly indicates that, whenever possible, extinct fossil taxa should be incorporated into the phylogenetic studies that are used for subsequent biogeographic analyses.

The effects of extinction on biogeography are not solely confined to their impact on phylogenetic results; they in fact decrease our ability to detect biogeographic congruence among representatives of the extant biota (Lieberman, 2000a, 2002b; Turner et al., 2009). The problem that biogeographic studies of extant taxa face is basically the same problem that paleontologists face because the fossil record is incomplete. An illustration of biogeographic congruence in an idealized two-clade case is shown in Fig. 9.12. Imagine that some of these taxa were extinct: for example, taxon C in the clade on the left-hand side of Fig. 9.12 and taxon B in the clade on the right-hand side of Fig. 9.12. A neontologist who only sampled living taxa would retrieve the pattern in Fig. 9.13. Lieberman (2002b) termed such a pattern artificially incongruent, because it does not reflect true biogeographic incongruence that arises from processes such as traditional dispersal and sympatric speciation; instead, it reflects an underlying sampling issue. Lieberman (2002b) focused on how such an artificially incongruent result would impact a biogeographic study based on BPA. Both area cladograms in Fig. 9.12 predict the relationship (A(B(C(D)))). By contrast, in Fig. 9.13 the area cladogram on the left, when scored in a BPA along with the cladogram on the right, would predict the relationship (C(A(B(D)))), while the one on the right would predict the relationship (B(A(C(D)))). (These two area cladograms in Fig. 9.13 may or may not be incongruent in a biogeographic perspective based on component analysis, depending on which different assumptions are used.) A similar situation involving artificial incongruence could of course arise with fossil taxa. Imagine that taxon C in the clade on the left-hand side of Fig. 9.12 and taxon B in the clade on the right-hand side of Fig. 9.12 were rare and thus not preserved in the fossil record.

It is important to note that simply the existence of *some* artificial incongruence by itself does not obviate the value of biogeographic studies. However, it is conceiv-
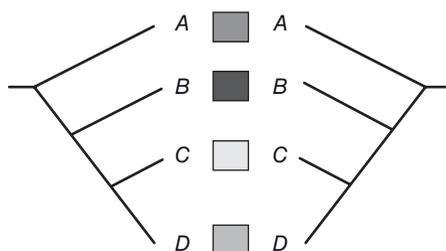
Figure 9.12. Biogeographic congruence in the case of two hypothetical clades where the shaded boxes denote the areas of geographic occurrence and the letters represent different terminal taxa, from Lieberman (2002b). From *Palaeogeography*, *Palaeoclimatology*, and *Palaeoecology*, used with permission of Elsevier.
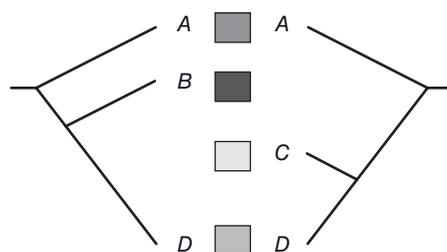
Figure 9.13. An example showing how extinction, in the case of the extant biota, or paleontological incompleteness, in the case of the fossil record, can cause artifactual biogeographic incongruence. In particular, consider that certain species present in certain geographic areas from the cladograms shown in Fig. 9.12 could not be recovered. In the case of a study based on extant organisms, this would be because these taxa were extinct; in the case of a study based on fossil organisms, it would be because these species were not preserved in the fossil record. The result is the same, apparent biogeographic incongruence that arises as a simple artifact, from Lieberman (2002b). From *Palaeogeography*, *Palaeoclimatology*, and *Palaeoecology*, used with permission of Elsevier.

able that if the artificial incongruence is too high in a biogeographic study, it might create noise and mask whatever signal is present in the data. At this time, we do not know what *too* high connotes, but Lieberman (2002b) posited that it could be problematic when more than half of the data appeared artificially incongruent due to sampling issues.

Lieberman (2002b) performed a series of simulation studies to better quantify the effects of artificial incongruence and see how extinction affects our ability to retrieve accurate biogeographic patterns in the extant biota. He used a computer program to evolve clades distributed across various areas of endemism; these were subjected to various extinction and speciation probabilities through time. Then the biogeographic patterns in these pruned clades were examined at a particular time slice, which was treated as the modern. Results suggest that artificial biogeographic incongruence increases as the age of a clade increases, and also as the extinction probability in that clade increases (even in the face of concomitantly climbing speciation rates). Of course caveats must be raised with any simulation study, and this one is no exception. Simulation studies are most valuable, not as a gauge of reality, but when they show how one variable may influence patterns as others are held fixed (a situation rarely possible in the real world). Based on these simulations, it would appear that clades with moderate to high extinction rates should probably be avoided in biogeographic studies that only consider the extant biota: for instance, clades with moderate or high extinction probabilities accumulated very high levels of artificial incongruence within a few tens of millions of years. We also note that clades with high speciation rates tend to have high extinction rates (Eldredge, 1979; Stanley, 1979, 1990; Vrba, 1980). Therefore, rapidly evolving groups may not always make the best subjects for biogeographic analyses. In particular, they will be most appropriate when they are quite young geologically. Artificial biogeographic incongruence also accumulates with clade duration, such that clades with a Mesozoic origin have the potential for showing significant artificial biogeographic incongruence (Lieberman, 2002b). This is especially true when extinction rates are high, but

it is even true when extinction rates (in the face of high speciation rates) are moderate. This may mean that extant taxa should only be used as candidates to study relatively ancient biogeographic events, like the break up of Pangea, if their extinction rates are quite low. In fact, in clades with high to moderate rates of extinction, simulations suggest that it may be hard to retrieve meaningful biogeographic patterns even if these clades originated only a few tens of millions of years ago. The results also suggest that it will be quite valuable, whenever possible, to incorporate fossil taxa into biogeographic studies. These fossil taxa will serve to partially obviate the potential issue of artificial incongruence by increasing sampling.

The importance of conducting biogeographic analyses on clades that originated relatively recently matches the predictions and recommendations of Wiley and Mayden (1985) and Brooks and McLennan (1991). In particular, they argued that the best subjects for biogeographic analysis were those clades that had experienced limited extinction (Lieberman, 2002b).

The parameters that determine our ability to study minimally artificially incongruent biogeographic patterns in the fossil record are of course not extinction probability and clade age; most fossil taxa are extinct and old. Instead, what matters for the recovery and sampling of fossil taxa is the completeness of the fossil record and the probability that individual species will be preserved (Lieberman, 2002b). Estimates on preservation probability are more difficult to derive than parameters like extinction rates, but a stab has been taken at deriving such values. It is not surprising that artificial incongruence rises as preservation probability falls. It would appear that preservation probability for good candidate groups to consider in paleobiogeographic studies should be better than 0.4 (i.e., if a clade had a true diversity of 10 species, more than 4 of its species would be fossilized); this is likely within the typical range of well-known fossil taxa like trilobites and ammonites, but whether other groups, particularly fossil vertebrates, reach these values remains to be determined (Lieberman, 2002b). This simulation result, like the previous one on extant organisms, suggests that whenever possible it would be highly beneficial to include fossil taxa that have still living relatives. (This is of course not possible in the case of trilobites and ammonites.) In effect, incorporating data from both the living and the fossil realms will maximize taxonomic sampling and lead to improvement in our ability to study biogeographic patterns. Using simulation studies, Turner et al. (2009) also concluded that it was important that biogeographic studies evenly sampled a group across its entire geographic range; in particular, failure to adequately sample the different biogeographic regions a clade occupies (or occupied) can lead to recovery of inaccurate biogeographic patterns.

An interesting perspective on the nature of the fossil record and biogeography was developed by Hunn and Upchurch (2001) and Upchurch et al. (2002). It relates to the issue of how clade duration affects our ability to study biogeography and also shares elements in common with our earlier discussion of how to define biogeographic areas. Hunn and Upchurch (2001) and Upchurch et al. (2002) argued that through time the biogeographic relationships among organisms may change as the geological relationships among the regions these taxa occupy change. For example, Upchurch et al. (2002) described a case involving various clades of nonavian dinosaurs; some of these clades persisted throughout much of the Mesozoic and thus experienced several Earth history regimes running the gamut from continental collision to continental rifting. An analysis that considered the paleobiogeographic

history of these clades in total might pick up a mix of signatures from these different Earth history regimes; the result might be jumbled overall patterns. Instead, Hunn and Upchurch (2001) and Upchurch et al. (2002) argued that it was best to consider the biogeographic history of these groups in the context of several distinct time slices. Although problems could result from this approach, as it might render some component taxa in a phylogeny paraphyletic, it seems interesting and Lieberman (2000a) endorsed a similar perspective. We note that Upchurch et al. (2002) only focused on patterns of vicariance and did not test for geodispersal but they did find that biogeographic patterns differed in Jurassic and Cretaceous dinosaurs.

How different time slices are defined could affect biogeographic results, just as it did in the case of area definition. It would be important to have a transparent way of defining how particular time slices should be chosen. For instance, they might be defined along the lines of intervals containing one presumed primary Earth history regime, though it would be important to avoid having this introduce any circularity into the study. Another potential way of avoiding the problem of changing Earth history regime through time, and its impact on paleobiogeography, is to focus on fossil clades that have relatively short geological durations, i.e., go extinct soon after they originate. In this manner, they would be less likely to experience several Earth history regimes throughout their evolutionary histories.

STATISTICAL APPROACHES TO BIOGEOGRAPHIC ANALYSIS

The application of statistical approaches to biogeographic analysis of clades is relatively new, and the extent to which they will ever provide a substitute for parsimony-based vicariance analyses is a question open to future discussion. Regardless of the outcome of this future debate, there are some interesting features of both likelihood and Bayesian approaches that can be usefully employed if certain assumptions are met in nature. Note that all model-based approaches to date use only molecular data. We review two of these methods and then discuss some of the more recent application of statistical methods to biogeographic questions.

The DEC Model of likelihood inference (Ree and Smith, 2008) is a bit like DIVA in that single clades are analyzed one at a time. It is different in that DEC requires a geologic model to implement likelihood calculations whereas DIVA, in its parsimony or Bayesian manifestations, is not dependent on an a priori history of the areas analyzed.

DEC builds a Q-matrix of instantaneous change using two parameters, the rate of dispersal from one area to another and the rate of extinction within a single area. Dispersal rates are generally additive, that is, if there are three areas (1, 2, and 3) and a species occupies two of these areas (1 and 2), then the dispersal rate to area 3 is a function of both the dispersal rate of 1 to 3 and the dispersal rate of 2 to 3. For species that are widespread in two or more areas, cladogenesis can occur between areas or within one area (note the assumption that the areas are pre-defined) and the sequence of vicariance events is assumed to be strictly dichotomous. Priors are determined following Ree (2005) by multiplying a flat prior for the ancestral range by the flat priors of between and within area vicariance. These priors are used to determine the overall prior for range inheritance scenarios. The method seeks to calculate the likelihood of a particular tree topology of areas using the

matrix of transition rates and the prior probability of range inheritance by integrating conditional likelihoods of range inheritance at the internal nodes in a manner similar to integrating conditional likelihoods of character states in a character analysis.

In simulations, the DEC approach seemed to work best when dispersal and local extinction are rare relative to speciation, at least at the one rate input into the simulation (Ree and Smith, 2008). The authors provide a working example by analyzing the speciation history of the Hawaiian species belonging to *Psychotria* of the coffee family.

The DEC approach is an interesting application of likelihood to the problem of biogeographic analysis of single groups. It appears to us to be restricted to those kinds of analyses that deal with geologic areas (such as the Hawaiian example of islands) rather than biotic areas. Biotic areas do not exist without biotas. They come into existence or are enlarged by dispersal and disappear with extinction.

Approximate Bayesian Computation of Hickerson et al. (2006) uses coalescence theory to test the proposition that two or more clades originated at the same time. If they did not, the conclusion is that they achieved biogeographic congruence by chance. Hickerson et al. (2006) employ this analysis to test the proposition that six sister species pairs of echinoids simultaneously vicariated 3.1 million years ago with the emergence of the Isthmus of Panama. Leaché et al. (2007) employed the same technique to hypothesize that there were two, rather than one, waves of dispersal and vicariance among the reptile and mammal faunas of Baja, California.

We find this type of analysis interesting. It directly addresses the points made by Simberloff et al. (1981) that no statistical test had been applied to vicariance analysis to demonstrate that congruence was nonrandom. And it is a potential way of approaching Lieberman's (2002b) concerns about artificial matching of vicariant patterns due to extinction. However, one must understand the assumptions underlying the application of coalescence analysis and have some assurance that the assumptions apply.

The first assumption is that both the ancestor and each of its descendants are effectively panmictic. We do not have to assume that all species have the same effective population size. We do, however, have to assume that effective population size can be determined by time to coalescence, and we must expect that effective population size is such that coalescence will occur after the vicariance event in question. We can approach the first assumption with FST analysis of the descendants. Highly structured species (populations with low gene flow) are not good candidates for analysis. The second assumption is simple to approach if the vicariance event is relatively recent in time, but, one must further assume that branch lengths are measures of time: that is, that a molecular clock applies. In addition, to actually match coalescence against a known time of vicariance or to compare between clades, the clock must be calibrated.

Given these assumptions, analysis proceeds in the following manner: (1) estimate the shape of the genealogy from the sequence data; (2) estimate relative time from branch lengths; (3) find the predicted speciation/vicariance event in relative time between clades, which will be later than the coalescence branch length between any single pair of sister species; (4) calibrate the clock; and (5) compare between clades. Relative time may also be used if one assumes that rates of evolution are constant (or constant enough) across lineages. This might be reasonable for some genes and

closely related taxa (such as the echinoderms analyzed by Hickerson et al., 2006), but is less likely to be true of mammals and reptiles in Baja where calibration was necessary.

Donoghue and Moore (2003) and Ree et al. (2005) suggest that likelihood analysis might be used to inject a greater deal of quantitative rigor into the study of biogeography. One notable example of an application of ML methods is McGuire et al.'s (2007) study of phylogenetic and biogeographic patterns in hummingbirds. To focus on the biogeographic aspects of their study, they were interested in reconstructing the ancestral states of hummingbird occurrence. McGuire et al. (2007) used these reconstructions to test several biogeographic hypotheses including whether hummingbirds originated in South America, their current seat of maximal diversity; how frequently the lineage had dispersed out of South America; and how often South America may have been recolonized from the other regions where they occur, Central and North America. Ancestral biogeographic states of nodes were estimated using a ML method that calculates the probability that each node had each of five possible states. These five states correspond to the five biogeographic regions considered in their study: North America; South America; Central America; Greater Antilles; and Lesser Antilles. The probability was conditional on the phylogenetic trees they generated using molecular sequence data, the total tree lengths, the relative branch lengths on the tree, and the biogeographic states of the terminal taxa.

The model McGuire et al. (2007) used assumed an equal rates model of character state transformation. The model takes into account phylogenetic uncertainty at particular nodes. It turns out that the estimated biogeographic state of any given node is very much affected by the length of the various branches on the tree (the essence of using branch lengths), and estimated branch lengths can differ substantially depending on which model and phylogenetic method are used to generate the input phylogeny.

McGuire et al.'s (2007) results for one of sub-clade of hummingbirds are shown in Fig. 9.14. They concluded that South America appears to comprise the ancestral biogeographic state of hummingbirds and there may have been as many as 30–34 independent dispersal events out of South America, with 28 of these lineages colonizing Central America (McGuire et al., 2007). Further, most of these events appear to have been relatively recent range extensions, indicating that the hummingbird lineages in Central America are new arrivals. Across all hummingbirds, there is only one unequivocal range expansion from Central America back into South America (not shown in Fig. 9.14). McGuire et al. (2007) also recognized that taxon sampling will have an important effect on ML state reconstructions, although this is also certainly true for parsimony-based approaches to reconstruct ancestral biogeographic states. This is particularly germane given that McGuire et al. (2007) discussed how essentially modern looking fossil hummingbirds are known from Europe, yet no taxa from Europe could be sampled in their molecular phylogenetic analysis (and the clade is today restricted to North and South America).

McGuire et al.'s (2007) study and conclusions are interesting. One potential concern that might be raised about the applicability of the ML methods McGuire et al. (2007) used is how much does the biogeographic character change model used, in their case an equal rates model, affect the biogeographic results; however, choosing a parsimony procedure to predict biogeographic character states makes similar
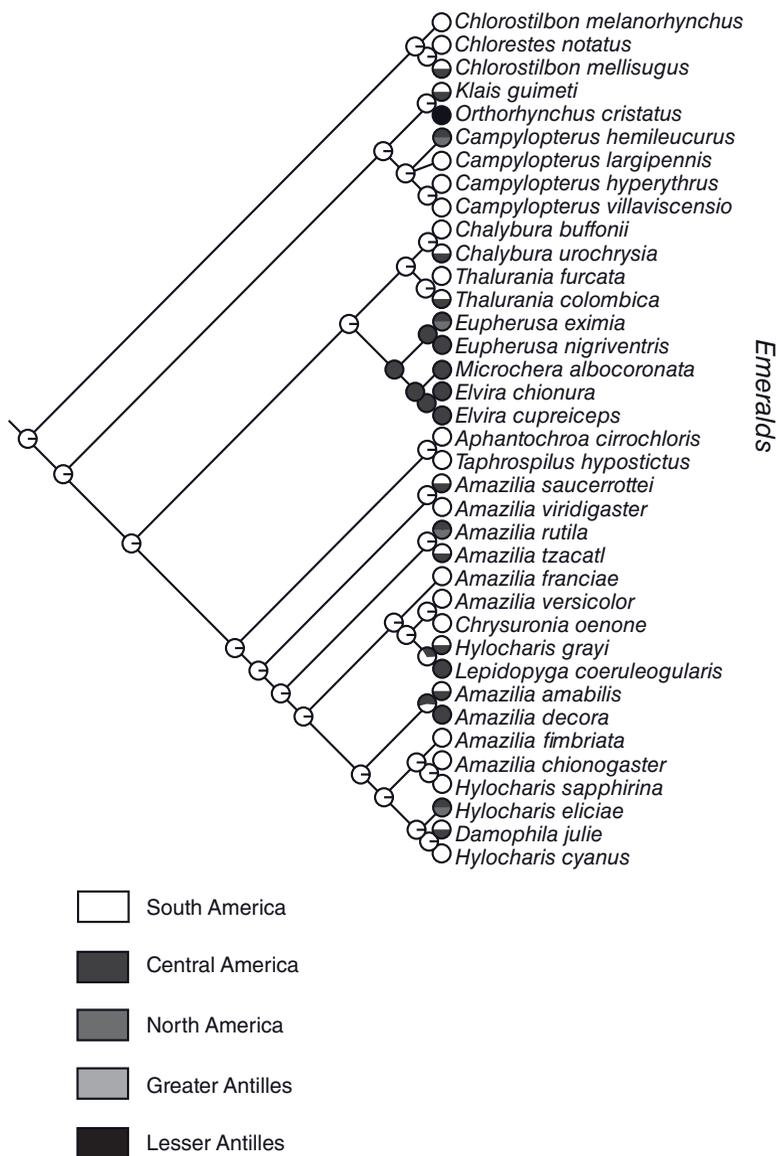
Figure 9.14. An example from McGuire et al.'s (2007) work on South American humming-birds showing the application of likelihood methods to the reconstruction and interpretation of biogeographic history. Used with permission of *Systematic Biology*, the Society of Systematic Biologists, Oxford University Press, and J. McGuire, University of California, Berkeley. See color insert.

assumptions, although it implements character estimation in a different way. Another more critical concern is the extent to which there are variations in estimated branch lengths for the phylogeny being studied. These may well have a major influence on biogeographic state reconstruction and are values derived from the models and parameters used to reconstruct the initial phylogenetic patterns. McGuire et al.

(2007) corrected for this by basing their biogeographic state estimates on a range of different branch lengths and phylogenetic topologies.

Statistical methods in biogeography are still in their infancy. There is much to investigate. Currently we do not have a good idea of how variations in branch lengths will affect a particular biogeographic pattern. As pointed out by Rees et al. (2005), work needs to be done to understand errors.

TRACKING BIOGEOGRAPHIC CHANGE WITHIN A SINGLE CLADE

There may be times when a scientist is not interested in searching for congruence among clades per se but rather in considering the biogeographic history of a single clade. There are several quantitative options that can be used to pursue such a study if a hypothesis of phylogeny is available and if the geographic distributions of its component species are known. Brooks and McLennan (1991, 2002) review many such studies in detail and provide the most comprehensive literature review available.

The simplest way of proceeding is to take the phylogeny and substitute the taxon names with their geographic distributions, producing an area cladogram. It is rarely if ever possible to identify the direct ancestor of any given species in the fossil or extant biota (Englemann and Wiley, 1977); instead, we are best able to identify hierarchical patterns of shared relationship and common ancestry. Therefore, to determine the biogeographic origin of a clade, or of any set of taxa within that clade, it is necessary to optimize the ancestral areas in a manner similar to that presented above or, if using statistical techniques, by their conditional or prior probabilities.

Once biogeographic character states are optimized to ancestral nodes, it is possible to trace the pattern of geographic change associated with cladogenesis in the group and get at questions like "what is the predominant mode of speciation in a particular group?" As mentioned already, there can be three possible transitions between the geographic state of an ancestral node ands its descendant nodes and terminal taxa: geographic range can stay the same; geographic range can contract; or geographic range can expand. If geographic range does not change between an ancestor and its descendant, this may be possible evidence for sympatric speciation, at least at the scale of geographic resolution considered; if geographic range contracts, it would be a potential example of vicariance; finally, if geographic range increases, it is a potential example of either traditional or geodispersal. Below we consider some specific examples.

An application of this approach is the paleontological study of Stigall Rode (2005b) on Paleozoic brachiopods. She used phylogenetic analysis, in conjunction with mapping biogeographic character states, to study the relationship between biogeographic patterns and evolution in Middle and Late Devonian (~380–360 million year old) brachiopods. This was a key interval in the history of animal life associated with one of the five major biodiversity crises. Brachiopods were extremely diverse and abundant in the Devonian ecosystems studied. Stigall Rode (2005b) used phylogenies of two clades to reconstruct the relative extent to which cladogenesis was associated with either vicariance or range expansion. She found an unusual preponderance of speciation associated with range expansion. Ultimately this increased dispersal, along with diminished vicariance, might have contributed to the

reduced rates of speciation that are known to have prevailed at this time (Stigall Rode, 2005b). The pattern within these brachiopod clades matches a more general pattern documented for other groups extant at this time including phyllocarid crustaceans and bivalves (Stigall, 2008). Falling speciation rates apparently played an important role in contributing to the overall decline in diversity witnessed during the late Devonian (Rode and Lieberman, 2004, 2005).

Lieberman (2003c, 2005) presented other examples of studies that used phylogenies of individual clades, in conjunction with mapping biogeographic character states, to study macroevolution. Noteworthy examples of this type of study focus on "charismatic" organisms such as ceratopsian dinosaurs including the well-known *Triceratops* (Sereno, 1997) and large carnivorous dinosaurs including the well-known *Tyrranosaurus* (Sereno et al., 1996; Sereno, 1999).

Another way of considering biogeography is in the context of coevolutionary relationships; for example, between a set of parasite species and their hosts. Brooks et al. (1981), Brooks (1985), and Brooks and McLennan (1991, 2002) present several compelling instances of how parasites and their hosts have not only coevolved with one another, but also with the Earth, particularly as the geographic regions they occupied underwent a series of climatic or geological changes. Indeed, Brooks et al. (1981), Brooks (1985, 1988, 1990), and Brooks and McLennan (1991, 2002) describe the similarities in the aims and methods of biogeographic and coevolutionary studies, and the interested reader is referred to these works and the references therein for a more detailed discussion of this topic. One nice example is provided by Hoberg et al. (2001). They studied *Taenia* tapeworms found in various species including humans, cows, and pigs and uncovered phylogenetic evidence that the tapeworm probably invaded the hominid clade long before the development of animal domestication and agriculture (which occurred around 10,000 years ago). In fact, there may have been two separate instances where tapeworms colonized hominids, and further, these invasions of hominids occurred long before *Homo sapiens* evolved. It is very likely that some of these tapeworms were acquired by hominids as one of our early ancestors scavenged bovids on the African savannah.

Brooks and Ferrao (2005) extended the union between coevolutionary and biogeographic frameworks. They argued convincingly, in our opinion, that usually the trigger for a newly emerging infectious disease, a disease that may produce high mortality on the infected population, is a biotic expansion or geodispersal event to a new host. This is because newly emerging infectious diseases often involve a disease, akin to a parasite, coming into contact with a new host that lacks the necessary adaptations to fend off the disease. This has special relevance to humans today because one of the mechanisms associated with, and contributing to, the present-day biodiversity crisis are invasive species: organisms actively or passively moved about by humans. (We will discuss this issue more fully below when we focus on biodiversity crises as biogeographic phenomena.) As we purposefully or inadvertently introduce new organisms into new regions, we run the risk of facilitating a geodispersal event by some organism containing an incipient disease that might now be able to make the jump to our own species (Brooks and Ferrao, 2005).

Including or excluding any particular taxon in the phylogeny can affect biogeographic studies because it determines which biogeographic states are input into the analysis (Lieberman, 2000a, 2002). Sampling has an obvious effect on analyses of multiple clades, but it is important to realize that it can also affect the biogeographic

analyses of individual groups (Stevens and Heesy, 2006). In particular, including only extant taxa without considering available extinct taxa results in an incomplete and perhaps biased picture. For example, living lungfishes are entirely freshwater and seem perfect candidates to test the hypothesis that they are remnants of a once continuous Pangean freshwater biota. Perhaps so, but the sobering fact is that most fossil lungfishes are marine (e.g., Lundberg and Chernoff, 1992; Ahlberg et al., 2001).

PHYLOGEOGRAPHY: WITHIN SPECIES BIOGEOGRAPHY

One research area in biogeography that has become very important in the last decade is the study of within species patterns of biogeographic differentiation: a discipline known as phylogeography (Avise, 2000). Of course phylogenetic analysis can be conducted at any level within the genealogical hierarchy, but before molecular systematic tools were available it was typically not possible to consider vicariance within individual species-lineages. Now, however, that is no longer the case. Research in this area can be valuable and exciting because it potentially allows scientists to study episodes of speciation as they actually unfold and bridge micro- and macro-evolution. Phylogeography can currently be divided into two broad research foci. Some phylogeographic studies take a focus basically equivalent to the study of among species biogeographic patterns. They search for congruence among the phylogenetic and biogeographic patterns of different individual clades, in this case, species lineages; they may also use phylogenies to focus on the evolution of individual species across geographic space. Early research in this area included many studies by Avise (e.g., 1992), Zink (e.g., 1996), and colleagues. One of the topics that they considered in detail was what role did the major environmental changes during the most recent ice ages, over roughly the last 10,000–100,000 years, play in promoting speciation. Interestingly, they found that many of the extant terrestrial species of North America and Europe actually have histories that significantly predate the late Pleistocene and extend back to the early Pleistocene, and are thus more than a million years old (Klicka and Zink, 1997; Zink et al., 2000). Isolation caused by climatic changes may have led to geographic variation in these taxa, but it did not promote speciation. Prior to these studies, in a tradition going back to Darwin (1859, 1872), it had been assumed that most modern species were quite young; further, it was assumed that they had been significantly modified evolutionarily by the changes in geographic range that the most recent ice ages must have caused. In the case of various modern freshwater and marine species, phylogeographic studies have indicated that their age of origination was even more ancient than terrestrial species, and likely extends back several millions of years (Wiley and Mayden, 1985; Mayden, 1988b; Lieberman, 2000b). There are so many phylogeographic studies that it would be impossible to list them comprehensively, but some studies that have taken such an approach include Taberlet and Bouvet (1994), Avise et al. (1998), Bermingham and Moritz (1998), Burton (1998), Lieberman (2000b), Zink et al. (2000), Bisconti et al. (2001), Sanmartin et al. (2001), Drovetski et al. (2004), Chatzimanolis and Caterino (2007), and Eidesen et al. (2007).

One potential note of caution that is warranted with any phylogeographic study is that these reveal patterns of geographic differentiation within current species distributions, but that does not necessarily mean that they reveal the geometry and

nature of future cladogenesis and speciation. These current patterns of geographic differentiation may ultimately lead to long-term divergence and speciation, but they may also be ephemeral patterns of differentiation that are not long lasting. It is reasonable to assume that most extant species have some significant history, and further, that species throughout most of their histories are relatively conservative, a notion derived from evidence supporting stasis and punctuated equilibria (Eldredge and Gould, 1972; Lieberman et al., 1995; Eldredge et al., 2005) and from basic expectations of niche conservatism derived from population genetics (Holt and Gaines, 1992). Then, it may well be that species have gone through various episodes where their component populations show genetic differentiation across geographic space but then later this pattern of genetic differentiation becomes homogenized and populations later become differentiated along alternate geographic lines. Thus, a phylogeographic study of an extant species does not necessarily reveal speciation in action and instead may reveal one in the latest of many cycles of ephemeral differentiation that a long-lived species has experienced. This is because various populations or clades within a species may exist, but as long as there is the possibility of tokogenetic relationships among these clades, and the isolation of these populations is not of sufficient duration, they may still ultimately merge with one another (in essence, the evolutionary species concept of Wiley, 1978).

There is another research area that is often treated as being a form of phylogeography that explicitly incorporates presumed models of population genetic change, estimates of population size, etc. to develop gene genealogies and evolutionary species trees. These species trees can then be used to consider issues related to biogeography. This approach sometimes employs coalescent theory (Hudson, 1990, 1992; Wakeley, 2008) and has been referred to as statistical phylogeography by Knowles and Maddison (2002) and Templeton (2004). Some recent applications in this area include Smit et al. (2007) and Carsten and Richards (2007).

THE BIOGEOGRAPHY OF BIODIVERSITY CRISES

It is well known that invasive species accidentally or purposely introduced by humans into new biogeographic areas are one of the primary causes (along with human-induced habitat modification) of the current biodiversity crisis. Lyell (1832) was among the very first to recognize the connection between human-induced species invasions and species extinctions, and he lamented the threat these invasions posed to Earth's biota (Lieberman, 2000a). Invasive species can create new competitive interactions, and in this and other ways they disrupt ecosystems (e.g., Elton, 1958; Vermeij, 1978; Wilson, 1988, 1993, 1994; Brown and Lomolino, 1998; Eldredge, 1998; Lieberman, 2000; Rode and Lieberman, 2004). They can clearly be viewed as a biogeographic phenomenon because through such invasive species humans are engineering the destruction of the very areas of endemism that helped to create life's rich diversity. Humans, by moving species around, are serving as great biotic homogenizing forces and are tearing down the natural walls of endemism that nurtured diversity. Humans are acting as agents of one of the biogeographic processes we have already described in this chapter: geodispersal. Indeed, just as vicariance

precipitates an increase in diversification, geodispersal, if maintained over long periods of time, might serve to quell diversification. Valentine and Moores (1970, 1972) and Valentine et al. (1978) amply documented the long-term effects that geology, through the mechanism of causing vicariance, has on the levels of biodiversity on the planet. They showed that for the last 500 million years there has been an excellent correlation between tectonic isolation and biological diversity. The greater the isolation of continental blocs, and the larger the number of independent continental blocs, the greater the overall levels of biodiversity at any given time. This makes sense given that increasing the number of continental blocs increases the potential for geographic isolation and thus speciation. In particular, the great rise in fossil diversity seen over the last 75 million years or so may have largely been accomplished by the breakup of Pangea and various smaller tectonic blocs.

The fossil record represents a natural laboratory to study the effects that invasive species have over deep time (Stigall and Lieberman, 2006a). It is of course not possible to use the fossil record to trace out the day-to-day competitive interactions that today cause local population extirpations, and our species' history of inducing significant species' invasions on the planet has been relatively brief. However, there are geological events that have caused episodes of geodispersal and species invasions that can be observed in the fossil record, and the long-term consequences of this profound geodispersal can be studied. These cases make it possible to investigate the long-term consequences of species invasions. One of the most famous examples involves the Great American Interchange (Wallace, 1876; Matthew, 1915, 1939; Webb, 1978; Vrba, 1993), already discussed.

Another instance where geodispersal appears to be specifically tied to a biodiversity crisis occurred during the Late Devonian interval, roughly 365 million years ago. McGhee (1996) argued that this biodiversity crisis was triggered not so much by increasing extinction rates, but rather by declining speciation rates, recognizing that if speciation rates declined for long enough the decrease in speciation rate would ultimately cause a major drop in biodiversity due simply to "background" extinctions. Because of this, although it is sometimes referred to as a mass extinction, the Late Devonian is more properly termed a biodiversity crisis. Rode and Lieberman (2004, 2005), Stigall Rode (2005a), Stigall and Lieberman (2006a), and Stigall Rode and Lieberman (2006b) documented several examples of clades that showed a Late Devonian decline in speciation rates; further, they demonstrated that this was associated with a decline in endemism at the global scale. The decline in endemism was triggered by widespread geodispersal that basically converted the relatively provincial and endemic Lower and Middle Devonian marine biota into the cosmopolitan Late Devonian marine biota. The geodispersal was caused by a series of tectonic collisions and climatic changes, including the initial stages of the assembly of the supercontinent Pangea, and coupled with global warming. The combination of global warming and tectonic changes precipitated several pronounced episodes of sea-level rise. The geodispersal that occurred can be visualized as a series of invasions by marine organisms between what were formerly isolated areas of endemism. There was a spike of invasions in the late Devonian and at the same time the average geographic range of species dramatically increased. The picture that Rode and Lieberman (2004, 2005), Stigall Rode (2005a), Stigall and Lieberman (2006a), and Stigall Rode and Lieberman (2006b) put together was one of global change,

dramatic geodispersal, and species invasions, and a decline in opportunities for vicariant speciation. The net effect was a biodiversity crisis that lasted many millions of years.

The forces that cause species invasions today are of course different from those mechanisms that drove species invasions back in the Devonian (although we note that climate change, albeit human mediated, is today playing some role in facilitating invasions, both natural and manmade). Still, we can use paleobiogeographic patterns in the fossil record to make predictions about what the long-term consequences of human-mediated species invasions, an analog to geodispersal, will be. Not only will we get elevating extinction, but the fossil record also suggests long-term declining speciation. In the end, humans may have an even more deleterious effect on global biodiversity than previously thought. As long as we continue our current activities and cause invasions (and geodispersal), we will cut off the evolutionary motor of vicariant speciation (Stigall and Lieberman, 2006a).

A BRIEF HISTORY OF THE EVENTS INFLUENCING OUR PRESENT CONCEPTS OF HISTORICAL BIOGEOGRAPHY

Ask someone not in the know to enumerate the factors that convinced Darwin that life had evolved, and you might be given a discourse on the minutiae of pigeon breeding or on the musings of Reverend Malthus, but the fundamental patterns Darwin (1859) himself identified were the geographical distributions of organisms and the fossil record. Given this, one might be tempted to connect the skein between biogeography and evolution (and thus ultimately phylogenetics) back to Darwin's epochal (1859) "Origin of the Species. ..." However, such a view is incomplete. Biogeography's connection with evolution goes back considerably further. And the treatment of biogeography in Darwin (1859) differed from Darwin's earlier views on the subject in the so-called Darwinian notebooks (Barrett et al., 1987). Darwin's shift, because of his stature in the field, ultimately cut off some promising connections between biogeography and evolution—connections that have re-emerged.

Fundamental Divisions in Biogeography, a Pre-Evolutionary Context, or What Causes Biogeographic Patterns, Vicariance or Dispersal?

One of the seemingly never-ending debates in biogeography has centered around which processes best explain the distribution of life forms on this planet. In particular, have they primarily been influenced by episodes of dispersal, where species continually radiate out from a central area? Or, have they primarily been influenced by vicariance, where ancestral species were once more widespread, and the ranges of their descendants became ever more fragmented, due to the emergence of climatic or geographic barriers within their pre-existing ancestral ranges? We have already considered, in Chapter 2, the different modes of speciation, and there we touched on the various types of allopatric speciation. In particular, we argued that vicariance is one type of allopatric speciation that occurs when the once continuous range of a species is divided by one or more barriers produced by climatic or geological causes. This geographic isolation leads to evolutionary divergence and eventually speciation. Some form of dispersal may have created the original, broad

distribution of the species, but it is not directly related to the subsequent divergence within that lineage. There are other types of allopatric speciation that are more directly conjoined to dispersal because they involve species dispersing over a pre-existing geographic barrier and thereby becoming isolated (Wiley and Mayden, 1985; Funk and Brooks, 1990; Brooks and McLennan, 1991).

Here we briefly trace the history of debate on the role of geographic isolation in evolution. Aspects of this debate extend well back into the eighteenth century and thus were played out in a largely "pre-evolutionary context." For instance, Linnaeus, the father of modern taxonomy, was one of the first to articulate a theory to explain the distribution of organisms that did not strictly rely on an interpretation derived from a literal reading of the Bible (Kinch, 1980; Mayr, 1982; Browne, 1983). His theory relied on organisms dispersing out from a central region, distributions were related to the ecological vagaries of species seeking out their preferred habitat.

By contrast, Buffon argued that different floras and faunas had been created in regions that were separated by geographic barriers such that dispersal from a central area could not explain biogeographic patterns (Nelson, 1978; Mayr, 1982; Lieberman, 2000a). In his view, species were not dispersing outward from a central region and species ranges seemed to be circumscribed by geographic barriers. Groups were originally widespread and homogeneous and over time became divided into narrower areas and thus more geographically heterogeneous. As a consequence, environmentally similar but isolated areas will contain different species (Brown and Lomolino, 1998). Indeed Darwin, writing in 1839 (Barrett et al., 1987) before he clearly and also publicly articulated his ideas on evolution, expressed puzzlement that distinct bird species were found on the different but environmentally similar Galapagos Islands. Such a pattern clearly would have found easy explanation in Buffon's ideas on the association between biotic and geographic differences.

Buffon's views rely fundamentally on vicariance, and he was not alone among pre-evolutionary biogeographers in emphasizing vicariance; in fact, the first biogeographers to study patterns in the fossil record, including Adolphe Brogniart and Alphonse de Candolle, the son of Augustin, also marshaled evidence for the notion that life appeared as a single, widespread population that gradually became fragmented into many groups distributed in many, narrower regions (Browne, 1983; Lieberman, 2000a).

Perhaps Augustin de Candolle was the first to realize that both the perspective of Linnaeus and the perspective of Buffon had merit. He recognized that there were factors that controlled biogeographic patterns at both small and large scales (Browne, 1983; Lieberman, 2000a). Small-scale factors included climate and temperature; however, these alone could not explain biogeographic patterns because regions with very similar habitats, if isolated from one another, would have very different types of organisms. Large-scale factors seemed to be reflected in the fact that unique floras and faunas were geographically isolated, suggesting that an independent geological history produced an independent biogeographic history (Nelson, 1978; Lieberman, 2000a). In a sense, de Candolle's recognition suggests the importance of hierarchies in biogeography. The island biogeography of MacArthur and Wilson (1967) with its emphasis on immigration and emigration might explain small-scale population level biogeographic patterns within limited regions, but it fails to adequately explain patterns involving several distinct species and clades distributed across many biogeographic regions.

Charles Lyell, in his *Principles of Geology, Volume II* (Lyell, 1832), built on the work of Augustin de Candolle, and also recognized that both vicariance and dispersal were important processes in the history of biotas. However, he went a step further and tried to unite the vicariance and dispersalist perspectives (Lieberman, 2000a). Although Lyell was a geologist, he was passionately interested in the subject of biogeography (Browne, 1983). As a uniformitarian, Lyell held that the Earth and its biota were influenced by a series of cycles. He focused in particular on cycles that would cause geographic barriers to form and then fall, causing the geographic ranges of many species in a region to fragment and then expand. He specifically implicated geological and climatic changes as the primary factors that formed or effaced barriers. Thus, he recognized that geological processes can cause not only congruent vicariance among groups distributed in a particular area but also congruent range expansion. This is significant because, as already discussed, both vicariance and certain types of range expansion can produce biogeographic congruence.

The Growing Evolutionary Perspective and the Continued Debate About Vicariance and Dispersal

Leopold von Buch (1825) was among the first to explicitly link biogeography and evolution in his scientific writings. Relying on biogeographic patterns in Canary Island plant lineages, he described how geographic isolation led to the formation of first separate varieties and then species (Kottler, 1978; Lieberman, 2000a). Von Buch (1825) was emphasizing the importance of what we today call allopatric speciation. Von Buch (1825) suggested that dispersal occurred over pre-existing barriers and caused geographic isolation, i.e., in modern parlance he was referring to peripheral isolates allopatric speciation (Kottler, 1978; Sulloway, 1979). Darwin's early writings (before circa 1844) in his "notebooks" (Barrett et al., 1987) also emphasized the role of geographic isolation. Grinnell (1974), Mayr (1976, 1982), Kottler (1978), Sulloway (1979), Richardson (1981), and Browne (1983) have all lucidly demonstrated how early on Darwin was thinking that the primary motor for evolutionary change was geographic isolation, and further that this isolation would be triggered by some form of geological or climatic change.

Wallace and Hooker held similar views and very modern ones at that. For instance, Wallace (1855) stated that a "country having species, genera, and whole families peculiar to it will be the necessary result of its having been isolated for a long period, sufficient for many species to have been created on the type of pre-existing ones" (quoted in Brooks, 1984:75). Wallace (1857) described vicariance in action (although not using that term) when speaking of the relationship between the biotas of New Guinea and Australia, and further argued that to understand the present state of a fauna we need an understanding of the geological history it experienced (Brooks, 1984; Lieberman, 2000a). In many respects, Wallace represents the most important early culmination of the evolutionary approach to biogeography. In the 1850s, he held that geographic barriers fundamentally determined diversity patterns and patterns of geographic distribution and further that speciation would primarily occur due to allopatry in the vicariance mode: barriers formed within pre-existing species ranges as a result of geological or climatic change. Hooker (1853) also emphasized the role of geographic isolation and vicariance in explaining biotic similarities and differences (Fichman, 1977; Brown and Lomolino, 1998).

Darwin's own views on geographic isolation made a radical shift, begun around 1844, that came to fruition in the *Origin* (Darwin, 1859, 1872). It is true that in places in the *Origin* he still argued that geographic barriers played a role in producing evolutionary divergence. However, this emphasis had become much more muted, and instead he came to rely much more on dispersal, and the competitive drive for organisms to ever move outward, as the major process that shaped biogeographic patterns. He no longer saw geographic isolation as the most important process driving speciation and evolution (Mayr, 1976; Sulloway, 1979); competition and a drive to move ever outward and adapt to new selective milieus had become much more dominant processes. This newer perspective clearly de-emphasized the role that climatic and geological processes play in causing evolution, and instead Darwin relied much more on causal biotic factors (Lieberman, 2000a, 2005).

The reasons for Darwin's shift are ultimately less important relative to our discussion than the consequences of this shift. Because of Darwin's standing and his influence, his intellectual drift toward increasingly emphasizing dispersal while de-emphasizing geographic isolation had a tremendous effect on the fields of biogeography and evolutionary biology. For example, Wallace came to contradict his earlier writings and argued that biotic dispersal was the pre-eminent biogeographic process (e.g., Wallace, 1876). This is especially significant given Wallace's earlier views, which strongly endorsed vicariance (e.g., Wallace, 1855). Incongruent biotic dispersal is also the type of dispersal endorsed as among the most important biogeographic processes by Darwin (1859, 1872; and also in the Darwinian notebooks, see Barrett et al., 1987, and Fichman, 1977, for comments). By the mid-1900s almost all biogeographic studies and writings focused exclusively on biogeographic dispersal from a center of origin across a barrier as the major causal mechanism responsible for animal and plant distributions (e.g., Darlington, 1959). Congruence, if seen, was not due to common descent but to dispersal from a common center or origin. Thus, there was no expectation that we might observe congruent phylogenetic descent relative to biogeographic and phylogenetic history but only commonalities associated with the successful negotiation of geographic barriers; these might be achieved over quite different time frames depending on dispersal ability.

It is worth nothing that this type of dispersal was also thought to be important by those scientists who held that geographic isolation played a fundamental role in speciation. For example, Wagner (1868, 1889) held that isolation, leading to differentiation, typically occurred when a species moved over a pre-existing geographic barrier, thereby becoming isolated (Mayr, 1976; Sulloway, 1979). Indeed, this is consonant with von Buch's (1825) and also Mayr's (1942, 1963) views on the subject (Lieberman, 2000a). In addition, it is the mode of speciation argued for in the original formulation of punctuated equilibria by Eldredge and Gould (1972). Each of these views corresponds to allopatric speciation occurring via peripheral isolation.

The early history of evolutionary biogeography, writ large, shows a cyclical oscillation from views emphasizing vicariance, and thus more consonant with Buffon's ideas, to views emphasizing dispersal, and thus more consonant with Linnaeus' ideas. Of course, there were still evolutionary biogeographers who continued to emphasize geographic isolation and thus at times vicariance as an important process in speciation after Darwin (1859, 1872), including Wagner (1868, 1889) and Gulick (1888), but the truth is the work was largely ignored and there was an increasing

emphasis on dispersal and sympatric speciation. This did not change until after Mayr's (1942, 1963) championing of the role of geographic isolation. Today of course most biologists hold that speciation is primarily a geographic process Ross (1972, 1986), Platnick and Nelson (1978), Rosen (1978, 1979), Wiley (1981a, 1988a, b), Brooks (1985), Vrba (1985), Wiley and Mayden (1985), Brooks and McLennan (1991, 2002), Zink (1991), Avise (1992), and Lieberman (2000a), among others. Of course, we should not assume that speciation always occurs allopatrically; the advantage of phylogenetic methods is that they make it possible to test the null hypothesis that a particular speciation event was allopatric and accept the alternative (sympatric speciation) if we reject the null.

The emergence of phylogenetic systematics did not lead all biogeographers to abandon dispersalist hypotheses, but the melding of phylogenetics and some of Croizat's biogeographic ideas led to the establishment of schools of thought (known variously as cladistic, vicariance, or historical biogeography) that emphasized vicariance over dispersal. Once these schools emerged, in some cases the pendulum swung so far back that some entirely denied the relevance of range expansion to historical biogeography. In particular, this came to be implemented by some cladistic biogeographers in the following manner: if the phylogenetic pattern of one clade does not fit the general pattern of vicariance, simply change the phylogeny to get it to fit. These are the so-called assumptions 1 and 2 of Nelson and Platnick (1981) and Humphries and Parenti (1986) discussed above. We believe now, however, that biogeographers are attempting to take a more mature look at dispersal and vicariance and put each in their proper context.

CHAPTER SUMMARY

- Phylogenetic biogeography is the scientific discipline that relates evolutionary change to geological and climate changes; as such it tries to investigate the coevolution of the Earth and its biota.
- Biogeographic studies are valuable for understanding a variety of topics including assessing mechanisms of evolution and studying the nature of past and present biodiversity crises.
- There is impressive evidence from many fields, including biogeography, that geology and climate (Earth history) are among the primary pacemakers of evolution.
- The early history of the fields of evolutionary biology and biogeography are closely intertwined.
- The primary signature that phylogenetic biogeography aims to study is evolutionary congruence across geographic space.
- Evolutionary congruence across geographic space can be produced by two processes: vicariance and geodispersal. It is essential to incorporate both of these types of processes into biogeographic studies.
- Because there are several potential sources of noise that obscure biogeographic congruence, it is necessary to use analytical techniques to tease out the signal of congruence; there is debate about the type of analytical techniques to use and different techniques have their different strengths, but we most fully

endorse a technique based on a modified version of Brooks Parsimony Analysis because it makes it possible to study both vicariance and geodispersal in a phylogenetic context.

· Critical advances in biogeography and evolution will come by integrating data from the fossil record and the extant biota because data from each of these areas has their respective strengths and weaknesses.

· There are still opportunities to make fundamental contributions in phylogenetic biogeography, and many topics need to be more fully explored including how biogeographic areas can be defined quantifiably and repeatably; how extinction and paleontological preservation affects our ability to study biogeography in the extant biota and the fossil record, respectively; how we can better integrate biogeographic studies on the extant biota and the fossil record; and what the real differences between various biogeography methods are in terms of philosophy, analytical protocol, and the results that they yield. Finally, we need more primary input data, including more results from biogeographic studies, to see which processes most consistently influence biogeography and evolution.

10

SPECIMENS AND CURATION

In the earlier chapters, we have covered what might be termed the philosophical and analytical aspects of phylogenetics and why we assert that phylogenetics is superior to other approaches to systematics. This and the following chapter will concentrate on other matters of more general interest to all systematists. We begin with specimens and curation. We will end by briefly discussing some examples of the uses of museum collections that directly speak to issues of economic and societal importance.

SPECIMENS, VOUCHERS, AND SAMPLES

A specimen is an individual organism examined by a systematist. Two or more specimens comprise a systematic series and are usually grouped to represent samples of different demes of a species or different species of a larger taxon. Most specimens are preserved in such a manner that they can be later identified by other investigators. Preservation techniques are different for different organisms, and several preservation techniques may be used within a discipline. The nature of the question asked usually determines the number of specimens collected or examined from existing collections. The list below summarizes a few possible areas of research commonly undertaken by systematists.

1. *Geographic Variation.* The investigator examines a number of series from all parts of the geographic range of one or more species. The number of specimens examined from each series usually comprises a statistical sample because variation within species is frequently a statistical exercise. Due regard must be

paid to individual variation. For example, if the species is sexually dimorphic, then adequate samples should be examined of each sex and comparisons made within sex if that is where the variation lies.

2. *Species-Level Studies.* An adequate sample of the variation of each species should be assessed when possible. To maintain good taxonomic practice, the type specimens of all nominal species should be examined, especially if a particular species is suspected to be comprised of two or more species. Samples should be drawn from critical areas of the ranges of species. For example, if two species are allopartic, then samples should be drawn from the adjacent borders of their ranges. See Chapter 2 for analytical techniques that are employed to make species decisions. Many of these techniques dictate sampling methods.

3. *Higher-Level Studies.* The nature of the samples and the number of specimens examined will vary depending on the questions asked and the number of specimens available. Adequate taxon sampling is often a problem in higher-level studies, especially if specialized preservation techniques are employed. For example, in molecular studies the sampling pool is limited by the number of tissues available, and inadequate sampling may lead to problems such as long-branch attraction that might yield anomalous results. Likewise, osteological studies may dictate low sample sizes simply because of a lack of prepared specimens. And sample size is always problematic when dealing with rare species or many fossils. One does the best one can when faced with such difficulties, relying on subsequent workers to correct mistakes.

The converse of the rare specimen conundrum is the situation where there are too many specimens. The investigator might be faced with thousands of specimens. There are diminishing returns of useful information in counting, measuring, and describing all available specimens, and the investigator must formulate a strategy of subsampling that will meet the goals of the study.

The Need for Voucher Specimens

Voucher specimens should be preserved in such a manner as to allow subsequent investigators the opportunity to identify the specimen as belonging to a particular species. Specimens housed in museums are voucher specimens because museum specimens are purposefully preserved in this manner. However, many investigators are preserving parts of whole specimens in a manner that allows them to study other aspects of their samples. Among the most common are tissue samples preserved in fluids or cryogenically so that molecular data can be taken. Specimens from which tissue samples (or karyotypes, etc.) are taken should always be preserved as vouchers and deposited in a museum. This is the only way that future investigators can access the collector's identification of the sample. If the voucher is too large (a great white shark) or the tissue taken from living specimens (whales), then every possible means should be employed to obtain a photo identification or, alternatively, some quality tag of identification (common species, or two or more investigators agree on the identification). For example, the U.S. Fish and Wildlife service conducts regular fisheries surveys and places the data in large databases. These data are useful from

a number of aspects, including documentation of the ranges of species. The data are associated with a quality tag for identification. The data for well-known species can thus be used with some confidence.

Access to Specimens

Investigators contemplating a study are faced with two choices: use existing specimen resources to complete a study or collect new specimens. The goals of the study will dictate this choice. Obviously, for example, a molecular study would require one to collect specimens if tissues are not available. However, even if new collecting is warranted, there is a wealth of information available in previously collected material that helps guide the investigator to localities where there is a high probability that new specimens can be obtained or will reveal areas where no previous collections have been made but where collecting is vital to the project. There are two basic sources of information as to where specimens that might be used in a study are housed.

Previous Literature

Most published papers have a specimen examined section that contains specimen information of some kind or another. Sometimes specimen localities are detailed, or a catalog number is mentioned and the catalog has detailed information. Earlier works, especially from the nineteenth century, can have misleading information or no useful information (locality: Atlantic Ocean), but subsequent revisionary work might provide more exact information.

Systematic Collections

Specimens deposited in museum and university collections may contain detailed information, including exact localities, dates of collection, etc. The younger the accession, the greater the chance that detailed information is available. A thorough knowledge of the taxonomic history of a group is essential to find all the records because entries are frequently cataloged under the names of synonyms.

Many phylogenetic and taxonomic studies are undertaken using only previously collected material. Indeed, why go to the expense of collecting when the specimens are already available? Museum collections can be viewed as large lending libraries of specimens. Investigators locate the specimens needed and request a loan of the material. Curators respond by fulfilling reasonable requests by qualified investigators.

Access to Specimens in the Age of the Internet

Museum collections have traditionally kept information on specimens in paper catalogs. Cross-indexing was also a paper record. Informing investigators as to museum holdings was cumbersome. In the second half of the twentieth century, museums began to capture their data electronically in the form of databases. (see Wiley and Peterson, 2003, for brief history and references). These efforts were

largely funded by granting agencies such as the National Science Foundation in the United States and the Comisión Nacional para el Conocimiento y Uso de la Biodiversidad in Mexico who understood the potential value of biological collections in addressing biodiversity, land-use, agricultural, and climate change issues. Programs specifically tailored to capturing museum data records were created. Captured data that can be easily sorted and requested by investigators can be transmitted via email (e.g., Soberón, 1999). Almost parallel to these efforts were efforts to provide direct access to museum records. Systems such as FishGopher were developed that made it possible to query records from several institutions and initiatives such as NEODAT worked to consolidate records so that they could be queried from a single Internet portal. However, a standard was missing that would allow simultaneous searches of heterogeneous databases (databases using different programs and with different data fields). What emerged was the "Darwin Core" of standard fields such as genus and species (http://wiki.tdwg.org/twiki/bin/view/DarwinCore/WebHome). This set of standard concepts was first produced in collaborative efforts and continues to be refined through various versions by the Taxonomic Data Working Group (TDWG: www.tdwg.org), an international consortium.

The establishment of the Darwin Core (DwC) was fundamental because it allowed the possibility to identify data common to two databases. What was then needed was a way to query those databases and return the information for each file of the DwC that they held in common. Early attempts used NASI/NISO Z39.50, an information transfer protocol that allows simultaneous queries of differently structured databases (Vieglais et al., 2000). This was fundamentally different from the earlier efforts, such as FishGopher that required all queried databases to use the same platform and data schema. However, Z39.50 was abandoned and a new transfer protocol was developed, Distributed Generic Information Retrieval (DiGIR), and this transfer protocol is being modified into the TDWG Access Protocol for Information Retrieval (TAPIR).

The production of standards for database information and retrieval make possible distributed database queries where the investigator can query multiple databases and retrieve all the records from these databases in a single, coherent format. In short, they make possible the concept of a single, global, virtual world museum of biodiversity information (Peterson et al., 2003). Uses of such a distributed network range from simple queries, identifying specimens for study, plotting geographic ranges, searching for misidentified specimens (records found outside the known range), and more synthetic activities detailed later in the chapter. It fills another critical role. European and North American collections in particular have vast holdings of material from other countries. Global access to specimen records affords the opportunity for scientists in these countries to access the biodiversity records of other countries if they have access to the Internet.

COLLECTING AND COLLECTION INFORMATION

There are three basic reasons to go to the field apart from the fact that field work is fun. The first reason is to collect specimens with the objective of sampling region

or taxa that are underrepresented, thus adding to our knowledge of the biodiversity of a group or region. The second is to collect and preserve specimens in a manner not represented in collections: in particular, sampling of tissues for cytogenetic, histological, or molecular study (with associated traditionally preserved vouchers). The third is to collect with a particular and specific systematic goal that requires filling in the geographic representation needed to complete the study. Collecting for specific systematic problems is undertaken for a variety of reasons, two of which are listed below:

1. The available specimens (from all sources) do not adequately cover the suspected geographic range of the group, or the number of available specimens is not adequate to answer the research question.
2. The characters of interest cannot be studied using the specimens available.

Once the decision is made to collect, the difference between a successful and unsuccessful field experience is frequently the amount of time spent in planning the trip. Poorly planned field trips are likely to result in misplaced efforts. Here are some points to consider:

1. Previous literature and field notes. Available literature can be consulted, and frequently this literature leads back to specific localities where the organisms have been collected and identifies the conditions found in the field at the date of collection. Perhaps their occurrence at different localities is seasonal. Localities that have not been previously collected from might be predictable based on previous collecting. Observations made at such localities will guide future field work. If the objective is specialized collecting (e.g., obtaining tissue samples), working at localities that have been previously collected at the same time of year are likely to yield the specimens needed.
2. Maps. Good maps are indispensable. Plot all known occurrences on the map before proceeding to the field.
3. Collecting regulations. Many taxa are regulated by specific laws that dictate that permits be obtained to collect them. Even if specific permits are not required, export and import permits of specimens between borders require permits. Investigators must obtain all required permits before going to the field to collect.
4. Collecting methods and preservation techniques. Attention to collecting methods is critical. Not only must the investigator be aware of the most efficient collecting methods, but he or she must also be aware of regulations and guidelines regarding humane collecting practices. In some countries such as the United States, institutions must follow guidelines set forth both institutionally and governmentally. These vary depending on the organism. Preservation methods should follow standards in the particular field and the aims of the collecting relative to the research to be performed.
5. Social and political considerations. Attitudes and beliefs of the peoples one is likely to contact in the field should be understood and respected.

Field Data

A certain amount of field data must be associated with specimens if they are to be of scientific value. These data should be directly associated with the sample and reflected in field notes taken at the time of collection. For example, one might tag the specimen if specimens are individually curated and the tag number is reflected in the field notes. Or one writes sufficient information on the herbarium sheet to ensure a link between the specimen and the field notes. The best policy is to minimally adhere to the standards of the repository where the specimens will eventually reside. Basic data include the following:

1. Collection information. Collecting localities are usually recorded in sequence in the field notes (Wiley uses initials, year, and collecting event, e.g., EOW 1979-1). The field number should be placed on the specimens, in the field notes, and on a map, if available.
2. Locality. The locality should be as specific as possible. For example, "U.S.; Kansas: Douglas County: 12 km W. Jct. U.S. Hwy 50 and Kansas Hwy 10 on K10," rather than "12 miles W of Lawrence, KS." The former gives an accurate distance from a known point; the latter depends on where the investigator meant be "Lawrence, KS," a city several miles in circumference. Regular use of GPS units can provide latitude and longitude, making later conversion unnecessary and providing value-added data. Of course, latitude and longitude are critical if the collections are made at sea or in terrestrial areas that lack geographic landmarks. In some cases (e.g., deep marine localities), the locality is actually a vector and not a point and the information is recorded from ship's log.
3. Date and time period. Avoid confusing dates by recording the data with the month abbreviated: 10Mar07 rather than 10/3/07 or 3/10/07. Time period of the collecting event should be on the 24 hour clock to avoid ambiguity.
4. Collectors. The names of collectors should be noted.
5. Collecting methods. The specific collecting methods should be listed.
6. Faunal or floral summary. If possible, all specimens collected during a collecting event should be recorded in the field notes along with any notes as to species seen but not collected.

The above list represents the minimum data necessary to ensure that the specimens collected have scientific value, but many investigators add more notes:

1. Site description or site picture. A brief description of the collecting site, with specific notes on microhabitats where specimens are collected will aid future investigators who revisit the site.
2. Specimen notes. Color notes might be taken. Some investigators may find it necessary to photograph the specimens to preserve this information.
3. Disposition of living or specially preserved specimens. This is especially important with tissue samples. Ensure that the vouchers and the tissues have identical identifying tags.

4. Ecological notes. A wide-ranging category from natural history notes of individual species to estimates of relative abundance, etc.

THE SYSTEMATICS COLLECTION

Systematic collections consist of series or lots of specimens that are properly documented to preserve their scientific value. Most systematic collections are housed in museums or universities and are generally separated into different collections that are curated in a similar manner. A systematic collection should be thought of in the same manner as a research library. It provides an accessible record (albeit incomplete) of the flora or fauna of the geographic regions of coverage in the same way that a library provides a record of literature on selected subject areas. Like a library a systematic collection must serve several functions:

1. Be organized in such a way that its holdings are accessible to users.
2. Be willing to make its holdings available to those qualified to study the specimens.
3. Make a commitment to provide proper long-term storage of specimens.

Many larger systematic collections strive for worldwide taxonomic coverage in particular groups, but most concentrate on particular groups and particular geographic areas. No collection can possibly provide collections for all groups in all areas, which makes the ability to access and cross-correlate records from different museums vital. The vast majority of specimens in any collection are preserved by traditional means particular to the collection or to the group collected. However, the holdings of specimens preserved in "nontraditional" ways (e.g., pollen, spores, frozen tissues, histological and karyological slides, images) are increasing as the value of specimens preserved in such ways increases.

Loans and Exchanges

The major mission of curators is to ensure that specimens housed in their collection are properly maintained. Another major mission is to ensure that they and their associated data are available to researchers. In many cases, researchers visit the facility to examine specimens, spending days, weeks, or even months at the facility. Unless they bring their own equipment, they are dependent on the facilities of the collection. In other cases, loan of specimens is requested. The person requesting the loan has certain responsibilities:

1. Request only those specimens needed to accomplish the project, and keep any single request reasonable.
2. Gear the request to those specimens that can be analyzed in a reasonable amount of time, usually dictated by the loan period (six months to one year is usual).
3. Request additional material only when the current loan is returned, or give good reason why you need to keep the material.

4. Maintain all borrowed specimens in the condition received unless permission is obtained in advance to dissect or otherwise manipulate the specimens.
5. Return the material at the time specified, or request a loan extension.

The lending institution also has certain responsibilities:

1. Make specimens available to all qualified researchers.
2. Provide a reasonable amount of time for the investigator to examine the material.
3. Recover the specimens in a reasonable amount of time.
4. Permit manipulation of common specimens, including dissection and other operations, if the researcher makes the case that such manipulation will yield data that outweight the "cost" of the manipulation.
5. Adjudicate conflicts between researchers to ensure that all have the opportunity to examine the same specimens.

Lending institutions have a special responsibility to holdings that are consumable, in particular, tissue specimens. Tissues will eventually be used up, and special consideration should be given to tissues of rare or endangered species. Having said this, we think it is a mistake to treat tissue collections as if they are proprietary collections. If the institution is going to maintain a tissue collection, then the tissues should be reasonably available to all qualified researchers. However, tissues are a consumable resource, so duplication of effort between investigators should be minimized and collaboration should be fostered to meet common goals.

Exchanges are one way collections build holdings. Recipients catalog specimens into their collections, take responsibility for their care, and, in doing so, broaden their research base to the mutual benefit of both institutions. The two institutions have certain responsibilities. A primary one is to provide documentation (collecting permits, import and export permits, CITES permits, etc.) that ensures the recipient collection that the specimens have been legally obtained.

Curation

Curation involves a series of activities from the initial receipt of specimens to the continuing processes of ensuring that specimens are properly maintained. Most systematists learn their curatorial practices from their senior colleagues. In most university and large free-standing museums, the day-to-day curation is actually performed by a growing number of professional collections managers. The curator may set policy and approve loans, but the bulk of the actual curation is in the hands of the collection manager, freeing the curator to pursue research. This allows the curator to delegate authority for the day-to-day care and maintenance of the collections under his or her care. The great variety of collections dictates a variety of curatorial practices, but the following activities are usually common to all.

Receipt of Specimens, Accessing the Collections, and Initial Sorting   Curation begins with the receipt of material. Official receipt is acknowledged when the entire collection is accessioned. Accessioning acknowledges taking official ownership. The

first step in the process is to determine if the specimens were legally obtained, and this means examining permits to ensure that all is in order. A properly accessioned collection consists of all specimens of all species taken in a particular collecting event and the field data associated with that collecting event.

Sorting and Identifying   After accession, the specimens are sorted, still keeping the field data associated with each specimen. The determination (identification to a particular taxon) of the sorted specimens is then attempted. Curators, collection managers, graduate students, and undergraduate students may help in this process, but ultimate responsibility for correct determination rests with the curators. Keys (Chapter 11) are often employed. A label with the determination, the person who made the determination, and enough data to link the specimen to the original field data is then associated with the specimen in a manner that will ensure that data are not lost. For example, the accession number serves to associate the specimens with the original field data until such time as the specimen can be cataloged.

Cataloging   Depending on the tradition for each group or collection, specimens are either given individual catalog numbers or lots of specimens of the same species are assigned a catalog number. A label or tag is fixed to the specimen or lot to ensure association with data in the catalog entry. Labels, tags, and other kinds of curation material should be of archival quality to ensure that they survive and are associated with the specimens in perpetuity.

Storage   Specimens should be stored in a manner that assures their accessibility and protection. This includes guards against humidity, temperature changes, UV light, and insect pests.

Arrangements of Collections

Particular attention should be paid to accessibility. A misplaced specimen is a lost specimen, especially in a large museum, as with books in large libraries. Some collections reflect the original garden of Linnaeus, specimens arranged in taxonomic order. This can be a great teaching tool, but it requires those who are retrieving and reshelving specimens to be familiar with the arrangement. Frequently, the taxonomic order is old, reflecting the practices of the past. This is especially true of large collections; it simply is not worth trying to rearrange holdings to reflect current ideas of phylogenic relationships or current taxonomy. Some curators have made the decision to not follow taxonomic arrangements, but place their specimens alphabetically by family, genus, and species (family names being somewhat stable), to provide easy access.

Type Specimens

Primary type specimens of species or infraspecific taxa (e.g., holotypes, lectotypes, neotypes) serve important functions in taxonomy and require special attention. Primary types are usually set aside from the rest of the collection and receive special curatorial care. Secondary type specimens (paratypes, etc.) may be curated with the

primary types or placed in the main collection. All type material should be marked in such a way that their status is apparent.

Catalogs

Catalogs contain most of the information about the specimens in collections. In some groups (e.g., fishes), specimens are cataloged by lot, each lot being composed of a number of individual specimens of the same species from the same collecting event. For many groups, specimens are cataloged individually. In either case, a label is fixed to the specimen (or placed in the jar of specimens) that associates the specimens with the catalog. For ease of use in the computer age, it is becoming more common for each specimen or lot to be assigned a universally unique identifier number (UUID) and perhaps even a bar code. UUIDs allow distributed computing systems to identify a record with reasonable confidence that the record is, in fact, unique.

Modern museum catalogs are electronic. In the early days of computer cataloging, the form of the databases and their capacity were limited, and this could make for frustratingly long searches. Technology has matured, and those collections that did not wait for perfection before converting to electronic data were rewarded. Powerful relational databases such as Specify are freely available (having been paid for by government grants) and can handle almost any information that one has available for a particular specimen, including images, data collected, links to published literature, links to other databases such as tying a specimen to a gene sequence, etc. Many museums spend considerable resources in retrospective cataloging, moving the data previously available only in paper form to electronic databases. Once the information is entered, it is relatively easy to make the data available online. This allows potential users to access the data and make requests for specimens.

What Is in a Catalog?   The information in catalogs varies considerably. Older catalogs frequently have minimal information, newer catalogs may have more. The minimum acceptable information by today's standards includes the following:

1. Museum number
2. Species name
3. Number of specimens if cataloged by lot
4. Locality where the specimen or lot was collected
5. Date of collection
6. Collector(s).

Additional information might include:

1. Group name (at least family)
2. Size range of specimens if by lot
3. Who determined identification
4. Original catalog number of an exchange

5. UUID and or barcode
6. Method of preservation
7. Tissue number of a voucher for a tissue sample (mandatory if a tissue voucher)
8. Reference to original field notes
9. Field number
10. Accession number
11. Habitat
12. Type of collecting gear
13. Cataloger
14. Remarks
15. Links to literature in which the specimen is cited
16. Links between tissues and vouchers
17. Links to data, such as a link from a voucher and tissue to a Genbank sequence.

The Responsibility of Curators

A curator has the primary responsibility to ensure that his or her collection is properly maintained. This calls for a certain amount of knowledge about specimens, preservation, upkeep, and the use of archival materials and understanding of modern curatorial practices. The curator must know enough to be able to make strategic decisions as to what to keep and what to exchange (or even discard) as specimens come into the collections. It is the curator who determines the overall quality and coverage of the collection and the strategy for collection growth.

THE IMPORTANCE OF MUSEUM COLLECTIONS

Most of the known past and present biodiversity of Earth is documented in natural history museum collections. By one estimate, some 3 billion specimens have been deposited in museums over the last 300 years (Krishtalka et al., 2002), and that number grows daily. This is the major resource for documenting biodiversity, actual specimens that can be examined and verified by future scientists.

Until recently the wealth of information contained in natural history museums has been used only by systematists. An analogy is appropriate: this is like libraries or languages only used or understood by specialists. The result is that the role of a natural history museum to society in general seems a bit obscure if not downright irrelevant. In the following sections, we would like to present some of the uses to which museum collections can be put that make them not only useful to systematists but also to other biologists and decision makers as they grapple with issues such as global climate change and the potential threat of invasive species.

As mentioned above, standards make it possible to query databases on different platforms at different museums. This provides systematists, biogeographers, and biodiversity scientists with unprecedented access to specimen data. Integrating these data with environmental information using GIS techniques can lead to insight into ecology and evolution. We will provide some examples of such inte-

grated "ecological niche modeling" studies below, but first we need to discuss some limitations that biodiversity scientists (including systematists) should be aware of when using these resources:

1. Taxonomy in the database may not reflect current taxonomy. This is especially true at the level of species where older names may be used. So, searches should include synonyms.
2. Misidentifications are common.
3. Georeferencing may be incorrect.
4. For critical work where use of the data results in publication, one should check the actual specimens. An amusing anecdote serves as a cautionary tale. Lozier et al. (2009) modeled (tongue-in-cheek) the predicted distribution of Sasquatch in western North America thereby demonstrating that one needs to scrutinize data before using it.

INTEGRATING BIODIVERSITY AND ECOLOGICAL DATA

For about the last 25 years biodiversity scientists have been developing tools that are meant to integrate the distributional data for a species and the environmental factors that prevail over the range of that species into an "ecological niche model" that is meant to ferret out those ecological factors that either predict of explain the observed distribution (for review, see Peterson, 2003). In general, the ecological factors are average temperature, land cover, or soil type. The niche models are produced via a variety of different algorithms that are generally termed *machine-learning* or *genetic algorithms*. There are a plethora of such algorithms ranging from relatively simple ones that search for a set of "limiting factors" (BIOCLIM, Nix, 1986) to relatively complex ones that take the niche as heterogeneous over space (GARP, Stockwell and Noble, 1992), to neural networks (e.g., Vander Zanden et al., 2004). The basic idea is to match the ecological conditions at the localities where specimens have been collected and to use this information to produce a predictive model of the ecology of the species. The ecological conditions are taken from electronic coverages and range from relatively coarse (one square degree) to relatively fine (on the scale of meters squared). In a GIS environment, each pixel would be associated with the values for each environmental coverage. And each locality for specimens of the species studied is also located on a specific pixel. The algorithm uses this information to construct an ecological niche model for the species as a whole by learning about the environmental conditions associated with each specimen record. The trick that makes the system predictive is this: the niche model can then be projected back on the landscape to see which pixels it predicts the species' niche should be present, even if the investigator does not have a specimen record from that pixel.

There are some issues to consider. Most of our data about biodiversity comes from what is known as "presence-only" data. That is, we know where we have collected samples, but just because we have not collected samples at a particular locality does not mean that the species is not present. True absence data are hard to come by. Algorithms such as GARP and MaxEnt are designed to work with

presence-only data. Many neural network algorithms require absence data as well as presence data, which is the reason why GAPR (Stockwell and Noble, 1992; Stockwell, 1999; Stockwell and Peters, 1999) and MaxEnt (Phillips et al., 2004) and other presence-only algorithms dominate in biodiversity studies at large scales.

A Simple Example: Range Predictions

Wiley et al. (2003) used the machine-learning program GARP to model the distributions of a number of marine fishes whose geographic ranges centered on the Gulf of Mexico. Environmental coverages included the World Ocean Atlas 1998 data set (including nine coverages such as temperature, salinity, and dissolved oxygen; NOAA, 1999) and bathymetry (Smith and Sandwell, 1997). The results for the relatively stenotopic shark *Etmopterus schultzi* is shown in Fig. 10.1. The yellow circles represent localities used to model the niche while the green circles represent known samples that were projected onto the landscape after the prediction was made to test the model. The results are statistically significant and point to areas where the species has not been sampled but should be found. See Wiley et al. (2003) for details of coverages and methods.

Figure 10.1. Prediction of geographic distribution of the shark *Etmopterus schultzi* in the Central Atlantic, Caribbean and Gulf of Mexico using GARP. Some point localities are used by GARP in concert with 9 WOA 98 environmental surface coverages and bathymetry. Other point localities are withheld from modeling and used to test the prediction. Blue denotes bottom depth, with lighter blue indicating relatively shallow waters. Pink to rust brown shading denotes number of model intersections: pink, 5–6; red, 7–9; rust brown, 10 intersections respectively. The inset shows details from off Louisiana. From Wiley et al. (2003), Oceanography, volume 16, number 3, Figure 2: 124, used with permission. See color insert.

Predicting Species Invasions

The use of museum data for predicting how widely an invasive species might spread if successfully introduced is built around the ability of the investigator to first build an ecological niche model and test it over the native range of the species and then project that niche model onto a new landscape and see where the niche might exist. Peterson (2003) provides a review of this research program.

Modeling does not forecast the probability of successful invasion, but it does seem to do a good job of forecasting the eventual distribution of successful invaders. For example, a niche model for the aquatic weed *Hydrilla verticillata* generated from only 20 native occurrence points (all that could be found!) in its native range in Asia successfully predicts the occurrence of this invasion in drainages in North America where it is a known invader and suggests that this noxious weed will reach an even larger distribution. This has economic consequences for inland water transportation and an unknown economic impact on native freshwater habitats (Langeland, 1996).

Asia is also vulnerable. Largemouth bass from North America have been extensively introduced in the southern islands of Japan. Niche models made from North American occurrence data from museums predicted 96 percent of the known Japanese occurrences and projected that the species could become established in the northern island of Hokkaido (Iguchi et al., 2004), a prediction that actually materialized in 2001 (Teranishi and Ohhama, 2004). Small mouth basses are not so well established, but the 10 known occurrences were predicted by North American niche models. The small mouth bass may be more of a threat to native fishes. Bass fishing is an industry that generated from $500 million to $1 billion in Japan, but the government has called for the removal of all nonnative species from Japanese waters through the Invasive Alien Species Act of 2005 (www.env.go.jp/en/nature/as.html).

Global Climate Change

One of the more comprehensive studies of the effects of global climate change is the study by Peterson et al. (2001a, b) on the effects of climate change on the birds, mammals, and butterflies of Mexico. The study included the compiled records of presence from some 45 natural history collections and included individual modeling of some 1,870 species of birds, 416 species of mammals, and 175 species of butterflies. The results under two climate change scenarios suggest that while extinctions and drastic range changes may be relatively rare, species turnover of communities would be high, suggesting considerable ecological perturbations. Less optimistic and more controversial is the more recent study by Thomas et al. (2004) that predicts that between 18 percent and 35 percent of species in their study areas may be fated to extinction by 2050 under 3 climate change scenarios (minimum, moderate, and maximum given no change in current trends).

CHAPTER SUMMARY

- Specimens and series of specimens are the backbone of systematic study.
- How and what to examine depends on the nature of the proposed investigation.

- All samples should be documented with voucher specimens, especially those that consist of part of the organism that cannot be immediately identified such as tissue samples. Photographic vouchers are acceptable in many circumstances.
- Electronic databases and the Internet are powerful tools to access specimens, but traditional resources must also be used.
- Field work should have specific objectives and require considerable preparation.
- Systematic collections must be well organized and available to qualified workers.
- Museum collections document the bulk of the biodiversity of the Earth and are valuable not only for systematics but also for such economically important issues as global climate change and the threat of invasive species.

11

PUBLICATION AND RULES OF NOMENCLATURE

Systematics is one of the few scientific disciplines where publications over 10 years old are still relevant. Systematists must be historical scholars and taxonomic lawyers as well as research scientists. Historical scholarship requires knowledge of the kinds of systematic literature and the metadata for that literature in the form of nomenclators (essentially lists of names and their origins) and other kinds of publications. Familiarity with various rules of nomenclature is needed if the systematist engages in any kind of descriptive work or revisions where past literature is pertinent to the research. The mark of a complete phylogeneticist is one who can combine critical phylogenetic analyses with solid taxonomic scholarship when needed.

KINDS OF SYSTEMATIC LITERATURE

Descriptions of New Species

The number of described species of macroeukaryotes is on the order of 1.4 million species (Wilson and Peter, 1988; Wilson, 1993), and the number of undescribed species is between 5–30 million (May, 1988). No one has any idea of the potential number of species of prokaryotes, and the number of protists is likely to be much higher than present estimates (Bass et al., 2007). Thus, we can expect that species descriptions will continue to be a major activity for systematists.

Species descriptions range from isolated descriptions to descriptions embedded in taxonomic revisions. Although commonly thought of as descriptive science, descriptions of new species require considerable synthesis, including a thorough

knowledge of related species, the history of taxonomy of the group, a careful examination of related species, and an examination of type specimens of related species and types representing synonyms of currently recognized species. Comments are often made about the relationships of the new species to others, and new species are frequently presented in keys that allow their identification relative to existing species.

Revisionary Studies

Revisionary studies range from synopses and reviews of the current state of the taxonomy of a group to monographs. Synopses and reviews summarize the current knowledge of a group, bringing together as much of the existing knowledge of the taxonomy of the group as can be gleamed from literature. As such, a proper synopsis is a valuable systematic tool, bringing together all of the scattered literature of a group. Classifications are usually found in revisionary works, but occasionally published separately in the older literature (e.g., Simpson, 1945, mammals; Wetmore, 1960, birds; Takhtajan, 1969, angiosperms, updated, 1997). These works can be considered a kind of synopsis.

Revisions and monographs are the most demanding of all systematic works that include the taxonomy of a group. At their best, they include:

1. The complete taxonomic history of a group.
2. Diagnoses and descriptions of each species.
3. Keys for identification.
4. All literature relating to the systematics of the group.
5. A revised classification.

Of course, if the reviser is a phylogeneticist, we can expect a phylogenetic hypothesis and that the revised classification will contain only monophyletic groups.

Keys

Keys are usually found in revisionary faunistic or floristic works. However, they may be published separately. The Internet provides an excellent venue for the electronic publication of keys to specific regions that can be easily updated with increases in knowledge or changes in the fauna and flora. Many museums make such electronic keys available through their websites. Keys to economically important pests are especially common (e.g., www.entomology.umn.edu/ladybird for ladybird beetles of Minnesota). Such keys afford details that are not easy to capture, or expensive to publish, in print form.

Faunistic and Floristic Works

These might be termed monographs for specific areas in contrast to monographs of specific groups. Usually, the study is also restricted to a particular group. Such a work may include:

1. Accounts of species that inhabit a defined area. These may range from very brief accounts to more complete accounts that include considerable taxonomic information such as the synonyms of species and even descriptions of new species.
2. History of previous work in the area for the organisms studied.
3. Keys to the identification of species.

Floristic works are one of the major venues of botanical publication. In addition to the basic systematic data of the plants themselves, Radford et al. (1974) suggest that the following topics be addressed:

1. Location and geography of the study area.
2. A history of botanical exploration.
3. A survey of applicable physiography and topology.
4. A summary of major biogeographic patterns placed in historical context.
5. A summary of present ecology.
6. A summary of pertinent pedologic and geographic data.
7. A summary of pertinent climatological data.
8. A review of previous works.
9. A description of present land use or abuse.
10. A list of cited references.

Atlases

An atlas may include illustrations of the species of a particular taxonomic group or geographic area. It may contain illustrations or distributional maps. An example is the Atlas of North American Freshwater Fishes, which includes not only a brief account of each species but also a photograph and a spot map of distribution (Lee et al., 1980 et seq.).

Catalogs

A catalog is a tabulation of species detailing varying amounts of information. Catalogs of type material are common and frequently include references to the original description, synonyms, and ranges. Many type catalogs are now placed online for access over the Internet. An example is the catalog of type specimens of the University of Florida Herbarium, which is searchable by scientific name, family, or collector (www.flmnh.ufl.edu/herbarium/types).

Checklists

A checklist is a list of species for a particular group, either for a specified area or globally. Checklists are especially popular for butterflies (online; www.naba.org) and birds (AOU checklist of North American Birds: www.aou.org) and are commonly included in field guides and faunal/floristic works.

Handbooks and Field Guides

Handbooks are usually designed to enable the nonspecialist to identify groups and species occupying a particular political or geographic region. Inclusiveness and coverage varies.

Taxonomic Scholarship

Formal revisionary works deal with most taxonomic problems and are published in peer-reviewed journals. Some problems, especially those concerning priority of names, appear only in specialized journals. In zoology, papers dealing with nomenclatural issues are published in *The Bulletin of Zoological Nomenclature*, a journal sponsored by the International Commission on Zoological Nomenclature (ICZN). In botany, the journal *Taxon*, sponsored by the International Association of Plant Taxonomists, is the major forum for nomenclatural discussion.

Phylogenetic Analyses

The purpose of phylogenetic analyses is to place species in the context of their common ancestry relationships. A phylogenetic analysis may be included in a revisionary study, but it is more and more frequent to see phylogenetic analysis performed against the backdrop of the current taxonomy of a group without the accompanying taxonomic revision. Such studies lead to new insights into the relationships of taxa and may inform subsequent revisions compiled in more traditional form.

ACCESS TO THE LITERATURE

An exhaustive literature search is usually in order when a taxonomic revision is contemplated. The reviser must account for all the names ever used for the taxa revised and the nomenclatural histories of these names. The best beginning is to consult the latest revision of the group, followed by a literature search. More and more journals are appearing on the Internet; however, much of the older literature has not been electronically captured. We make note of the various bibliographic aids that now appear in electronic form, but they do not entirely replace more traditional, paper-based, literature searches.

 As in many aspects of science, the Internet has revolutionized literature searches. However, any search using only the Internet should be considered incomplete. Not all of the resources are online, and due respect should be paid to library searches, especially for older or obscure literature. Fortunately, libraries are cooperating to an extent unheard of in the past. For example, if you can find the reference, you can frequently get a PDF of the paper.

Literature in Zoology

1. Biological Abstracts (online). Many subdisciplines covered since 1926. Not as comprehensive in taxonomic material as specialized bibliographic services that concentrate on botany or zoology.

2. PASCAL and predecessors (online from 1973). Covers many disciplines, formerly *Bulletin Signalétique*. Coverage of literature in about 100 languages, most with English and French titles and with French abstracts. Pertinent are biosciences and geology.

3. *Berichte über die gesamte Biologie*. German equivalent of biological abstracts covering the period since 1926.

4. Harvard University Herbarium. Links to a number of other Internet sites with literature information.

Zoological bibliographies include the following:

1. *Zoological Record*. The *Zoological Record* originally published by the Zoological Society of London with the cooperation of the Natural History Museum, London, is now published through Thompson Scientific by subscription. The *Record* summarizes the systematic literature from 1864 to present. The last 25 years are available online. Coverage in the early volumes is spotty while coverage in later volumes is fairly complete. The taxonomy follows the accepted taxonomy of the year of publication.

2. *Archiv für Naturgeschichte* (now *Zeitschrift für Wissenschaftliche Zoologie, Abteilung B*). The bibliography covers the systematic literature for 1832 to present.

In addition to the various bibliographies listed above, direct access to names, authors, and dates for genera and species may be gained from various nomenclators. Those proposing new names should always consult all appropriate nomenclators to avoid homonomy. The older nomenclators, in print versions, are shown in Table 11.1. An updated online version of *Nomenclator Zoologicus* (print volumes edited by Neave between 1939 and 1950) is currently in final stages of development (version 0.86 as of 2008) and covers zoological names from 1758 to 2004. In addition to nomenclators, there are detailed accounts of taxonomic names for particular groups. For example, the Catalog of Fishes accounts for all known generic names and specific epitaphs of fishes and is available online (Eschmeyer and Fricke, 2009).

Literature in Botany

A large number of botanical resources are online, either free or by subscription (institutional, individual, or part of society membership). Many of the nomenclators and indices are now available in electronic form, and we emphasize access to these sources, many of which appear in earlier paper editions. Botanical bibliographies include the following:

1. *Taxonomic Literature: A Selective Guide to Botanical Publications and Collections with Dates, Commentaries and Types, 2nd edition*, edited by Frans Stafleu, Richard Cowan, and later also Erik Mennega. One of the *Regnum vegetable* series of the International Association of Plant Taxonomy. Online

version available free to members of the International Association of Plant Taxonomy.

2. Other useful works published under the auspices of the International Association of Plant Taxonomy include the *International Code of Botanical Nomenclature*, *Index Nominum Genericorum*, *Names Currently in Use*, as well as other works. Most are available online to members of the association.

3. *Kew Bibliographic Databases.* Provides online access to *The Kew Record of Taxonomic Literature*, the *Plant Micromorphological Bibliographic Database* and the *Economic Botany Bibliographic Database*.

There are numerous botanical nomenclators and indices. Some are listed below:

1. *Nomenclator Botanicus* (Steudel). Covers plant names from 1753 to 1840. The latest edition (1840–1841) is available online through the Biodiversity Heritage Library.

2. *Nomenclator Botanicus* (Pleiffer). Covers names of higher plant taxa through 1859.

3. *International Plant Names Index.* An online collaboration of the Royal Botanic Gardens, Kew, the Harvard University Herbaria, and the Australian National Herbarium that is a database of names and associated bibliographic details of seed plants, ferns, and fern allies. It provides access to the *Index Kewensis*, the *Gray Card Index*, and the *Australian Plant Names Index*. The *Index Filicum* is also included.

4. *Index Nominum Genericorum* (Farr et al., 1979 et seq.). A generic nomenclator covering all plants including algae and fossils, as well as fungi, with type species indicated. (Available online.)

5. *Index Muscorum* (Wijk et al., 1959–1969). Covers the names of mosses from 1801 to present, with post 1969 supplements published in *Taxon*.

6. *Index of Mosses Database.* An online database project of the names of mosses of the Missouri Botanical Gardens with bibliographic links.

TABLE 11.1. Nomenclators for animals. Blackwelder (1967:234) conveniently orders various nomenclators by period of coverage. See Blackwelder (1972) for other references that emphasize Vertebrata. Many of these volumes are available online.

Period	Reference	Coverage
1758–1800; 1801–1850	Sherborn (1902; 1922–1933)	Genera and species
1758–1842	Agassiz (1846, 1848)	Genera
1758–1873	Marschall (1873)	Genera
1758–1882	Scudder (1882)	Genera
1758–1926	Schulze et al., (1926–1954)	Genera
1758–1945	Neave (1939–1940, 1950)	Genera
1801–1910	Waterhouse (1902, 1912)	Genera

PUBLICATION OF SYSTEMATIC STUDIES

Publication is the final step in systematic research. The format of systematic papers varies with the type of study and the journal to which the manuscript is submitted. Journals usually contain a section in each issue that instructs prospective authors on matters of style and content, and these should be carefully read and followed. The *CSE Manual for Authors, Editors and Publishers* (Style Manual Committee, 2006) is an invaluable resource for the prospective author. The format for a systematic work will usually contain the following parts:

1. Title. The title should be informative without being overly long. Many workers scan or electronically search for words expected to appear in the title. An obscure title leads to an obscure paper. If the name of the group or species is not immediately recognizable, then the name of the major group and family should also appear in the title.

2. Author's name and address. Authors should be consistent and use the same form of their name in their publications. The address should be the address of the institution at which the author was affiliated at the time period at which the research was performed, with the current mailing address following. This ensures that the institution receives proper credit.

3. Key words. Effective electronic searches depend on good use of key words. Choose them wisely, using the guidelines provided by the journal.

4. Abstract. A good abstract is crucial to the paper. In an age of increasing number of publications, workers heavily depend on the title and the abstract to determine if the paper is worth reading in detail. A good abstract will interest the prospective reader to read the entire paper.

5. Introduction. The scope and purpose of the paper should be presented along with the historical background leading to the study. This frequently takes the form of a literature review directed at the specific problem. In revisionary studies the author hopes to capture the history, in brief, of taxonomic work on the group.

6. Materials and methods. This section should explain the protocols used in the study in sufficient detail so that the work can be repeated. Frequently, this takes the form of referencing standard protocols, but if the investigator deviates from standard protocols, these deviations should be spelled out explicitly. In revisionary work, the section should also present a list of the specimens examined (or reference to an appendix containing this information) in sufficient detail such that subsequent workers can re-examine these specimens (e.g., by citing institutions and catalog numbers). In molecular studies, the primers used should be presented or referenced and the voucher specimens listed. In addition, the gene sequences should be deposited in GenBank.

7. Body of the text. The body includes the results, conclusions, and discussion, as appropriate. Various aspects of more formal taxonomic sections of the body will be detailed in subsequent sections.

8. Acknowledgments. Those who contributed to the author's efforts should be acknowledged. This should include institutions and staff that contributed specimens or data, persons who helped with techniques, the collection of

data, formal (if known) and informal reviewers and those who helped prepare the manuscript. The agency providing funding (if any) should be acknowledged.

9. References cited or bibliography. All studies cited in the paper appear in this section. The author should check the journal for format before submission.

10. Appendices. Appendices are frequently useful for detailing character descriptions, listing specimens examined, detailing collection localities, and presenting other material.

Major Features of the Formal Taxonomic Work

Certain features of a species description or group revision have a more formal manner of presentation than do papers that do not contain the formal presentation of names meant to conform to nomenclatural codes. In general, aspects include name presentation, formal diagnosis or description, synonymies, and material examined. Additional components may include comparisons, distributional data, etymology, a key, and illustrative material. There are several formats for formal taxonomic presentations, and these tend to differ from group to group and journal to journal, so the reviser should understand the practices particular to his own field. Below is one of these many formats, presented for illustrative purposes only:

1. Presentation of the valid/correct name
2. Reference to figures
3. Synonymy
4. Type material examined
5. Other material examined
6. Diagnosis
7. Description
8. Comparisons not covered in the diagnosis
9. Distribution
10. Etymology
11. Key if appropriate.

Name Presentation

Names are minimally presented as the name of the taxon followed by the author and year of publication for those names covered by the appropriate code and thus subject to the rules of priority. Frequently references to figures and common names are included. If the taxon is new, this is noted without author designation (because the author(s) are those of the paper). The examples below are from Naumann (1977, Tinthiini, sesiid moths) and Vari (1978, teraponid teleost fishes).

Tinthiini LeCerf, 1917
(Figure 11a, male genitalia, Figure 6)
Bidyanus bidyanus (Mitchell, 1838)
Silver Perch
Figure 77

Note that Mitchell's name is put in parentheses because he originally placed this species in another genus. Other name presentation may include the complete reference citation and reference to type material, as for example, this name presentation by Ee and Berry (2009) for a species of *Croton* (Euphorbiaceae).

Croton jamaicensis van Ee & PE. Berry, sp. nov.—TYPE: JAMAICA. St. Catherine: Healthshire, near Salt Island, 1 Sep. 1908, *Wm. Harris & N.L. Britton 10520*. (holotype: NY!; isotypes: BM!, P!, UCWI!).

If new names are proposed, they should be explicitly designated by using phrases such as "new species" or "nov. sp." The zoological rules require this for the lower categories where priority applies.

Synonomies

Synonyms are the various names that have been validly published and applied to the same taxon. The senior synonym is usually the correct name of the taxon, or if the rules of priority do not apply, the preferred name of the taxon. Rules governing synonymy can be found in each nomenclatural code. Here we are concerned with their presentation in publication.

A major part of any revision or description of a new species is a presentation of names that have been applied to the taxon. This is an important part of the revision because it links the published concept of the taxon to past concepts and thus permits the past literature to be placed in context. Synonomies come in two basic forms. Complete synonomies purport to give every reference to every name ever applied to the taxon. Such synonomies are rare. Abbreviated synonomies purport to list those names that have directly affected the taxonomy of the group or species and provide entry into the pertinent taxonomic literature, which might include not only past revisions and descriptions, but also guides and floral or faunal works. Although only valid synonyms affect the history of the name, synonomies should also include unacceptable names and mistakes in identity with suitable annotation to indicate their nature.

Species synonomies (the most commonly encountered formal synonomies) minimally should include (1) the original form of the name, (2) the author and date of publication, and may contain (3) reference and page number. There are two basic kinds of formats: by date or by name. Here are two species and one genus example.

Example 1. Arrangement by Date of Publication, References in Literature Cited/Bibliography (from Bartram, 1977, A Bony Fish)

Proterus elongates Wagner, 1863

1893 *Proterus elongates* Wagner: 645
1881 *Notagogus macropterus* Vetter: 46
1895 *Proterus speciosus* Wagner: Woodward: 184, pl. 3, fig. 5
1941a *Proterus speciosus* Wagner: Eastman: 407, pl. 13, fig. 1

Example 2. Arrangement by Date of Name, Reference in Literature Cited (from Rindge, 1972, A Moth)

Plataea calcaria (Pearsall)

Apricrena calcaria Pearsall, 1911, p. 205. Barnes and McDunnough, 1917, p. 122.
Plataea triangularia Barnes and McDunnough, 1916, p. 27, pl. 3, fig. 18 (holotype maes); 1917, p. 115; 1918, p. 151 (placed as synonym of *calcaria*).
Plataea dulcinia Dyar, 1923, p. 23.
Plataea dulcinea [sic.]: McDunnough 1938, p. 170 (placed as synonym of *calcaria*).

Example 3. Arrangement by Date, Reference Included, Synonymy in Paragraph Style (from Koponen, 1968, Mosses)

7.   Genus *Orthomnion* Wils. 1857
     Orthomnion Wilson in Mitten, Kew Journ. Bot. 9:368. 1857.—*Mnium** *Orthomnion* (Wils., Mitten, Journ. Linn. Soc. London Suppl. Bot. 1:142. 1859.—Typus: *Orthomnion bryoides* (W. Griff.) Norkett (cf. Norkett, 1958).

Generic synonomies differ from specific synonomies principally in that (1) only available synonyms are listed and (2) the type species is listed. Frequently the reason why the type species is a type is also noted. The generic synonomy from Wiley (1976, garfishes) is an example where types are annotated.

Lepisosteus Lacépède

Lepisosteus Lacépède 1803:331 (type species *L. gravialis* by subsequent designation, Jordan and Evermann, 1896:109).
Sarchirus Rafinesque 1818a:418 (type species *S. vittatus* by subsequent designation, Jordan, 1877:9).
Cylindrosteus Rafinesque 1820:72 (type species *C. platostomus* by subsequent designation, Jordan, 1877:11).
Lepidosteus (Lacépède): Koenig, 1825:12; Agassiz, 1843:2 (amended spelling of *Lepisosteus*).

Material Examined

Material examined can be listed in several places. Type specimens are frequently listed immediately after the synonymy and should include the type locality, but placement depends on the particular journal. The style of presentation should be as abbreviated as possible and frequently only gives the name and the catalog number of the specimen (with museum acronym). In some cases the specimens may not be listed at all, but reference is made as to how that information can be accessed.

The Diagnosis

Diagnoses in revisionary work have a different function than diagnoses used for conveying phylogenetic characters of monophyletic groups. The purpose of a diag-

nosis in revisionary work is to set aside the taxon from other taxa, that is, to distinguish the taxon from other taxa. Thus, the diagnosis may convey more than synapomorphies or autapomorphies, but may include any characters useful in demonstrating that the taxon is different from other taxa. Differential diagnosis, where the characters are directly compared with closest relatives, are the most informative. A good diagnosis will allow other systematists to identify members of the taxon in the most concise manner possible without having to consult the more detailed description. If there is no key, the diagnosis takes on an even more important role. Avoid the tendency to write a diagnosis like a description; this obscures the functions of both. Length is not important; clarity is paramount.

Botanists must write at least the diagnosis in botanical Latin, and it is not comparative. Stearn (2004) allows one to accomplish this task and provides a wealth of information about botanical Latin and a very useful dictionary. A full description written in Latin is permitted, but rare. Below, is an example of part of the diagnosis of *Croton jamaicensis* Ee & Berry (2009) for the benefit of zoologists who are not familiar with Latin diagnoses.

Arbuscula monoica 2–5 m; foliis ovato-lanceolatis, acuminatis, basi eglandulosis triplinervis, penninervis, …

The Description

The description is an account of the characters studied. Descriptions may be exhaustive or concise. More efficient descriptions are sometimes possible by treating them hierarchically. That is, do not repeat characters found in all species of a genus in the description of each species, cover them in the description of the genus, etc. The style of the journal and traditions within the field of study usually dictate the breadth of the description, and the best guide is to read the literature to get a feel for the expectations in the discipline. Redescriptions are cast in a similar manner as descriptions.

One can help later revisers if one adopts a standard style for both diagnoses and descriptions. In zoology the form adopted may vary according to the traditions of the group. In botany suggestions for standard plant descriptions are found in many of the botanical systematics books such as Radford (1986), Simpson (2006), Judd et al. (2008), and Stussey (2009).

Illustrations and Graphics

Illustrations and graphics illustrating characters may be considered part of the description and are cited as such in many revisions (see above). Apart from color illustrations usually rendered by professional illustrators, the choice of medium is line drawings, stippled/shaded drawings, and photographs. The choice should be based on the best medium to illustrate the characters. Systematists who do not have the services of a professional illustrator must pick up illustration skills, including how to use such mechanical devices as camera lucidas and digital cameras, with attendant issues of lights, etc. There are several excellent books on biological illustration including Hodges (2003), Briscoe (1995), Wood (1994), and Zweifel (1988). Illustrations and graphs take on a new meaning in field guides, floras, and faunas where descriptions are necessarily short and discuss only those characters that aid in identifying species.

Comparisons and Discussion

Within the descriptive format, the comparison section may be used to contrast the characteristics of closely related species in more detail than that covered in a differential diagnosis. This is the section where it is appropriate to discuss phylogenetically informative characters. The discussion section can be used in various ways, e.g., (1) to discuss rationale for describing the taxon, (2) in paleontology, to bring readers to the attention of fragmentary specimens that elude exact identification or that were incorrectly identified in the past, (3) to point to gaps in collection coverage and geographic regions that need to be explored, or (4) to call attention to new or previously unknown characters. These sections also provide a venue for discussing the relationships of the species to other taxa.

Distributional Data

The range of a taxon should be described in words. Additionally, a distribution map can be presented based on the specimens actually examined by the author. Obviously, a taxon with a narrow distribution may not warrant a map, the verbal description would suffice. There are several different map formats.

Figure 11.1. Distribution of the largemouth bass, *Micropterus salmoides*, in its presumed native range, with an ecological forecast for North America using GARP from Iguchi et al. (2004). Point data (dots and triangles) gathered from 11 museum databases via FishNet and direct access. Dots and triangles are locality data, with dots used for niche modeling and triangles for testing the resulting models. Dark red represents the joint predictions of potential range from 10 models, light red represents the joint prediction of 7–9 models. The 10 models visualized were determined by objective criteria, as discussed in Iguchi et al. (2004). From Wiley (2007). Transactions of the American Fisheries Society 136, Fig. 1, p. 1131. Used with permission of the American Fisheries Society. See color insert.

1. Spot maps. These are the most accurate type of map and spot maps are the preferred format for taxonomic publication. Each locality is shown as a symbol on the map, and each species has a different symbol (Fig. 11.1). With increasing use of GIS technology and accessibility of electronic databases, production of such maps is now much easier. However, for the taxonomic paper, every record presented on such maps should be comprised of specimens actually examined by the author(s).

2. Boundary-line maps. The entire range of a taxon is shown with a polygon that encompasses all known records examined by the investigator. In less formal works (such as field guides), the polygon may encompass all records thought to be reliable, not simply those personally examined by the investigator. Again, GIS technology makes the production of such maps much easier.

3. Pictorial maps. These maps combine both geographic and morphological data. They are popular for some groups and are used effectively to portray geographic variation in selected characters.

Etymology

The origin of a name is a point of taxonomic scholarship. Descriptions of new taxa should include the origin (and gender if a new genus) of the names. It is a service to the community to also include the etymology of all genera and species in a revision. In many groups the etymology has been presented by previous revisers and can be consulted. Two common sources that assist in determining etymology are Stearn (2004), Brown (2000), and Borror (1960). Note that etymology is required by the bacterial code and recommended for the botanical code.

Keys

Keys are devices used to identify specimens. There are two major kinds of keys. *Structured keys* (single-access keys) present an abbreviated list of contrasting characters that lead to the next set of contrasting characters until the specimen is identified. Technically, they are decision trees. *Polyclave keys* (multiaccess keys, expert-systems) present a list of characteristics. The investigator picks characters from the list that match the specimen and then submit the list to a computer program that matches the list of characters to a data bank of taxa and characters and returns the identification to the investigator based on matches.

Although some may present more than two choices, the most common variety are *dichotomous keys*. Clarity and convenience are the most important qualities of a good-structured key. These qualities can be incorporated into a natural key (a key that follows a natural classification), but they are more frequently realized in an artificial key. Simple artificial keys are preferred over complex natural keys. Classifications, not keys, are the vehicle for presenting hypotheses of relationship. The most convenient structured keys are arranged in a series of dichotomous choices, or couplets. Each member of a couplet is a lead. Two common types of keys likely to be seen in the literature are indented keys and bracket keys. Below we present examples of each type based on parts of an original key to identification of kingbirds taken from Brodkorb (1968).

Indented Key

   A. Belly white
       B. Tail with white tip …………………………………. *T. tyrannus*
      BB. Tail without white tip ………………………… *T. dominicensis*
  A.A. Belly yellow
       C. Exposed culmen equal to or greater than
          tarsus ………………………………………………. *T. melancholicus*
     CC. Exposed culmen shorter than tarsus
       D. Outer primaries shorter than sixth, outer
          web of lateral rectix brownish with narrow
          gray edging ………………………………………… *T. vociferans*
     DD. Outer primary longer than sixth, outer
          web of lateral rectix white to shaft ……………………… *T. verticalis*

Bracket Key

1a. Belly white …………………………………………………….. 2
  b. Belly yellow …………………………………………………… 3
2a. Tail with white tip ……………………………………. *T. tyrannus*
  b. Tail without white tip …………………………………. *T. dominicensis*
3a. Exposed culmen equal or shorter than tarsus …………… *T. melancholichus*
  b. Exposed culmen shorter than tarsus ………………………………… 4
4a. Outer primary shorter than 6th …………………………… *T. voliferans*
  b. Outer primary longer than 6th …………………………… *T. verticalis*

The dichotomous bracket key is by far the most common, and in our opinion, useful type of structured key. Indented keys have two practical difficulties: (1) the contrasting leads are frequently separated, making direct comparison of the alternate characters less convenient and (2) they take up a larger amount of page space. Some structured keys, especially those meant for nonspecialists, are liberally illustrated. Good dichotomous keys have several characteristics:

1. Each couplet is composed of strictly contrasting leads of two to three character contrasts.
2. The style is telegraphic.
3. The characters that are used are readily observable if at all possible.
4. Character contrasts should not call for value judgments ("lighter" or "darker" without reference to another shade on the same organism).
5. Some provision should be made for ages and sexes, or the key should be clearly labeled as only applicable for a particular sex or age group.
6. Characters that are apt to rely on expert judgment should be illustrated.

Polyclave keys are usually computer-based and associated with a dedicated software package, of which there are a number of both free and commercial products.

Modern polyclave keys are based on a matrix of taxa and characters from the total morphology of the species in the matrix. These keys have both advantages and disadvantages. Two advantages are recognized by Simpson (2006). First, because the character matrix draws from many morphological systems, one can use the key even if the specimen lacks more traditional key characters (such as floral morphology in plants) or is incomplete. Second, if the specimen cannot be identified to species, at least the choices can be narrowed. We would add a third: this kind of key appears to us to be pre-adapted to the Internet and can be constantly updated by editing the taxon-character matrix as new species and more morphological characters are discovered and studied. The major disadvantages are (1) unless it is a paper-based key you will need access to a computer and (2) there are not many such keys available.

BLAST: A "Polyclave Key" for Molecular Data. The kinds of keys discussed above are morphology-based. But what of nucleotide sequences? In most standard systematic work, molecular data are (or should be) associated with voucher specimens that can be identified through traditional taxonomic practices. However, in many groups there are now a sufficient number of molecules for some groups that have been sequenced so as to permit some level of identification of fragmentary specimens, larvae in search of matches with adults, and even commercial products (is this tuna or catfish?) such that some mention should be made of "DNA Keys." The major data repository is GenBank and the major tool is BLAST. A nucleotide sequence is submitted to the BLAST server (nucleotide blast on nucleotide collection) and sequences similar (or perhaps identical) that are in the database will be returned. It is then up to the investigator to explore the returned alignments and make a determination of identity. Of course, one can also query for gene identity or BLAST other sorts of molecular data.

THE RULES OF NOMENCLATURE

The history of proposing rules for the naming of organisms stretches back at least to Linnaeus' *Philosophia Botanica* (1751). A. P. de Candolle (1813) and Strickland (1842) proposed codes for plants and animals, respectively. By 1867 Alphonse de Candolle (son of A. P.) convened the First International Botanical Congress in Paris and the result was the "Paris Code" of 1867, but it was not until a number of such congresses that a truly international code was adopted in 1930. The current botanical code (International Code of Botanical Nomenclature, 2005) was adopted in Vienna. On the zoological side, various countries adopted their own codes of nomenclature (e.g., Strickland, 1842) until 1899 when a first truly international draft was proposed and later adopted in 1901 by the Fifth International Congress of Zoology. It was published in 1905 under the auspices of the Sixth Congress. The current version is the fourth edition (International Code of Zoological Nomenclature, 1999) published by the International Trust for Zoological Nomenclature. Bacteriologists chafed under the International Code of Botanical Nomenclature, and this resulted in the adoption of a separate Bacteriological Code (International Code of Nomenclature of Bacteria, 1990), whose history dates back to 1930 and the current revision was adopted in 1990 and published by the American Society for Microbiology in 1992. The last major code, the International Code of Virus Classification and Nomenclature, was adopted in 2002.

In addition to the four major codes, there are codes governing such things as cultivated plants. All four major codes are available on the Internet, so we shall not cover them in detail. The major principles adopted by the major codes are similar in spirit if not in wording.

1. Independence. Each is independent of the other. However, taxon coverage may overlap between the bacteriological codes and the botanical codes.

2. Stability. Three of the codes adopt the *Principle of Priority*, which states that the earliest use of the name of a group will be adopted unless petition is made to use a later name. In general, when two names are available for the same taxon, the older name has priority of use unless set aside by a decision by the appropriate commission. The code governing viruses is the exception.

3. Conservation of names. In the interests of stability, each code has provisions for the conservation of names and mechanisms to assure that names changes are minimized for names within the scope of the code.

4. New names. Each code specifies the rules for naming newly recognized or newly discovered taxa that are allocated to those ranks governed by the rules. These rules specifically cover what must be stated in terms of differentiating the new taxon, the form and ending of the name, the nature of the type, and the method of publication.

5. Names are Latin or Latinized. Because classic Latin is a dead language, the use of Latin names is language neutral.

6. Homonymy. The same name cannot be used for two taxa governed by the same code.

7. Scope. Each code specifies the taxonomic scope of their rules. For example, names above the family rank in zoology are not governed by the ICZN.

8. The basis of names. Types specimens (or cultures) are the basis of naming.

9. Retroactive. Each code is retroactive from the date of its last revision.

10. Each code states that its major function is to determine the application of names of taxa allocated to particular ranks and to leave to the community the biological interpretation of those names. Thus, the codes are, or attempt to be, "biology neutral."

Basic Nomenclatural Concepts

Priority   Priority is established by date of publication. Rules for effective publication are covered by each code and are likely to change as electronic publication becomes common. In general the date of publication is that date on which the printed description becomes available (default in the absence of evidence is the date appearing on the publication).

Correct Name and Valid Name   Different meaning in different codes. In botany the correct name is the one and only name that is used for a particular taxon. In zoology, the valid name is the one and only name that is used for a taxon. In botany a valid name is one that has been correctly published. In zoology, an available name is one that has been correctly published.

Synonyms   Names that are effectively published are available names (zoological term). Two available names for the same taxon are synonyms. Usually the older name is the correct/valid name, but if the taxon is split, the other name may be available for use depending on how the split is made. Names that are not effectively published are considered *nomen nuda* and are without meaning in both the zoological and the botanical codes.

Homonyms   The same name applied to two taxa are homonyms. The senior homonym is the earlier name, and the junior homonym is the later name. For example, *Clastes* Cope was proposed for a genus of garfishes, but it was already a name for a genus of spiders. In this case, the gar genus required a replacement name, *Clastichthyes* (Whitney, 1940). Codes differ on what constitutes a homonym, for example orthographic variants are not homonyms in zoology but are homonyms in botany.

Conserved Names (Nomen conservadum)   Although priority usually dictates which name is correct/valid, this can be set aside by a decision of the appropriate international commission. When this decision is made, the senior synonym is set aside and the selected junior synonym becomes the valid/correct name. The decision is usually made because the younger name has been in common usage. An often cited example is conservation of the name *Tyrannosaurus rex* Osborn, 1905 in place of *Manospondylus gigas* Cope, 1892. There are lists of both conserved and rejected names.

Limits of Priority   Priority in zoology and botany extends back in time to specific taxonomic works which date the beginning of naming for each code. In zoology nomenclature begins with edition 10 of Linnaeus' *Systema Naturae*, considered published on 1 January 1758. In botany the beginning of nomenclature differs for different groups. Most plant nomenclature, including algae, begin with the first edition of Linneaus' *Species Plantarum* (1 May 1753), but some nonvascular plants and all fungi have different start dates (consult the ICBN). Bacteriological names begin with *Species Plantarum* as with most plants. The virus code does not recognize priority as a principle.

Names and Name Endings   In general, all codes use binominal nomenclature for taxa ranked as species, singular nouns in the nominative case for generic names and plural nouns in the nominative case for those taxa ranked above the level of genus for those names governed by the codes. Name endings are specified for the various codes and the levels at which specific endings are mandatory differs between codes.

Types   Where the rules apply, the names of taxa are based on types, not on character properties or biological properties. Ultimately there is a specimen, illustration, "work of a specimen" or other such physical referent that can be examined. This does not make the system typological. The ultimate purpose of a type specimen is not to describe the variation or limits of variation of a species or other taxon, but to provide an example of what the original describer examined when he or she named the species. That is, it provides a direct physical link to the original describer or to a reviser who later designated the types that underlie a name. Thus when there

is a dispute regarding the name that should be applied to a species or whether two species are one or one species is two (or more), the reviser can examine at least one example of what the original describer thought was being named. If two names are based on the same type, then they are *objective* (zoology) or *nomenclatural* (botany) synonyms. If a species has been described several times with different types, then the various names are *subjective* (zoology) or *taxonomic* (botany) synonyms.

We end our brief discussion of the various rules by stressing, again, one of the primary reasons for their success. These rules are about the formation and use of names and how names can be published in a manner that is acceptable to the communities. They are not rules governing the biological meaning of the names of taxa. The fact is, we do not know whether the majority of named taxa are monophyletic, paraphyletic, or polyphyletic. The goal of the phylogenetic research program is to refer all valid/correct names of higher taxa to monophyletic groups and to refer all species binominals to evolving lineages, making nomenclature canonical with the best evolutionary theory possible.

CHAPTER SUMMARY

- Systematics advances through the publication of descriptions, revisions, and phylogenetic analyses.
- There are a variety of kinds of publications that serve as outlets for systematic research.
- Taxonomic scholarship is aided by numerous kinds of literature and literature summaries that are specific to the particular field.
- Publication of systematic research usually follows a set format, and this format is discipline or journal specific.
- The formation and use of names are governed by codes of nomenclature that are specific for particular groups of organisms and independent of other codes.
- Codes do not cover the biological meanings of names.

LITERATURE CITED

Abdel Ghany, A. G. A., and E. A. Zaki. 2003. Isolation, characterization and phylogenetic analysis of copia-like retrotransposons in the Egyptian colon *Gossypium barbadense* and its progenitors. *African J. Biotech.* 2:165–168.

Abe, F. R., and B. S. Lieberman. 2009. The nature of evolutionary radiations: A case study involving Devonian trilobites. *Evol. Biol.* 36:225–234.

Adams, B. J. 1998. Species concepts and the evolutionary program in modern nematology. *J. Nematology* 30:1–21.

Adams, D. C., F. J. Rohlf, and D. E. Slice. 2004. Geometric morphometrics: Ten years of progress following the "revolution." *Italian J. Zool.* 71:5–16.

Adams, E. N., III. 1972. Consensus techniques and the comparison of taxonomic trees. *Syst. Zool.* 21:390–397.

Adanson, M. 1763. *Familles des Plantes*. Vincent, Paris.

Adrain, J. M., G. D. Edgecombe, and B. S. Lieberman (eds.). 2001. *Fossils, Phylogeny, and Form*. Kluwer Academic/Plenum, New York.

Agapow, P. M., O. R. P. Bininda-Edmonds, K. A. Crandall, J. L. Gittleman, G. M. Mace, et al. 2004. The impact of species concept on biodiversity studies. *Q. Rev. Biol.* 79:161–79.

Agassiz, L. R. 1833–1844. *Recherches sur les Poissons Fossiles*. Imprimérie de Petitpierre, Nauchâtel.

Agassiz, L. 1846 and 1848. *Nomenclatoris Zoologici*. Index Universalis, Soloduri, Sumotibus Jent et Gassman.

Aguilar, J. F., J. A. Roselló, and G. N. Feliner. 1999. Molecular evidence for the compilospecies model of reticulate evolution in *Armeria* (Plumbaginaceae). *Syst. Biol.* 48:735–754.

Aguilar-Aguilar, R., R. Contreras-Mdina, and G. Salgado-Maldonado. 2003. Parsimony analysis of endemicity (PAE) of Mexican hydrological basins based on helminth parasites of freshwater fishes. *J. Biogeography* 30:1861–1872.

Ahlberg, P. E., Z. Johanson, and E. B. Daeschler. 2001. The late Devonian lungfish *Soederberghia* (Sarcopterygii, Dipnoi) from Australia and North America, and its biogeographical implications. *J. Vertebrate Paleontol.* 21:1–12.

Akaike, H. 1973. Information theory as an extension of the maximum likelihood principle. In: *Second International Symposium on Information Theory* (B. Petrov and F. Csaki, eds.). Akademiai Kiado, Budapest: 267–281.

Alexander, H. J., J. S. Taylor, S. Sze-Tsun Wu, and F. Breden. 2006. Parallel evolution and vicariance in the guppy (*Poecilia reticulata*) over multiple spatial and temporal scales. *Evolution* 60:2352–2369.

Almeida, M. T., and F. A. Bisby. 1984. A simple method for establishing taxonomic characters from measurement data. *Taxon* 33:405–409.

Anderson, F. E., and D. L. Swofford. 2004. Should we be worried about long-branch attraction in real data sets? Investigations using metazoan 18S rDNA. *Mol. Phylogen. Evol.* 33:440–451.

Anderson, J. S. 2001. The phylogenetic trunk: Maximal inclusion of taxa with missing data in an analysis of the Lepospondyli (Vertebrata, Tetrapoda). *Syst. Biol.* 50:170–193.

Archibald, J. D. 2009. Edward Hitchcock's pre-Darwinian (1840) "Tree of Life." *J. Hist. Biol.* 42:561–592.

Archie, J. W. 1985. Methods for coding variable morphological features for numerical taxonomic analysis. *Syst. Zool.* 34:326–345.

Archie, J. W. 1989. A randomization test for phylogenetic information in systematic data. *Syst. Zool.* 38:239–252.

Arnold, E. N. 1981. Estimating phylogenies at low taxonomic levels. *Z. Zool. Syst. Evolut.-forsch.* 19:1–35.

Arratia, G. 1999. The monophyly of Teleostei and setm-group teleosts. In: *Mesozoic Fishes 2. Systematics and Fossil Record* (G. Arratia and H. P. Schultze, eds.). Verlag Dr. Friedrich Pfeil, München, Germany: 265–334.

Ashlock, P. D. 1971. Monophyly and associated terms. *Syst. Zool.* 20:63–69.

Ashlock, P. D. 1972. Monophyly again. *Syst. Zool.* 21:430–438.

Ashlock, P. D. 1985. A revision of the *Bergidea* group: A problem in classification and biogeography (Hemiptera-Heteroptera: Lygaeidae). *J. Kansas Entomol. Soc.* 57:675–688 (dated 1984).

Avise, J. C. 1992. Molecular population structure and the biogeographic history of a regional fauna: A case history with lessons for conservation biology. *Oikos* 63:62–76.

Avise J. C. 2000. *Phylogeography: The History and Formation of Species*. Harvard University Press, Cambridge, MA.

Avise, J. C., and R. M. Ball. 1990. Principles of genealogical concordance in species concepts and biological taxonomy. *Oxford Surv. Evol. Biol.* 7:45–67.

Avise, J. C., D. Walker, and G. C. Johns. 1998. Speciation durations and Pleistocene effects on vertebrate phylogeography. *Proc. Royal Soc.* London, Series B 265:1707–1712.

Ax, P. 1987. *The Phylogenetic System: The Systematization of Organisms on the Basis of Their Phylogenesis*. Wiley-Interscience, New York.

Balinsky, B. I. 1970. *An Introduction to Embryology*. W. B. Saunders, Philadelphia.

Barkley, T. M., P. DePriest, V. Funk, R. W, Kiger, W. J. Kress, J. McNeill, J. Moore, D. H. Nicolson, D. W. Stevenson, and Q. D. Wheeler. 2004a. A review of the International Code of Botanical Nomenclature with respect to its compatibility with phylogenetic classification. *Taxon* 53: 159–161.

Barkley, T. M., P. DePriest, V. Funk, R. W. Kiger, W. J. Kress, and G. Moore. 2004b. Linnean nomenclature in the 21st century: A report from a workshop on integrating traditional nomenclature and phylogenetic classification. *Taxon* 53: 153–158.

Barrett, M., M. J. Donoghue, and E. Sober. 1991. Against consensus. *Syst. Zool.* 40:486–493.

Barrett, P. H., P. J. Gautrey, S. Herbert, D. Kohn, and S. Smith (eds.). 1987. *Charles Darwin's Notebooks, 1836–1844*. Cornell University Press, Ithaca, NY.

Bartram, A. W. H. 1977. The Microsemiidae, a Mesozoic family of holostean fishes. *Bull. British Mus. Nat. Hist. (Geol.)* 2a(2):137–234.

Bass, D., T. A. Richards, L. Matthai, V. Marsh, and T. Cavalier-Smith. 2007. DNA evidence for global dispersal and probable endemicity of protozoa. *BMC Evol Biol.* 7(1):162 17854485.

Baum, B. R. 1988. A simple procedure for establishing discrete characters from measurement data, applicable to cladistics. *Taxon* 37:63–70.

Baum, D. A., and K. L. Shaw. 1995. Genealogical perspectives on the species problem. In: *Experimental and Molecular Approaches to Plant Biosystematics* (P. C. Hoch and A. G. Stephenson, eds.). Monogr. Syst, Missouri Bot. Gard. 53:289–303.

Bealer, G. 1999. Property. In: *The Cambridge Dictionary of Philosophy*. 2nd edition (R. Audi, ed.). Cambridge University Press, New York: 751–752.

Beard, K. C. 1998. East of Eden: Asia as an important center of taxonomic origination in mammalian evolution. *Bull. Carnegie Mus. Nat. Hist*. 34:5–39.

Beard, K. C. 2002. East of Eden at the Paleocene/Eocene boundary. *Science* 295:2028–2029.

Beatty, J., and W. L. Fink. 1979. Review of: *Simplicity* by E. Sober. *Syst. Zool.* 28:643–65l.

Beddard, F. E. 1895. *A Textbook of Zoogeography*. Cambridge University Press, Cambridge, UK.

Bely, A. E., and G. A. Wray. 2001. Evolution of regeneration and fission in annelids: Insights from *engrailed-* and *orthodenticle*-class gene expression. *Development* 128:2781–2791.

Bennett, K. D. 1997. *Evolution and Ecology: The Pace of Life*. Cambridge University Press, Cambridge, UK.

Bermingham, E., and C. Moritz. 1998. Comparative phylogeography: Concepts and applications. *Mol. Ecol.* 7:367–369.

Bininda-Emonds, O. R. P. (ed.). 2004. *Phylogenetic Supertrees: Combining Information to Reveal the Tree of Life*. Computational Biology 4.

Bisconti, M., W. Landini, G. Bianucci, G. Cantalamessa, G. Carnevale, et al. 2001. Biogeographic relationships of the Galapagos terrestrial biota: Parsimony analyses of endemicity based on reptiles, land bards and *Scalesia* land plants. *J. Biogeog.* 28:495–510.

Blackburn, T. M., and K. J. Gaston. 1998. Some methodological issues in macroecology. *Am. Nat.* 51:6814–6883.

Blackwelder, R. E. 1967. *Taxonomy: A Text and Reference Book*. John Wiley & Sons, New York.

Blackwelder, R. E. 1972. *Guide to Taxonomic Literature of Vertebrates*. Iowa State University, Ames, IA.

Bock, W. J. 1969. The concept of homology. *Ann. New York Acad. Sci.* 167:111–115.

Bock, W. J. 1974. Philosophical foundations of classical evolutionary taxonomy. *Syst. Zool.* 22:375–392.

Bonde, N. 1977. Cladistic classification as applied to vertebrates. In: *Major Patterns in Vertebrate Evolution* (M. K. Hecht, P. C. Goody, and B. M. Hecht, eds.). Plenum Press, New York: 741–804.

Bookstein, F. L. 1991. *Morphometric Tools for Landmark Data: Geometry and Biology*. Cambridge University Press, New York.

Bookstein, F. L. 1994. Can biometrical shape be a homologous character? In: *Homology: The Hierarchical Basis of Comparative Biology* (B. K. Hall, ed.). Academic Press, San Diego: 197–227.

Borror, D. J. 1960. *Dictionary of Word Roots and Combining Forms*. Mayfield, Mountain View, CA.

Borror, D. J., D. W. Delong, and C. A. Triplehorn. 1976. *An Introduction to the Study of Insects*. Holt, Rinehart and Winston, New York.

Bowler, P. J. 1996. *Life's Splendid Drama*. University of Chicago Press, Chicago.

Boyd, R. 1991. Realism, antifoundationalism, and the enthusiasm for natural kinds. *Philosophical Studies* 61:127–148.

Boyd, R. 1999. Homeostasis, species, and higher taxa. In: *Species: New Interdisciplinary Essays* (R. A. Wilson, ed.). MIT Press, Cambridge, MA: 141–185.

Bremer, K. 1978a. The genus *Leysera* (Compositae). *Botaniska Notiser* 131:369–383.

Bremer, K. 1978b. *Oreoleysera* and *Antithrixia,* new and old South African genera of the Compositae. *Botaniska Notiser* 131:449–453.

Bremer, K. 1988. The limits of amino acid sequence data in angiosperm phylogenetic reconstruction. *Evolution* 42:795–803.

Bremer, K. 1990. Combinable component consensus. *Cladistics* 6:369–372.

Bremer, K. 1992. Ancestral areas: A cladistic reinterpretation of the center of origin concept. *Syst. Biol.* 41:436–445.

Bremer, K. 2002. Gondwanan evolution of the grass alliance of families (Poales). *Evolution* 56:1374–1387.

Bremer, K., and H.-E. Wanntorp. 1979. Hierarchy and reticulation in systematics. *Syst. Zool.* 28:624–627.

Brett, C. E., and G. C. Baird. 1995. Coordinated stasis and evolutionary ecology of Silurian to Middle Devonian faunas in the Appalachian Basin. In: *New Approaches to Speciation in the Fossil Record* (D. H. Erwin and R. L. Anstey, eds.). Columbia University Press, New York: 285–315.

Bridgman, P. 1927. *The Logic of Modern Physics*. McMillan, New York.

Briscoe, H. M. 1995. *Preparing Scientific Illustrations*. Springer-Verlag, New York.

Brodkorb, P. 1968. *Part five – Birds. Vertebrates of the United States*. McGraw-Hill, New York.

Brooks, D. R. 1981. Hennig's parasitological method: A proposed solution. *Syst. Zool.* 30:229–249.

Brooks, D. R. 1985. Historical ecology: A new approach to studying the evolution of ecological associations. *Ann. Missouri Bot. Gard.* 72:660–680.

Brooks, D. R. 1988. Scaling effects in historical biogeography: A new view of space, time, and form. *Syst. Zool.* 37:237–244.

Brooks, D. R. 1990. Parsimony analysis in historical biogeography and coevolution: Methodological and theoretical update. *Syst. Zool.* 39:14–30.

Brooks, D. R., and A. Ferrao. 2005. The historical biogeography of co-evolution: Emerging infectious diseases are evolutionary accidents waiting to happen. *J. Biogeography* 32:1291–1299.

Brooks, D. R., and K. E. Folinsbee. 2005. Paleobiogeography: Documenting the ebb and flow of evolutionary diversification. In: *Paleobiogeography: Generating New Insights into the Coevolution of the Earth and Its Biota* (B. Lieberman and A. L. Stigall Rode, eds.). Paleontological Society Special Papers No. 11, Yale University Press, New Haven, CT: 15–44.

Brooks, D. R., and D. A. McLennan. 1991. *Phylogeny, Ecology, and Behavior: A Research Program in Comparative Biology*. University of Chicago Press, Chicago.

Brooks, D. R., and D. A. McLennan. 2002. *The Nature of Diversity*. University of Chicago Press, Chicago.

Brooks, D. R., T. B. Thorson, and M. A. Mayes. 1981. Fresh-water stingrays (Potamotrygoni-dae) and their helminth parasites: Testing hypotheses of evolution and coevolution. In: *Advances in Cladistics* (V. A. Funk and D. R. Brooks, eds.). New York Botanical Garden, New York: 149–175.

Brooks, D. R., and M. G. P. Van Veller. 2003. Critique of parsimony analysis of endemicity as a method of historical biogeography. *J. Biogeography* 30:819–825.

Brooks, D. R., and E. O. Wiley. 1986. *Evolution as Entropy: Toward a Unified Theory of Biology*. University of Chicago Press, Chicago.

Brooks, J. L. 1984. *Just Before the Origin*. Columbia University Press, New York.

Brower A. V. Z. 1999. Delimitation of phylogenetic species with DNA sequences: A critique of Davis and Nixon's population aggregation analysis. *Syst. Biol.* 48:199–213.

Brown, J. H., and M. V. Lomolino. 1998. *Biogeography*, 2nd edition. Sinauer, Sunderland, MA.

Brown, J. H., and B. A. Maurer. 1989. Macroecology: The division of food and space among species on continents. *Science* 243:1143–1150.

Brown, R. W. 2000. *Composition of Scientific Words*. Smithsonian Institution Press, Washington, DC.

Browne, J. 1983. *The Secular Ark: Studies in the History of Biogeography*. Yale University Press, New Haven, CT.

Brundin, L. 1966. Transantarctic relationships and their significance, as evidenced by chironoimid midges. *Kungliga Svenska Vetenskapsakadamien Handlingar* 4(11):1–472.

Brundin, L. Z. 1988. Phylogenetic biogeography. In: *Analytical Biogeography* (A. A. Myers and P. S. Giller, eds.). Chapman and Hall, New York: 343–369.

Bryant, D. 2003. A classification of consensus methods for phylogenies. In *BioConsensus* (M. Janowitz, F. J. Lapointe, F. R. McMorris, B. Mirkin, and F. S. Roberts, eds.). DIMACS Series in Discrete Mathematics and Theoretical Computer Science, Vol. 61. Amer. Math. Soc. Providence, RI: 163–184.

Buch, L. von. 1825. *Physicalische Beschreibung der Cancrischen Inseln*. Koeniglichen Akademie der Wissenschaften, Berlin.

Buck, R., and D. Hull. 1966. The logical structure of the Linnean Hierarchy. *Syst. Zool.* 15:97–111.

Buffon, Comte de, G. L. L. 1749. *Histoire Naturelle*, Vol. 1. Paris.

Burton, R. S. 1998. Intraspecific phylogeography across the Point Conception biogeographic boundary. *Evolution* 52:734–745.

Bush, G. L. 1975. Models of animal speciation. *Ann. Rev. Ecol. Syst.* 6:339–364.

Cain, A. J. 1958. Logic and memory in Linnaeus's system of taxonomy. *Proc. Linn. Soc. London* 169:144–163.

Cain, A. J., and G. A. Harrison. 1958. An analysis of the taxonomist's judgment of affinity. *Proc. Zool. Soc. London* 131:85–98.

Camin, J. H., and R. R. Sokal. 1965. A method for deducing branching sequences in phylogeny. *Evolution* 19:311–326.

Candolle, A. P., de. 1813. *Théorie Elémentaire de la Botanique*. Detréville, Paris.

Cannatella, D. C., and K. de Queiroz. 1989. Phylogenetic systematics of the anoles: Is a new taxonomy warranted? *Syst. Zool.* 38:57–69.

Carine, M. A., and R. W. Scotland. 1999. Taxic and transformational homology: Different ways of seeing. *Cladistics* 15:121–129.

Carpenter, J. M. 1988. Choosing among equally parsimonious cladograms. *Cladistics* 4:291–296.

Carpenter, J. M. 1994. Successive weighting, reliability and evidence. *Cladistics* 10:215–220.

Carpenter, J. M. 2003. Critique of pure folly. *Bot. Rev.* 69:79–92.

Carpenter, J. M., J. E. Strassmann, S. Turillazzi, C. R. Hughes, C. R. Solis, and R. Cervo. 1993. Phylogenetic relationships among paper wasp social parasites and their hosts (Hymenoptera: Vespidae: Polistinae). *Cladistics* 9:129–146.

Carson, H. L. 1957. The species as a field for recombination. In: *The Species Problem* (E. Mayr, ed.). AAAS, Washington, DC: 23–38.

Carsten, B. C., and C. L. Richards. 2007. Integrating coalescent and ecological niche modeling in comparative phylogeography. *Evolution* 61:1439–1454.

Cavalli-Sforza, L. L., and A. W. F. Edwards. 1967. Phylogenetic analysis: Models and estimation procedures. *Evolution* 32:550–570.

Chan, F. Y., J. Robinson, A. Brownlie, R. A. Shivdasani, A. Donovan, C. Brugnara, J. Kim, B. C. Lau, H. E. Witkowska, and L. I. Zon. 1997. Characterization of adult alpha- and beta-globin genes in the zebrafish. *Blood* 89:688–700.

Chappill, J. A. 1989. Quantitative characters in phylogenetic analysis. *Cladistics* 5:217–234.

Chase, M. W., D. E. Soltis, R. G. Olmstead, D. Morgan, D. H. Les, B. D. Mishler, et al. 1993. Phylogenetics of seed plants: An analysis of nucleic sequences from the plastid gene *rbc. L. Ann. Missouri Bot. Gard.* 80:528–580.

Chatterjee, H. J. 2006. Phylogeny and biogeography of gibbons: A dispersal-vicariance analysis. *Intl. J. Primatol.* 27:699–712.

Chatzimanolis, S., and M. S. Caterino. 2007. Limited phylogeographic structure in the flightless ground beetle, *Calathus ruficollis*, in southern California. *Diversity and Distributions* 13:498–509.

Chesser, R. T., and R. M. Zink. 1994. Modes of speciation in birds: A test of Lynch's method. *Evolution* 48:490–497.

Chin-Yih, Ou., C. A. Ciesielski , G. Myers, C. I. Bandea , C.-C. Luo, B. T. M. Korber , J. I. Mullins, G. Schochetman, R. L. Berkelman, A. N. Economou, J. J. Witte, L. J. Furman, G. A. Satten, K. A. Maclnnes, J. W. Curran, and H. W. Jaffe. 1992. Molecular epidemiology of HIV transmission in a dental practice. *Science* 256(5060):1165–1171.

Claridge, M. F., H. A. Dawah, and M. R. Wilson. 1997. *Species: The Units of Diversity*. Chapman and Hall, New York.

Coleman, K. A., and E. O. Wiley. 2001. On species individualism: A new defense of the species-as-individuals hypothesis. *Philo. Sci.* 68:498–517.

Colless, D. H. 1980. Congruence between morphological and allozyme data for *Menidia* species: A reappraisal. *Syst. Zool.* 29:288–299.

Collier, G. E. 1979. Synopsis of the genus *Fundulus* (Cyprinodontidae: Pisces). *J. Amer. Killifish Assoc.* 12:2–14.

Congreve, C. R., and B. S. Lieberman. 2010. Phylogenetic and biogeographic analysis of deiphonine trilobites. *J. Paleontol.* 84:128–136.

Conti, E., T. Eriksson, J. Schonenberger, K. J. Systsma, and D. A. Baum. 2002. Early Tertiary out of-India dispersal of Crypteroniaceaea: Evidence from phylogeny and molecular dating. *Evolution* 56:1931–1942.

Coope, G. R. 1979. Late Cenozoic Coleoptera: Evolution, biogeography, and ecology. *Ann. Rev. Ecol. Syst.* 10:247–267.

Coyne, J. A., and A. Orr. 2004. *Speciation*. Sinauer Assoc., Sunderland, MA.

Cracraft, J. 1974. Phylogenetic models of classification. *Syst. Zool.* 23:71–90.

Cracraft, J. 1982. Geographic differentiation, cladistics and vicariance biogeography: Reconstructing the tempo and mode of evolution. *Amer. Zool.* 22:411–424.

Cracraft, J. 1983. Species concepts and speciation analysis. *Current Ornithol.* 1:159–187.

Cracraft, J. 1988. Deep-history biogeography: Retrieving the historical pattern of evolving continental biotas. *Syst. Zool.* 37:221–236.

Cranston, P. S., and C. J. Humphries. 1988. Cladistics and computers: A chironomid conundrum? *Cladistics* 4:72–92.

Cronquist, A. 1978. Once again, what is a species? In: *Biosystematics in Agriculture* (J. A. Romberger, ed.). Allenheld, Osman & Company, Montclair, NJ: 3–20.

Croizat, L. 1964. *Space, Time, and Form, the Biological Synthesis*. Published by the author, Caracas.

Croizat, L., G. Nelson, and D. E. Rosen. 1974. Centers of origin and related concepts. *Syst. Zool.* 23:265–287.

Crowson, R. A. 1970. *Classification and Biology*. Atherton Press, New York.

Cunningham, C. W. 1997. Is congruence between data partitions a reliable predictor of phylogenetic accuracy? Empirically testing an iterative procedure for choosing among phylogenetic methods. *Syst. Biol.* 46:464–478.

Dahn, R. D., M. C. Davis, W. N. Pappano, and N. H. Shubin. 2007. Sonic hedgehog function in chondrichthyan fins and the evolution of appendage patterning. *Nature* 445:311–314.

Darlington, P. J., Jr. 1959. Area, climate, and evolution. *Evolution* 13:488–510.

Darwin, C. 1859. *On the Origin of Species by Means of Natural Selection; or the Preservation of Favored Races in the Struggle for Life*. Reprinted 1st edition. Cambridge, MA: Harvard University Press.

Darwin, C. 1872. *On the Origin of Species by Means of Natural Selection; or the Preservation of Favored Races in the Struggle for Life*. Reprinted 6th edition. New American Library, New York.

Daubin V., N. A. Moran, and H. Ochman. 2003. Phylogenetics and the cohesion of bacterial genomes. *Science* 301(5634):829–832.

Davis, J. I., and K. C. Nixon. 1992. Populations, genetic variation, and the delimitation of phylogenetic species. *Syst. Biol.* 41:421–435.

Davis, M. B. 1983. Quaternary history of deciduous forsts of eastern North America and Europe. *Ann. Missouri Bot. Gard.* 70:550–563.

Davis, M. C., R. D. Dahn, and N. H. Shubin. 2007. An autopodial-like pattern of Hox expression in the fins of a basal actinopterygian fish. *Nature* 447:473–477.

Davis, P. H., and V. H. Heywood. 1965. *Principles of Angiosperm Taxonomy*. P. Van Nostrand, New York.

de Candolle, A. 1813. *Théorie Élémentaire de la Botanique*. Déterville, Paris.

Degnan, J. H., and N. A. Rosenberg. 2009. Gene trees discordance, phylogenetic inference and the multispecies coalescent. *Trends Ecol. Evol.* 24:332–340.

de Grave, S. 2001. Biogeography of Indo-Pacific Potoniinae (Crustacea, Decapoda): A PAE analysis. *J. Biogeogr.* 28:1239–1254.

de Jussieu, A. L. 1789. *Genera Plantarum*. Herissant, Paris.

de Queiroz, K. 1988. Systematics and the Darwinian revolution. *Phil. Sci.* 55:238–259.

de Queiroz, K. 1995. The definition of species and clade names: A reply to Ghiselin. *Biol. Philo.* 10:223–228.

de Queiroz, K. 1997. The Linnean Hierarchy and the evolutionization of taxonomy, with emphasis on the problem of nomenclature. *Aliso* 15(2):125–144.

de Queiroz, K. 1998. The general lineage concept of species, species criteria, and the process of speciation. In: *Endless Forms: Species and Speciation* (D. J. Howard and S. H. Berlocher, eds.). Oxford University Press, Oxford, UK: 57–75.

de Queiroz, K., and M. J. Donoghue. 1988. Phylogenetic systematics and the species problem. *Cladistics* 4:317–338.

de Queiroz, K., and M. J. Donoghue. 1990. Phylogenetic systematics and species revisited. *Cladistics* 6:83–90.

de Queiroz, K., and J. Gauthier. 1992. Phylogenetic taxonomy. *Ann. Rev. Ecol. Syst.* 23:449–480.

de Queiroz, K., and J. Gauthier. 1994. Toward a phylogenetic system of biological nomenclature. *Trends Ecol. Evol.* 9:27–31.

de Pinna, M. C. 1991. Concepts and tests of homology in the cladistic paradigm. *Cladistics* 7:367–394.

de Pinna, M. 1996. Teleostean monophyly. In: *Interrelationships of Fishes* (M. L. J. Stiasney, L. R. Parenti, and G. D. Johnson, eds.). Academic Press, San Diego:147–162.

Dickerson, R. E., and I. Geis. 1983. *Hemoglobin: Structure, Function, Evolution and Pathology*. Benjamin/Cummings, Menlo Park, CA.

Dobzhansky, T. 1935. A critique of the species concept in biology. *Philo. Sci.* 2:344–355.

Dobzhansky, T. 1937. *Genetics and the Origin of Species*. Columbia University Press, New York.

Dobzhansky, T. 1962. *Mankind Evolving. The Evolution of the Human Species*. Yale University Press. New Haven, CT.

Dobzhansky, T. 1970. *Genetics of the Evolutionary Process*. Columbia University Press, New York.

Donoghue, M. J. 1985. A critique of the biological species concept and recommendations for a phylogenetic alternative. *The Bryologist* 88:172–181.

Donoghue, M. J., J. A. Doyle, J. Gauthier, A. G. Kluge, and T. Rowe. 1989. The importance of fossils in phylogeny reconstruction. *Ann. Rev. Ecol. Syst.* 20:431–460.

Donoghue, M. J., and B. R. Moore. 2003. Toward an integrative historical biogeography. *Integrative and Comparative Biology* 43:261–270.

Doolittle, R. F. (ed.). 1990. *Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences*. Academic Press, San Diego.

Doyle, J. 1995. The irrelevance of allele tree topologies for species delimitation, and a non-topological alternative. *Syst. Bot.* 20:574–588.

Drovetski, S. V. 2003. Plio-Pleistocene climatic oscillations, Holarctic biogeography and speciation in an avian subfamily. *Journal of Biogeography* 30:1173–1181.

Ebach, M. C., C. J. Humphries, and D. M. Williams. 2003. Phylogenetic biogeography deconstructed. *Journal of Biogeography* 30:1285–1296.

Echelle, M. M., M. Stevenson, A. F. Echelle, and L. G. Hill. 1972. Diural periodicity in the plains killifish *Fundulus zebrinus kansae*. *Proc. Okla. Acad. Sci.* 51:3–7.

Edgar, R. C. 2004. MUSCLE: A multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 5(1):113.

Ee, B. W. van, and P. E. Berry. 2009. A phylogenetic and taxonomic review of *Croton* (Euphorbiaceae s.s.) on Jamaica including the description of *Croton jamaicensis*, a new species of section Eluteria. *Syst. Bot.* 34:129–140.

Eernisse, D. J., and A. G. Kluge. 1993. Taxonomic congruence versus total evidence, and amniote phylogeny inferred from fossils, molecules, and morphology. *Mol. Biol. Evol.* 10:170–195.

Eidesen, P. B., T. Carlsen, U. Molau, and C. Brochmann. 2007. Repeatedly out of Beringia: *Cassiope tetragona* embraces the Arctic. *J. Biogeogr.* 34:1559–1574.

Eldredge, N. 1971. The allopatric model and phylogeny in Paleozoic invertebrates. *Evolution* 25:156–167.

Eldredge, N. 1973. Systematics of lower and lower middle Devonian species of the trilobite *Phacops* Emmrich in North America. *Bull. Am. Mus. Nat. Hist.* 151:285–338.

Eldredge, N. 1979. Alternative approaches to evolutionary theory. In: *Models and Methodologies in Evolutionary Theory* (J. H. Schwartz and H. B. Rollins, eds.). *Bull. Carnegie Mus. Nat. Hist.* 13:7–19.

Eldredge, N. 1985. *Unfinished Synthesis: Biological Hierarchies and Modern Evolutionary Thought*. Oxford University Press, New York.

Eldredge, N. 1986. Information, economics, and evolution. *Ann. Rev. Ecol. Syst.* 17:351–369.

Eldredge, N. 1989. *Macroevolutionary Dynamics*. McGraw-Hill, New York.

Eldredge, N. 1993. What, if anything, is a species? In: *Species, Species Concepts, and Primate Evolution*. (W. H. Kimble and L. B. Mastin, eds.). Plenum Press, New York: 3–20.

Eldredge, N. 1998. *Life in the Balance*. Princeton University Press, Princeton, NJ.

Eldredge, N., and J. Cracraft. 1980. *Phylogenetic Patterns and the Evolutionary Process*. Columbia University Press, New York.

Eldredge, N., and S. J. Gould. 1972. Punctuated equilibria: An alternative to phyletic gradualism. In: *Models in Paleobiology* (T. J. M. Schopf, ed.). Freeman Cooper, San Francisco: 82–115.

Eldredge N., and S. N. Salthe. 1984. Hierarchy and evolution. *Oxf. Surv. Evol. Biol.* 1:182–206.

Eldredge, N., and I. Tattersall. 1975. Evolutionary models, phylogenetic reconstruction, and another look at hominid phylogeny. In: *Approaches to Primate Paleobiology* (F. S. Szalay, ed.). Karger, Basel: 218–242.

Eldredge, N., J. Thompson, P. Brakefield, S. Gavrilets, D. Jablonski, J. Jackson, R. Lenski, B. S. Lieberman, M. McPeek, and W. Miller III. 2005. The dynamics of evolutionary stasis. *Paleobiol.* 31:133–145.

Elton, C. S. 1958. *The Ecology of Invasions by Animals and Plants*. Methuen, London, UK.

Endler, J. A. 1977. *Geographic Variation, Speciation, and Clines*. Princeton University Press, Princeton, NJ.

Endler, J. A. 1989. Conceptual and other problems in speciation. In: *Speciation and Its Consequences* (D. Otte and J. A. Endler, eds.). Sinauer Associates, Sunderland, MA: 625–648.

Enghoff, H. 1995. Historical biogeography of the Holarctic: Area relationships, ancestral areas, and dispersal of non-marine animals. *Cladistics* 11:223–263.

Englemann, G. F., and E. O. Wiley. 1977. The place of ancestor-descendant relationships in phylogeny reconstruction. *Systematic Zoology* 26:1–11.

Ereshefsky, M. 2001. *The Poverty of the Linnean Hierarchy: A Philosophical Study of Biological Taxonomy*. Cambridge University Press, Cambridge.

Erwin, T. L. 1979. Thoughts on the evolutionary history of ground beetles: Hypotheses generated from comparative faunal analyses of lowland forest sites in temperate and tropical regions. In: *Carabid Beetles – Their Evolution, Natural History, and Classification* (T. L. Erwin, G. E. Ball, D. R. Whitehead, eds.). W. Junk, The Hague, Netherlands: 539–592.

Erwin, T. L. 1981. Taxon pulses, vicariance, and dispersal: An evolutionary synthesis illustrated by carabid beetles. In: *Vicariance Biogeography: A Critique* (G. Nelson, D. E. Rosen, eds.). Columbia University Press, New York: 159–183.

Eschmeyer, W. N., and R. Fricke (eds.). 2009 et seq. *Catalog of Fishes* electronic version; http://research.calacademy.org/ichthyology/catalog/fishcatsearch.html.

Faith, D. P. 1991. Cladistic permutation tests for monophyly and nonmonophyly. *Syst. Zool.* 40:366–375.

Faith, D. P., and J. W. O. Ballard. 1994. Length differences and topology-dependent tests: A response to Källersjö et al. *Cladistics* 10:57–64.

Faith, D. P., and P. S. Cranston. 1991. Could a cladogram this short have arisen by chance alone? On permutation tests for cladistic structure. *Cladistics* 7:1–28.

Farr, E. R., J. A. Leussink, and F. A. Stafleu (eds.). 1979. Index Nominum Genericorum (Plantarum). *Regnum Veg.* 100–102:1–1896.

Farris, J. S. 1969. A successive approximations approach to character weighting. *Syst. Zool.* 18:374–385.

Farris, J. S. 1970. Methods for computing Wagner trees. *Syst. Zool.* 19:83–92.

Farris, J. S. 1973. A probability model for inferring evolutionary trees. *Syst. Zool.* 22:250–256.

Farris, J. S. 1974. Formal definitions of paraphyly and polyphyly. *Syst. Zool.* 23:548–554.

Farris, J. S. 1976. Phylogenetic classification of fossils with recent species. *Syst. Zool.* 25:271–282.

Farris, J. S. 1977. On the phenetic approach to vertebrate classification. In: *Major Patterns in Vertebrate Evolution* (M. K. Hecht, P. C. Goody, and B. M. Hecht, eds.). Plenum Press, New York: 823–850.

Farris, J. S. 1980. The efficient diagnoses of the phylogenetic system. *Syst. Zool.* 29:386–401.

Farris, J. S. 1983. The logical basis of phylogenetic analysis. In: *Advances in Cladistics* proceedings of the second meeting of the Willi Hennig Society (N. I. Platnick and V. A. Funk, eds.). Columbia University Press, New York: 1–36.

Farris, J. S. 1988. *HENNIG86 Reference: Documentation for Version 1.5*. Port Jefferson Station, New York.

Farris, J. S. 1989a. Hennig86: A PC-DOS program for phylogenetic analysis. *Cladistics* 5:163.

Farris, J. S. 1989b. The retention index and the rescaled consistency index. *Cladistics* 5:417–419.

Farris, J. S. 1999. Likelihood and consistency. *Cladistics* 15:199–204.

Farris, J. S. 2010. Systemati foundering. *Cladistics* 26:1–15.

Farris, J. S., M. Källersjö, V. A. Albert, M. Allard, A. Anderberg, B. Bowditch, C. Bult, J. M. Carpenter, T. M. Crowe, J. De Laet, K. Fitzhugh, D. Frost, P. Goloboff, C. J. Humphries, U. Jondelius, D. Judd, P. O. Karis, D. Lipscomb, M. Luckow, D. Mindell, J. Muona, K. Nixon, W. Presch, O. Seberg, M. E. Siddall, L. Struwe, A. Tehler, J. Wenzel, Q. Wheeler, and W. Wheeler. 1995. Explanation. *Cladistics* 11:211–218.

Farris, J. S., M. Källersjö, A. G. Kluge, and C. Bult. 1994a. Permutations. *Cladistics* 10:65–76.

Farris, J. S., M. Källersjö, A. G. Kluge, and C. Bult. 1994b. Testing significance of incongruence. *Cladistics* 10:315–319.

Farris, J. S., V. A. Albert, M. Källersjö, D. Lipscomb, and A. G. Kluge. 1996. Parsimony jackknifing outperforms neighbor-joining. *Cladistics* 12:99–124.

Farris, J. S., A. G. Kluge. 1998. A the brief history of three-taxon analysis. *Cladistics* 14:349–362.

Farris, J. S., A. G. Kluge, and M. J. Eckhart. 1970. A numerical approach to phylogenetic systematics. *Syst. Zool.* 19:172–189.

Felsenstein, J. 1978a. The number of evolutionary trees. *Syst. Zool.* 27:27–33.

Felsenstein, J. 1978b. Cases in which parsimony or compatibility methods will be positively misleading. *Syst. Zool.* 27:401–410.

Felsenstein, J. 1981a. A likelihood approach to character weighting and what it tells us about parsimony and compatibility. *Biol. J. Linnean Soc.* 16:183–196.

Felsenstein, J. 1981b. Evolutionary trees from DNA sequences: A maximum likelihood approach. *J. Mol. Evol.* 17:368–376.

Felsenstein, J. 1985a. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* 39:783–791.

Felsenstein, J. 1985b. Confidence limits on phylogenies with a molecular clock. *Syst. Zool.* 34:152–161.

Felsenstein, J. 2004. *Inferring Phylogenies*. Sinauer Assoc., Sunderland, MA.

Felsenstein, J. 2007. PHYLIP (Phylogeny Inference Package) version 3.67. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.

Feng, D. F., and R. F. Doolittle. 1987. Progressive sequence alignment as a prerequisite to correct phylogenetic trees. *J. Mol. Evol.* 25:351–360.

Feynman, R. 1965. *The Character of Physical Law*. MIT Press, Cambridge, MA.

Fichman, M. 1977. Wallace, zoogeography and the problem of land bridges. *J. Hist. Biol.* 10:45–63.

Fink, W. L., and M. L. Zelditch. 1995. Phylogenetic analysis of ontogenetic shape transformations: A reassessment of the piranha genus *Pygocentrus* (Teleostei). *Syst. Biol.* 44:343–360.

Fitch, W. M. 1970. Distinguishing homologous from analogous proteins. *Syst. Zool.* 19:99–113.

Fitch, W. M. 1971. Toward defining the course of evolution: Minimum change for a specific tree topology. *Syst. Zool.* 20:406–416.

Folinsbee, K. E., and D. R. Brooks. 2007. Miocene hominoid biogeography: Pulses of dispersal and differentiation. *J. Biogeogr.* 34:383–397.

Forey, P. L. 2001. The PhyloCode: Description and commentary. *Bull. Zool. Nomenclature* 58:81–96.

Forey, P. L. 2002. What's all this fuss about PhyloCode? *Palaeontology Newsletter* 44:20–32.

Forey, P. L., R. A. Fortey, P. Kenrick, and A. B. Smith. 2004. Taxonomy and fossils: A critical appraisal. *Phil. Trans. R. Soc. Lond. B* 359:639–653.

Forster, M. R., and E. Sober. 1994. How to tell when simpler, more unified, or less ad hoc theories will provide more accurate predictions. *British J. Phil. Sci.* 45:1–35.

Fortey, R. A. 1997. Classification. In: *Treatise on Invertebrate Paleontology, Part O, Arthropoda 1, Trilobita, revised. Volume 1: Introduction, Order Agnostida, Order Redlichiida* (R. L. Kaesler, ed.). Geological Society of America and University of Kansas. Boulder, CO, and Lawrence, KS.

Foster, D. R., P. K. Schoonmaker, and S. T. A. Pickett. 1990. Insights from paleoecology to community ecology. *Trends Ecol. Evol.* 5:119–122.

Frey, J. K. 1993. Modes of peripheral isolate formation and speciation. *Syst. Biol.* 42(3):373–381.

Frost, D. R., and A. G. Kluge. 1994. A consideration of epistemology in systematic biology, with special reference to species. *Cladistics* 10:259–294.

Frost, D. R., and D. M. Hillis. 1990. Species in concept and practice: Herpetological applications. *Herpetologica* 46:87–104.

Frost, D. R., A. G. Kluge, and D. M. Hillis. 1992. Species on contemporary herpetology: Comments on phylogenetic inference and taxonomy. *Herpetological Rev.* 23:46–54.

Funk, V. A. 1981. Special concerns in estimating plant phylogenies. In: *Advances in Cladistics: Proceedings of the First Meeting of the Willi Hennig Society* (V. A. Funk and D. R. Brooks, eds.). New York Botanical Garden, New York: 73–86.

Funk, V. A. 1985. Phylogenetic patterns and hybridization. *Ann. Missouri Bot. Gard.* 72:681–715.

Funk, V. A., and D. R. Brooks. 1990. Phylogenetic systematics as the basis of comparative biology. *Smithson. Contr. Bot.* 73:1–45.

Gaffney, E. S. 1979. Tetrapod monophyly: A phylogenetic analysis. *Bull. Carnegie Mus. Nat. Hist.* 13:92–105.

Gao, K., and M. A. Norell. 1998. Taxonomic revision of *Carusia* (Reptilia: Squamata) from the Late Cretaceous of the Gobi Desert and phylogenetic relationships of anguimorphan lizards. *Amer. Mus. Nat. Hist. Novitates* 3230:1–51.

Gao, and N. H. Shubin. 2001. Late Jurassic salamanders from northern China. *Nature* 410:574–577.

Garcia-Cruz, J., and V. Sosa. 2006. Coding quantitative character data for phylogenetic analysis: A comparison of five methods. *Syst. Bot.* 31:302–309.

Gatesy, J. 2000. Linked branch support and tree stability. *Syst. Biol.* 49:800–807.

Gaur, L. K., A. L. Hughes, E. R. Heise, and J. Gutknecht. 1992. Maintenance of DQB1 polymorphisms in primates. *Mol. Biol. Evol.* 9:599–609.

Gaut, B. S., and P. O. Lewis. 1995. Success of maximum likelihood phylogeny inference in the four-taxon case. *Mol. Biol. Evol.* 12:152–162.

Gauthier, J. 1986. Saurischian monophyly and the origin of birds. *Mem. California Acad. Sci.* 8:1–47.

Gauthier, J., A. G. Kluge, and T. Rowe. 1988. Amniote phylogeny and the importance of fossils. *Cladistics* 4:105–209.

Gegenbaur, C. 1873. Ueber das Archipterygium. *Jenaische Z. Med. Nat. Wiss.* 7:131–141.

Getz, W. M., and V. Kaitala. 1989. Ecogenetic models, competition, and heteropatry. *Theor. Pop. Biol.* 36:34–58.

Ghiselin, M. T. 1966. *The Triumph of the Darwinian Method*. University of California Press, Berkeley, CA.

Ghiselin, M. T. 1974. A radical solution to the species problem. *Syst. Zool.* 23:536–544.

Ghiselin, M. 1980. Natural kinds and literary accomplishments. *Mich. Q. Rev.* 19:73–88.

Ghiselin, M. T. 1984. "Definition," "Character," and Other Equivocal Terms. *Syst. Zool.* 33:104–110.

Ghiselin, M. T. 1989. Individuality, history and laws of nature in biology. In: *What the Philosophy of Biology Is* (M. Ruse, ed.). Kluwer Academic Publ., Dordrecht: 53–66.

Ghiselin, M. T. 1995. Ostensive definitions of the names of species and clades. *Biol. Phil.* 10:219–222.

Ghiselin, M. T. 1997. *Metaphysics and the Origin of Species*. State University of New York Press, Albany, NY.

Ghiselin, M. T. 2002. Species concepts: The basis for controversy and reconciliation. *Fish and Fisheries* 3:151–160.

Ghiselin, M. T. 2005. Homology as relationship of correspondence between parts of individuals. *Theory Biosci.* 124:91–103.

Ghiselin, M. T. 2007. Is the Pope a Catholic? *Biol. Phil.* 22:283–291.

Gift, N., and P. F. Stevens. 1997. Vagaries in the delimitation of character states in quantitative variation - An experimental study. *Syst. Biol.* 46:112–125.

Gilmour, J. S. L. 1940. Taxonomy and philosophy. In: *The New Systematics* (J. S. Huxley, ed.). Clarendon Press, Oxford: 461–474.

Gingerich, P. D. 1979. Paleontology, phylogeny, and classification: An example from the mammalian fossil record. *Syst. Zool.* 28:451–464.

Glasby, C. J., and B. Alvarez. 1999. Distribution patterns and biogeography analysis of austral Polychaeta (Annelida). *J. Biogeogr.* 26:507–533.

Goldman, N. 1988. Methods for discrete coding of morphological characters for numerical analysis. *Cladistics* 4:59–71.

Goloboff, P. A. 1993. Estimating character weights during tree search. *Cladistics* 9:83–91.

Goloboff, P. A. 1997. Self-weighted optimization: Tree searches and state reconstruction under implied transformation costs. *Cladistics* 13:225–245.

Goloboff, P. A. 1998. Tree searches under Sankoff parsimony. *Cladistics* 14:229–237.

Goloboff, P. A. 1999a. *NONA (NO NAME)*, ver. 2. Published by the author, Tucuman, Argentina.

Goloboff, P. A. 1999b. Analyzing large data sets in reasonable times: Solutions for composite optima. *Cladistics* 15:415–428.

Goloboff, P. A., J. S. Farris, and K. Nixon. 2000. *TNT* (Tree analysis using New Technology) BETA. Published by the authors, Tucuman, Argentina (with subsequent updates).

Goloboff, P. A., C. I. Mattoni, and A. S. Quinteros. 2006. Continuous characters analyzed as such. *Cladistics* 22:1–13.

Gomez, J. M. D., and F. Lobo. 2006. Historical biogeography of a clade of *Liolaemus* (Iguania: Liolaemidae) based on ancestral areas and dispersal-vicariance analysis (DIVA). *Paps. Avulsos de Zool. (Sao Paulo)* 46:261–274.

Good, D. A., and D. B. Wake. 1992. Geographic variation and speciation in the torrent salamanders of the genus *Rhyacotriton* (Caudata: Rhyacotritonidae). *Univ. Calif. Pub. Zool.* 126:1–91.

Goode, G. B., and T. H. Bean. 1895. Oceanic ichthyology. *Smithson. Contribut. Knowl.* 31, plate 66.

Goodman, M., J. Czelusniak, G. W. Moore, A. E. Romero-Herrera, and G. Matsuda. 1979. Fitting the gene lineage into its species lineage, a parsimony strategy illustrated by cladograms constructed from globin sequences. *Syst. Zool.* 28:132–163.

Goodman, M., G. W. Moore, and G. Matsuda. 1975. Darwinian evolution in the genealogy of haemoglobin. *Nature* 253:603–608.

Gould, S. J. 1977. *Ontogeny and Phylogeny*. Harvard University Press, Cambridge, MA.

Gould, S. J. 1999. A division of worms. *Nat. Hist.* 108:18–22.

Grady, J. M., and W. H. LeGrande. 1992. Phylogenetic relationships, modes of speciation, and historical biogeography of the madtom catfishes, genus *Noturus* Rafinesque (Siluriformes: Ictaluridae). In: *Systematics, Historical Ecology, and North American Freshwater Fishes* (R. L. Mayden, ed.). Stanford University Press, Stanford, CA: 747–777.

Graham, R. 1986. Response of mammalian communities to environmental changes during the late Quaternary. In: *Community Ecology* (J. Diamond and T. Case, eds.). Harper and Row, New York: 300–313.

Grande, L. 2010. An empirical synthetic pattern study of gars (Lepisosteiformes) and closely related species, based on skeletal anatomy. The resurrection of Holostei. *Amer. Soc. of Ichthy. Herp. Spec. Pub.* 6:1–871.

Grande, L., and W. E. Bemis. 1998. A comprehensive phylogenetic study of amid fishes (Amiidae) based on comparative skeletal anatomy, an empirical search for interconnected patterns of natural history. *Soc. Vert. Paleo. Mem.* 4:1–690.

Grant, V. 1981. *Plant Speciation*. Columbia University Press, New York.

Graybeal, A. 1998. Is it better to add taxa or characters to a difficult phylogenetic problem? *Syst. Biol.* 47:9–17.

Greenwood, P. H. 1984. African cichlid fishes and evolutionary theories. In: *Evolution of Fish Species Flocks* (A. A. Echelle and I. Kornfield, eds.). University of Maine at Orono Press, Orono: 121–154.

Greenwood, P. H., R. S. Miles, and C. Patterson (eds.). 1973. *Interrelationships of Fishes*. Suppl. 1, Zool. J. Linnean Soc. London, Vol. 53.

Greg, J. R. 1950. Taxonomy, language and reality. *Amer. Nat.* 84:419–435.

Griffiths, G. C. D. 1974. On the foundations of biological systematics. *Acta Biotheor.* 23:85–131.

Griffiths, P. E. 1999. Squaring the circle: Natural kinds with historical essences. In: *Species: New Interdisciplinary Studies* (R. Wilson, ed.) MIT Press, Cambridge, MA: 209–228.

Grinnell, G. 1974. The rise and fall of Darwin's first theory of transmutation. *J. Hist. Biol.* 7:259–273.

Gulick, J. T. 1888. Divergent evolution through cumulative segregation. *J. Linn. Soc., Zool.* 20:189–274.

Haeckel, E. 1866. *Generelle Morpholigie der Organismen, II.* Georg. Reiner, Berlin.

Halas, D., D. Zamparo, D. R. Brooks. 2004. A historical biogeographical protocol for studying biotic diversification by taxon pulses. *J. Biogeogr.* 31:1–12.

Hall, B. K. 1992. *Evolutionary Developmental Biology.* Chapman and Hall, London.

Hallam, A. 1977. Jurassic bivalve biogeography. *Paleobiology* 3:58–73.

Hallam, A. 1981a. Relative importance of plate movements, eustasy, and climate in controlling major biogeographical changes since the early Mesozoic. In: *Vicariance Biogeography: A Critique* (G. Nelson and D. E. Rosen, eds.). Columbia University Press, New York: 303–330.

Hallam, A., 1981b. Response. In: *Vicariance Biogeography: A Critique* (G. Nelson and D. E. Rosen, eds.). Columbia University Press, New York: 339–340.

Hallam, A. 1983. Early and mid-Jurassic molluscan biogeography and the establishment of the central Atlantic seaway. *Palaeogeog., Palaeoclimat., Palaeoecol.* 43:181–193.

Hallam, A. 1994. *An Outline of Phanerozoic Biogeography*, vol. 10. Oxford University Press, Oxford.

Hardison, R. C. 1996. A brief history of hemoglobins: Plant, animal, protist, and bacteria. *Proc. Natl. Acad. Sci., USA* 93:5675–5679.

Harlan, J. R., and J. M. J. De Wet. 1963. The compilospecies concept. *Evolution* 17:497–501.

Härlin, M. 1998. Taxonomic names and phylogenetic trees. *Zool. Scripta* 27:381–390.

Härlin, M. 1999. Phylogenetic approaches to nomenclature: A comparison based on a nemertean case study. *Proc. Royal Soc. London* 266:2201–2207.

Härlin, M. 2003a. Taxon names as paradigms: The structure of nomenclatural revolutions. *Cladistics* 19:138–143.

Härlin, M. 2003b. On the relationship between content, ancestor, and ancestry in phylogenetic nomenclature. *Cladistics* 19:144–147.

Härlin, M., and P. Sundberg. 1998. Taxonomy and philosophy of names. *Biol. Philos.* 13:233–244.

Harper, C. W., Jr. 1976. Phylogenetic inference in paleontology. *J. Paleontol.* 50:180–193.

Harvey, P. H., and M. D. Pagel. 1991. *The Comparative Method in Evolutionary Biology.* Oxford University Press, Oxford, UK.

Hasegawa, M., H. Kishino, and T. Yano. 1985. Dating of the human–ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.* 22:160–174.

Hastings, W. K. 1970. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* 57:97–109.

Haszprunar, G. 1992. The types of homology and their significance for evolutionary biology and phylogenetics. *J. Evol. Biol.* 5:13–24.

Haszprunar, G. 1998. Parsimony analysis as a specific kind of homology estimation and the implications for character weighting. *Mol. Phylogen. Evol.* 9:333–339.

Hecht, M. K., and J. L. Edwards. 1977. The methodology of phylogenetic inference above the species level. In: *Major Patterns in Vertebrate Evolution* (M. K. Hecht, P. C. Goody, and B. M. Hecht, eds.). Plenum Press, New York: 3–51.

Hedges, S. B., K. D. Moberg, and L. R. Maxon. 1990. Tetrapod phylogeny inferred from 18S and 28S ribosomal RNA sequences and amniote phylogeny. *Mol. Biol. Evol.* 7:607–633.

Hein, J. J. 1990. A unified approach to alignment and phylogenies. *Methods Enzymol.* 183:626–645.

Hein, J. 1994. TreeAlign. In: *Methods in Molecular Biology, Vol. 25 Computer Analysis of Sequence Data, Part III* (A. N. Griffin and H. G. Griffin, eds.). Humana Press, Totoaw, NJ: 349–364.

Hembree, D. I. 2006. Amphisbaenian palaeobiogeography: Evidence of vicariance and geodispersal patterns. *Palaeogeogr. Palaeoclimat. Palaeoecol.* 235:340–354.

Hendricks, J. R., and B. S. Lieberman. 2008. New phylogenetic insights into the Cambrian radiation of arachnomorph arthropods. *J. Paleontol.* 82:585–594.

Hendy, M. D., and D. Penny. 1984. Cladograms should be called trees. *Syst. Zool.* 33:245–247.

Hennig, W. 1950. *Grundzuge einer Theorie der phylogenetischen Systematik*. Deutscher Zentralverlag, Berlin.

Hennig, W. 1953. Kritische Bermerkungen zum phylogenetischen System der Insekten. *Beit. Entomol.* 3:1–85.

Hennig, W. 1965. Phylogenetic systematics. *Ann. Rev. of Entomol.* 10:97–116.

Hennig, W. 1966. *Phylogenetic Systematics*. University of Illinois Press, Urbana.

Hennig, W. 1969. *Die Stammesgeschichte der Insekten*. E. Kramer, Frankfurt am Main.

Hennig, W. 1975. Cladistic analysis or cladistic classification? A reply to Ernst Mayr. *Syst. Zool.* 24:244–256.

Hennig, W. 1981. *Insect Phylogeny*. John Wiley & Sons, New York.

Hennig, W. 1983. Stammesgeschichte der Chordaten. *Forsch. Zool. Systemat. Evol.* 2:1–208.

Hermsen, E. J., and J. R. Hendricks. 2008. W(h)ither fossils? Studying morphological character evolution in the age of molecular sequences. *Ann. Missouri Bot. Gard.* 95:72–100.

Hickerson, M. J., G. Dolman, and C. Moritz. 2006. Comparative phylogeographic summary statistics for testing simultaneous vicariance. *Mol. Ecol.* 15:209–223.

Higgins, D. G., A. J. Bleasby, and R. Fuchs. 1992. CLUSTAL V: Improved software for multiple sequence alignment. *Comput. Appl. Biosci.* 8:189–191.

Higgins, D. G., and P. M. Sharp. 1988. Clustal: A package for performing multiple sequence alignment on a microcomputer. *Gene* 73:237–244.

Highton R. 1998. Is *Ensatina eschscholtzii* a ring-species? *Herpetologica* 54:254–278.

Highton R. 2000. Detecting cryptic species using allozyme data. In: The *Biology of Plethodontid Salamanders* (R. C. Bruce, R. G. Jaeger, and L. D. Houck, eds.). Kluwer Academic/ Plenum, New York: 215–241.

Highton R., and R. B. Peabody. 2000. Geographical protein variation and speciation in salamanders of the *Plethodon jordani* and *Plethodon glutinosus* complexes in the southern Appalachian mountains with the description of four new species. In: *The Biology of Plethodontid Salamanders* (R. C. Bruce, R. G. Jaeger, and L. D. Houck, eds.). Kluwer Academic/ Plenum, New York: 31–94.

Hill, C. R., and P. R. Crane. 1982. Evolutionary cladistics and the origin of angiosperms. In: *Problems of Phylogenetic Reconstruction* (K. A. Joysey and A. E. Friday, eds.). Academic Press, New York: 269–361.

Hillis, D. M. 1987. Molecular versus morphological approaches to systematics. *Ann. Rev. Ecol. Syst.* 18:23–42.

Hillis, D. M. 1991. Discriminating between phylogenetic signal and random noise in DNA sequences. In: *Phylogenetic Analysis of DNA Sequences* (M. M. Miyamoto and J. Cracraft, eds.). Oxford University Press, New York: 278–294.

Hillis, D. M. 1994. Homology in molecular biology. In: *The Hierarchical Basis of Comparative Biology* (B. K. Hall, ed.). Academic Press, San Diego: 339–368.

Hillis, D. M. 1995. Approaches for assessing phylogenetic accuracy. *Syst. Biol.* 44:3–16.

Hillis, D. M. 2006. Constraints in naming parts of the tree of life. *Mol. Phylo. Evol.* 42:331–338.

Hillis, D. M., and J. J. Bull. 1993. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. *Syst. Biol.* 42:182–192.

Hillis, D. M., J. J. Bull, M. E. White, M. R. Badgett, and I. J. Molineux. 1992. Experimental phylogenetics: Generation of a known phylogeny. *Science* 255:589–592.

Hillis, D. M., and J. P. Huelsenbeck. 1994. Support for dental HIV transmission. *Nature* 369:24–25.

Hillis, D. M., J. P. Huelsenbeck, and C. W. Cunningham. 1994. Application and accuracy of molecular phylogenies. *Science* 264:671–677.

Hillis, D. M., B. K. Mable, A. Larson. S. K. David, and E. A. Zimmer. 1996. Nucleic Acids IV: Sequencing and cloning. In: *Molecular Systematics*, 2nd edition (D. M. Hillis, C. Moritz, and B. K. Mable, eds.). Sinauer Associates, Sunderland, MA: 321–381.

Hoberg, E. P., N. L. Alkire, A. de Queiroz, and A. Jones. 2001. Out of Africa: Origins of the *Taenia* tapeworms in humans. *Proc. Royal Soc. of London, Ser. B*, 268:781–787.

Hodges, E. R. S. (ed.). 2003. *The Guild Handbook of Scientific Illustration*. John Wiley & Sons, New York.

Holder, M., and P. O. Lewis. 2003. Phylogeny estimation: Traditional and Bayesian approaches. *Natur. Rev. Genetics* 4:275–284.

Holt, R. D., and M. S. Gaines. 1992. Analysis of adaptation in heterogeneous landscapes: Implications for the evolution of fundamental niches. *Evolutionary Ecol.* 6:433–447.

Hooker, J. D. 1853. *The Botany of the Antarctic Voyage of H. M. Discovery Ships "Erebus" and "Terror" in the years 1839–1843. II. Flora Novae-Zelandiae. Part I. Flowering Plants*. Lovell Reeve, London.

Hosbach, H. A., T. Wyler, and R. Weber. 1983. The *Xenopus laevis* globin gene family: Chromosomal arrangement and gene structure. *Cell* 32:45–53.

Hoskin, C. J., M. Higge, K. R. McDonald, and C. Moritz. 2005. Reinforcement drives rapid allopatric speciation. *Nature* 437:1353–1356.

Hovenkamp, P. 1997. Vicariance events, not areas, should be used in biogeographical analysis. *Cladistics* 13:67–79.

Howard, D. J., and S. H. Berlocher (eds.). 1998. *Endless Forms: Species and Speciation*. Oxford University Press, Oxford, UK.

Hudson, R. R. 1990. Gene genealogies and the coalescent process. *Oxford Sur. Evol. Biol.* 7:1–44.

Hudson, R. R. 1992. Gene trees, species trees, and segregation of ancestral alleles. *Genetics* 131:509–512.

Huelsenbeck, J. P. 1991a. Tree-length distribution skewness: An indicator of phylogenetic information. *Syst. Zool.* 40:257–270.

Huelsenbeck, J. P. 1991b. When are fossils better than extant taxa in phylogenetic analysis? *Syst. Zool.* 40:458–469.

Huelsenbeck, J. P. 1995. Performance of phylogenetic methods in simulation. *Syst. Biol.* 44:17–48.

Huelsenbeck, J. P., and K. A. Crandall. 1997. Phylogeny estimation and hypothesis testing using maximum likelihood. *Ann. Rev. Ecol. Syst.* 28:437–466.

Huelsenbeck, J. P., and D. M. Hillis. 1993. Success of phylogenetic methods in the four-taxon case. *Syst. Biol.* 42:247–264.

Huelsenbeck, J. P., and F. Ronquist. 2001. MRBAYES: Bayesian inference of phylogeny. *Bioinformatics* 17:754–755.

Huidobro, L., J. J. Morrone, J. L. Villalobos, and F. Álvarez. 2006. Distributional patterns of freshwater taxa (fishes, crustaceans and plants) from the Mexican transition zone. *J. Biogeogr.* 33:731–741.

Hull, D. L. 1964. Consistency and Monophyly. *Syst. Zool.* 13:1–11.

Hull, D. L. 1966. Phylogenetic Numericlature. *Syst. Zool.* 15:14–17.

Hull, D. L. 1967. Certainty and circularity in evolutionary taxonomy. *Evolution* 21:174–189.

Hull, D. L. 1968. The operational imperative: Sense and nonsense in operationism. *Syst. Zool.* 17:438–457.

Hull, D. L. 1976. Are species really individuals? *Syst. Zool.* 25:174–191.

Hull, D. L. 1978. A matter of individuality. *Phil. Sci.* 45:335–360.

Hull, D. L. 1979. The limits of cladism. *Syst. Zool.* 28:416–440.

Hull, D. L. 1980. Individuality and selection. *Ann. Rev. Ecol. Syst.* 11:311–332.

Hull, D. L. 1981. Historical narratives and integrating explanations. In: *Pragmatism and Purpose: Essays Presented to Thomas Goudge* (L.W. Sumner, J. G. Slater, and F. Wilson, eds.). University of Toronto Press, Toronto: 172–188.

Hull, D. L. 1983. Karl Popper and Plato's metaphor. In: *Advances in Cladistics: Proceedings of the Second Meeting of the Willi Hennig Society* (N. I. Platnick and V. A. Funk, eds.). Columbia University Press, New York: 177–189.

Humphries, C. J. 1979. A revision of the genus *Anacyclus* L. (Compositae: Anthermidae). *Bull. British Mus. Nature. Hist. (Bot.)* 7:83–142.

Humphries, C. J. 1980. Cytogenic and cladistic studies in Anacyclus (Compositae: Anthemideae). *Nordic J. Bot.* 1:93–96.

Humphries, C. J. 1983. Primary data in hybrid analysis. *Advances in Cladistics: Proceedings of the Second Meeting of the Willi Hennig Society* (N. I. Platnick and V. A. Funk, eds.). Columbia University Press, New York: 89–104.

Humphries, C. J., and L. Parenti. 1986. Cladistic Biogeography. *Oxford Mon. Biogeog.* 2:1–98.

Humphries, C. J., and L. R. Parenti. 1999. *Cladistic Biogeography*, 2nd edition. *Interpreting Patterns of Plant and Animal Distributions*. Oxford Biogeography Series No. 12. Oxford University Press, Oxford, UK.

Hunn, C. A., and P. Upchurch. 2001. The importance of time/space in diagnosing the causality of phylogenetic events: Towards a "chronobiogeographical paradigm." *Syst. Biol.* 50:391–407.

Huntley, B., T. Webb III. 1989. Migration: Species' response to climatic variations caused by changes in the earth's orbit. *J. Biogeogr.* 16:5–19.

Huson, D. H., and D. Bryant. 2005. Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.* 23:254–267.

Iguchi, K., K. Matsuura, K. M. McNysett, A. T. Peterson, R. Scachetti-Pererra, K. A. Powers, D. A. Veglais, E. O. Wiley, and T. Yodo. 2004. Predicting invasions of North American basses in Japan using native range data and a genetic algorithm. *Trans. Amer. Fisheries Soc.* 133:845–854.

International Code of Botanical Nomenclature (St. Louis Code). 2000. Regnum Vegetabile 138. Koeltz Scientific Books, Königstein. (W. Greuter, Chair, available online).

International Code of Nomenclature of Bacteria (1990 Revision). 1992. American Society of Mircobiology Press (P. A. H. Sneath, author; code is available online).

International Code of Zoological Nomenclature, 4th edition. 1999. The International Trust for Zoological Nomenclature, London.

Jackman T. R., and D. B. Wake. 1994. Evolutionary and historical analysis of protein variation in the blotched forms of salamanders of the *Ensatina* complex (Amphibia: Plethodontidae). *Evolution* 48:876–897.

Jardine, N. 1969. A logical basis for biological classification. *Syst. Zool.* 18:37–52.

Johnson, G. D. 1984. Percoidei: Development and relationships. In: *Ontogeny and Systematics of Fishes* (H. G. Moser, W. J. Richards, D. M. Cohen, M. P. Fahay, A. W. Kendall, and S. L. Richardson, eds.). Amer. Soc. Ichthyol. Herpetol.Spec. Publ. No. 1:464–498.

Jollie, M. 1973. *Chordate Morphology*. R. E. Krieger, Huntington, NY.

Jordan, D. S. 1905. The origin of species through isolation. *Science* 22:545–563.

Jordan, D. S., and R. W. Evermann. 1900. The fishes of North and Middle America, Part IV. Bull. U. S. Nat. Mus. No. 47, PL. CLXXVII.

Judd, W. S., C. S. Campbell, E. A. Kellogg, P. F. Stevens, and M. J. Donoghue. 2008. *Plant Systematics: A Phylogenetic Approach*, 2nd edition. Sinauer Associates, Sunderland, MA.

Jukes, T. H., and C. R. Cantor. 1969. Evolution of protein molecules. In: *Mammalian Protein Metabolism III* (H.-N. Munro, ed.). Academic Press, New York: 21–132.

Kallersjo, M., J. S. Farris, A. G. Kluge, and C. Bult. 1992. Skewness and permutation. *Cladistics* 8:275–287.

Katoh, K., and H. Toh. 2008. Recent developments in the MAFFT multiple sequence alignment program. *Briefings in Bioinformatics* 9:286–298.

Kavanaugh, D. H. 1972. Hennig's principles and methods of phylogenetic systematics. *The Biologist* 54:115–127.

Kearney, M., and J. M Clark. 2003. Problems due to missing data in phylogenetic analyses including fossils: A critical review. *J. Vert. Paleont.* 23:263–274.

Key, K. H. L. 1967. Operational homology. *Syst. Zool.* 16:275–276.

Kharchenko N. V., A. E. Piskunov, S. Roeser, E. Schilbach, and R.-D. Scholz. 2004. Astrophysical supplements to the ASCC-2.5. II. Membership probabilities in 520 Galactic open cluster sky areas. *Astronomische Nachrichten* 325:740–748.

Kimura, M. 1980. A simple method for estimating evolutionary rate of base substitution through comparative studies of nucleotide sequences. *J. Mol. Evol.* 16:111–120.

Kinch, M. P. 1980. Geographical distribution and the origin of life: The development of early nineteenth century British explanations. *J. Hist. Biol.* 13:91–119.

Kiriakoff, S. G. 1959. Phylogenetic systematics versus typology. *Syst. Zool.* 8:117–118.

Kirkpatrick, S. C. D. Gelatt, Jr., and M. P. Vecchi. 1983. Optimization by simulated annealing. *Science* 220:671–680.

Kitcher, P. 1984. Species. *Phil. Sci.* 51:308–333.

Kitching, I. J., P. L. Forey, C. J. Humphries, and D. M. Williams. 1998. *Cladistics. The Theory and Practice of Parsimony Analysis*, 2nd edition. Syst. Assoc. Publ. 11. Oxford University Press, New York.

Kjer, K. M. 1995. Use of ribosomal-RNA secondary structure in phylogenetic studies to identify homologous positions – an example of alignment and data presentation from the frogs. *Mol. Phylogen. Evol.* 4:314–330.

Klicka, J., and R. M. Zink. 1997. The importance of recent Ice Ages in speciation: A failed paradigm. *Science* 277:1666–1669.

Kluge, A. G. 1988. Parsimony in vicariance biogeography: A quantitative method and a Greater Antillean example. *Syst. Zool.* 37:315–328.

Kluge, A. G. 1990. Species as historical individuals. *Biol. Philo.* 5:417–431.

Kluge, A. G. 1993. Three-taxon transformation in phylogenetic inference: Ambiguity and distortion as regards to explanatory power. *Cladistics* 9:246–259.

Kluge, A. G., and J. S. Farris. 1969. Quantitative phyletics and the evolution of Anurans. *Syst. Zool.* 18:1–32.

Kluge, A., and J. Farris. 1999. Taxic homology = overall similarity. *Cladistics* 15:205–212.

Knowles, L. L., and W. P. Maddison. 2002. Statistical phylogeography. *Mol. Ecol.* 11:2623–2635.

Koponen, T. 1968. Generic revision of Mniaceae Mitt. *Ann. Bot. Fenn.* 5:117–151.

Kornet, D. J. 1993. Permanent splits as speciation events: A formal reconstruction of the internodal species concept. *J. Theor. Biol.* 164:407–435.

Kottler, M. J. 1978. Charles Darwin's biological species concept and theory of geographic speciation: The transmutation notebooks. *Ann. Sci.* 35:275–297.

Kreiser, B. R. 2001. Mitochondrial cytochrome b sequences support recognition of two cryptic species of plains killifish, *Fundulus zebrinus* and *Fundulus kansae*. *Am. Midland Natur.* 146:199–209.

Kripke, S. 1980. *Naming and Necessity*. Harvard University Press, Cambridge, MA.

Krishtalka, L. 1993. Anagenetic angst. Species boundries in Eocene primates. In: *Species, Species Concepts, and Primate Evolution* (W. H. Kimble and L. B. Mastin, eds.). Plenum Press, New York: 331–334.

Krishtalka, L., T. Peterson, D. Viegalis, J. Beach, and E. Wiley. 2002. The green internet: A tool for conservation science. In: *Conservation in the Internet Age: Strategic Threats and Opportunities* (J. Levitt, ed.). Island Press, Washington, DC: 143–164.

Lamarck, J. B. 1809. *Philosophie Zoologique*. Dentu, Paris.

Langeland, K. A. 1996. *Hydrilla verticillata* (L.F.) Royle (Hydrocharitaceae), The perfect aquatic weed. *Castanea* 61:293–304.

Lankester, E. R. 1870. On the use of the term homology in modern zoology and the distinction between homogenetic and homoplastic agreements. *Ann. Mag. Nat. Hist.* 6:34–43.

Lanyon, S. M. 1985. Detecting internal inconsistencies in distance data. *Syst. Zool.* 34:397–403.

Lapointe, F.-J., and G. Cucumel. 1997. The average consensus procedure: Combination of weighted trees containing identical or overlapping sets of taxa. *Syst. Biol.* 46:306–312.

Laubichler, M. D. 2000. Homology in development and the development of the homology concept. *Amer. Zool.* 40:777–788.

Laurin, M. 2005. The advantages of phylogenetic nomenclature over Linnean nomenclature. In: *Animal Names* (A. Minelle, G. Ortalli, and G. Sanga, eds.). Instituto Veneto di Scienze, Lettere ed Arti, Venezia: 67–97.

Leaché, A., S. A. Crews, and M. J. Hickerson. 2007. Two waves of diversification in mammals and reptiles of Baja California revealed by hierarchical Bayesian analysis. *Biol. Lets.* 3:646–650.

Lee, D. S., C. R. Gilbert, C. H. Hocutt, R. E. Jenkins, D. E. McAllister, J. R. Stauffer (eds.). l980 et seq. *Atlas of North American Freshwater Fishes*. North Carolina State Museum of Natural History, Raleigh.

Lehman, H. 1967. Are biological species real? *Phil. Sci.* 34:157–167.

Lewis, P. O. 2001a. Phylogenetic systematics turns over a new leaf. *Trends in Ecol. and Evol.* 16:30–37.

Lewis, P. O. 2001b. A likelihood approach to estimating phylogeny from discrete morphological character data. *Syst. Biol.* 50:913–925.

Lieberman, B. S. 1992. An extension of the SMRS concept into a phylogenetic context. *Evol. Theory* 10:157–161.

Lieberman, B. S. 1993. Systematics and biogeography of the "*Metacryphaeus* Group," (Trilobita, Devonian) with a comment on adaptive radiations and the geological history of the Malvinokaffric Realm. *J. Paleont.* 67:549–570.

Lieberman, B. S. 1994. Evolution of the trilobite subfamily Proetinae and the origin, evolutionary affinity, and extinction of the Middle Devonian proetid fauna of Eastern North America. *Bull. Am. Mus. Nat. Hist.* 223:1–176.

Lieberman, B. S. 1997. Early Cambrian paleogeography and tectonic history: A biogeographic approach. *Geology* 25:1039–1042.

Lieberman, B. S. 1998. Cladistic analysis of the Early Cambrian olenelloid trilobites. *J. Paleontol.* 72:59–78.

Lieberman, B. S. 1999. Systematic revision of the Olenelloidea (Trilobita, Cambrian). *Bull. Yale University Peabody Mus. Nat. His.* 45:1–150.

Lieberman, B. S. 2000a. *Paleobiogeography: Using Fossils to Study Global Change, Plate Tectonics, and Evolution*. Plenum Press/Kluwer Academic, New York.

Lieberman, B. S. 2000b. Applying molecular phylogeography to test paleoecological hypotheses: A case study involving *Amblema plicata* (Mollusca, Unionidae). In: *Evolutionary Paleoecology* (W. D. Allmon and D. Bottjer, eds.). Columbia University Press, New York: 83–103.

Lieberman, B. S. 2001. Phylogenetic analysis of the Olenellina (Trilobita, Cambrian). *J. Paleont.* 75:96–115.

Lieberman, B. S. 2002a. Phylogenetic analysis of some basal Early Cambrian trilobites, the biogeographic origins of the eutrilobita, and the timing of the Cambrian radiation. *J. Paleont.* 76:672–688.

Lieberman, B. S. 2002b. Biogeography with and without the fossil record. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 178:39–52.

Lieberman, B. S. 2003a. Unifying theory and methodology in biogeography. *Evol. Biol.* 33:1–25.

Lieberman, B. S. 2003b. Biogeography of the Cambrian radiation: Deducing geological processes from trilobite evoluton. *Special Paps. in Palaeontol.* 70:59–72.

Lieberman, B. S. 2003c. Paleobiogeography: The relevance of fossils to biogeography. *Ann. Rev. Ecol. Syst.* 34:51–69.

Lieberman, B. S. 2004. Revised biostratigraphy, systematics, and paleobiogeography of the trilobites from the Middle Cambrian Nelson Limestone, Antarctica. *Univ. Kansas Paleontol. Contribs*. 14:1–23.

Lieberman, B. S. 2005. Geobiology and paleobiogeography: Tracking the coevolution of the Earth and its biota. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 219:23–33.

Lieberman, B. S., and N. Eldredge. 1996. Trilobite biogeography in the Middle Devonian: Geological processes and analytical methods. *Paleobiology* 22:66–79.

Lieberman, B. S., and G. Kloc. 1997. Evolutionary and biogeographic patterns in the Asteropyginae (Trilobita, Devonian). *Bull. Am. Mus. Nat. Hist.* 232:1–127.

Lieberman, B. S., and A. L. Melott. 2007. Considering the case for biodiversity cycles: Reexamining the evidence for periodicity in the fossil record. *PLoS One* 2(8) e759:1–9.

Lieberman, B. S., G. D. Edgecombe, and N. Eldredge. 1991. Systematics and biogeography of the "*Malvinella* Group," Calmoniidae (Trilobita, Devonian). *J. Paleont.* 65:824–843.

Lieberman, B. S., C. E. Brett, and N. Eldredge. 1995. Patterns and processes of stasis in two species lineages from the Middle Devonian of New York State. *Paleobiology* 21:15–27.

Linnaeus, C. 1751. *Philosophia Botanica*. G. Kiesewetter, Stockholm.

Linnaeus, C. 1753. *Species Plantarum*. L. Salvius, Stockholm.

Liu, F.-G. R., M. M. Miyamoto, N. P. Freire, P. Q. Ong, M. R. Tennant, T. S. Young, and K. F. Gugel. 2001. Molecular and morphological supertrees for eutherian (placental) mammals. *Science* 291:1786–1789.

Livezey, B. C. 1989. Phylogenetic relationships and incipient flightlessness of the extinct Auckland Islands Merganser. *Wilson Bull.* 101:410–435.

Lomolino, M. H., B. R. Riddle, R. J. Whittaker, and J. H. Brown. 2010. *Biogeography*, 4th edition. Sinauer Associates, Sunderland, MA.

Losos, J. B., and R. E. Glor. 2003. Phylogenetic comparative methods and the geography of speciation. *Trends in Ecol. and Evol* 18:220–227.

Lovtrup, S. 1977. *Phylogeny of Vertebrata*. John Wiley & Sons, New York.

Lozier, J., P. Aniello, and M. J. Hickerson. 2009. Predicting the distribution of Sasquatch in western North America: Anything goes with ecological niche modelling *J. Biogeo.* 36:1623–1627.

Lundberg, J. G., and B. Chernoff. 1992. A Miocene fossil of the Amazonian fish *Arapaima* (Teleostei, Arapaimidae) from the Magdalena River region of Colombia-biogeographic and evolutionary implications. *Biotropica* 24:2–14.

Lydekker, R. 1896. *A Geographical History of Mammals*. Cambridge University Press, Cambridge, UK.

Lyell, C. 1832. *Principles of Geology, Vol. 2*, 2nd edition. University of Chicago Press, Chicago.

Lynch, J. D. 1989. The gauge of speciation: On the frequencies of modes of speciation. In: *Speciation and its Consequences* (D. Otte and J.A. Endler, eds.). Sinauer Assoc., Sunderland, MA: 527–553.

MacArthur, R. H., and E. O. Wilson. 1967. *The Theory of Island Biogeography*. Princeton University Press, Princeton, NJ.

MacLeod, N. 2001. Landmarks, localizability and the use of morphometrics in phylogenetic analysis. In: *Fossils, Phylogeny and Form* (J. Adrain, G. Edgecombe, and B. Lieberman, eds.). Kluwer Academic/Plenum Press, New York: 197–233.

MacLeod, N. 2002. Phylogenetic signals in morphometric data. In: *Morphometrics, Shape, and Phylogenetics* (N. MacLeod and P. Forey, eds.). Taylor and Francis, London: 100–138.

MacLeod, N., and Pl. L. Forey (eds.). 2002. *Morphology, Shape and Phylogeny*. Systematics Assoc. Spec. Vol. Ser. 64:1–308.

Maddison, D. R. 1991. Discovery and importance of multiple islands of most-parsimonious trees. *Syst. Zool.* 40:315–32.

Maddison, D. R., and W. P. Maddison. 1992–2008. *McClade (now version 4)*. Sinauer Assoc. Sunderland, MA.

Maddison, W. P., and D. R. Maddison. 2009. Mesquite: A modular system for evolutionary analysis. Version 2.71. (www.mesquiteproject.org).

Maddison, W. P. 1991. Squared-change parsimony reconstructions of ancestral states for continuous-valued characters on a phylogenetic tree. *Syst. Zool.* 40:304–314.

Maddison W. P. 1993. Missing data versus missing characters in phylogenetic analysis *Syst. Biol.* 42:576–581.

Maddison, W. P. 1989. Reconstructing character evolution on polytomous cladograms. *Cladistics* 5:365–377.

Maddison, W. P. 1997. Gene trees in species trees. *Syst. Biol.* 46:523–536.

Maddison, W. P., and D. R. Maddison. 1992. *Mac-Clade: Interactive Analysis of Phylogeny and Character Evolution*. Sinauer Associates, Sunderland, MA.

Maddison, W. P., M. J. Donoghue, and D. R. Maddison. 1984. Outgroup analysis and parsimony. *Syst. Zool.* 33:83–103.

Maguire, K. C., and A. L. Stigall. 2008. Paleobiogeography of Miocene Equinae of North America: A phylogenetic biogeographic analysis of the relative roles of climate, vicariance, and dispersal. *Palaeogeo. Palaeoclimat. Palaeoecol.* 267:175–184.

Mahner, M., and M. A. Bunge. 1997. *Foundations of Biophilosophy*. Springer-Verlag, New York.

Mallet, J. 1995. The species definition for the modern synthesis. *Trends Ecol. Evol.* 10:294–299.

Margush, T., and F. R. McMorris. 1981. Consensus n-trees. *Bull. Math. Biol.* 43:239–244.

Marschall, A. 1873. *Nomenclator Zoologicus*. Vindobonae, typis C Ueberreuter (M. Salzer).

Martin, J., D. Blackburn, and E. O. Wiley. 2010. Are node-based and stem-based clades equivalent? Insights from graph theory. *PLoS Curr.* 2010 November 18; 2: RRN1196. doi:10.1371/currents.RRN1196.

Matthew, W. D. 1915. Climate and evolution. *Ann. New York Acad. Sci.* 24:171–318.

Matthew, W. D. 1939. *Climate and Evolution*, 2nd edition. New York Academy of Sciences, New York.

May, R. M. 1988. How many species are there on Earth? *Science* 241(4872):1441–1449.

Mayden, R. L. 1985. Biogeography of Ouachita highland fishes. *Southwest. Natur.* 30:195–211.

Mayden, R. L. 1987. Historical ecology and North American highland fishes: A research program in community ecology. In: *Community and Evolutionary Ecology of North American Stream Fishes* (W. J. Matthews and D. C. Heins, eds.). University of Oklahoma Press, Norman: 210–222.

Mayden, R. L. 1988a. Systematics of the *Notropis zonatus* species group, with description of a new species from the interior highlands of North America. *Copeia* 1988:153–173.

Mayden, R. L. 1988b. Vicariance biogeography, parsimony, and evolution in North American freshwater fishes. *Syst. Zool.* 37:329–355.

Mayden, R. L. 1992. An emerging revolution in comparative biology and the evolution of North American freshwater fishes. In: *Systematics, Historical Ecology and North American Freshwater Fishes* (R. L. Mayden, ed.). Stanford University Press, Stanford, CA: 864–890.

Mayden, R. L. 1997. A hierarchy of species concepts: The denouement in the saga of the species problem. In: *Species, the Units of Biodiversity* (M. F. Claridge, H. A. Dawah, and M. R. Wilson, eds.). Chapman and Hall, London: 381–424.

Mayden, R. L. 1999. Consilience and a hierarchy of species concepts: Advances towards closure on the species puzzle. *J. Nematol.* 31:95–116.

Mayden, R. L. 2002. On biological species, species concepts and individualtion in the natura. world. *Fish Fisheries* 3:171–196.

Mayden, R. L., and R. M. Wood. 1995. Systematics, species concepts, and the evolutionary significant unit in biodiversity and conservation biology. *Am. Fisheries Soc. Sympo.* 17:58–117.

Mayr, E. 1942. *Systematics and the Origin of Species*. Columbia University Press, New York.

Mayr, E. 1957. Difficulties and importance of the biological species concept. In: *The Species Problem* (E. Mayr, ed.). AAAS, Washington, DC: 371–388.

Mayr, E. 1959. Darwin and the Evolutionary Theory in Biology. In: *Evolution and Anthropology: A Centennial Appraisal* (B. J. Meggers, ed.). Anthropological Society of Washington: 1–10.

Mayr, E. 1963. *Animal Species and Evolution*. Harvard University Press, Cambridge, MA.

Mayr E. 1968. Theory of biological classification. *Nature* 220:545–548.

Mayr, E. 1969. *Principles of Systematic Zoology*. McGraw-Hill, New York.

Mayr, E. 1970. *Populations, Species, and Evolution: An Abridgment of Animal Species and Evolution*. Belknap Press, Cambridge, MA.? Harvard University Press

Mayr, E. 1974. Cladistic analysis or cladistic classification. *Z. Zool. Syst. Evolut.-forsch.* 12:94–128.

Mayr E. 1976. *Evolution and the Diversity of Life: Selected Essays*. Belknap Press, Cambridge, MA.

Mayr, E. 1982. *The Growth of Biological Knowledge*. Harvard University Press, Cambridge, MA.

Mayr, E. 2000. The biological species concept. In: *Species Concepts and Phylogenetic Systematics* (Q. D. Wheeler and R. Meier, eds.). Columbia University Press, New York: 17–29.

Mayr, E., and P. D. Ashlock. 1991. *Principles of Systematic Zoology*, 2nd edition. McGraw-Hill, New York.

McDade, L. A. 1990. Hybrids and phylogenetic systematics I. Patterns of character expression in hybrids and their implications for cladistic analysis. *Evolution* 44:1685–1700.

McDade, L. A. 1992. Hybrids and phylogenetic systematics II. The impact of hybrids on cladistic analysis. *Evolution* 46:1329–1346.

McDade, L. A. 1997. Hybrids and phylogenetic systematics III. Comparison with distance methods. *Syst. Bot.* 22:669–683.

McGhee, G. R., Jr. 1996. *The Late Devonian Mass Extinction*. Columbia University Press, New York.

McGuire, J. A., C. C. Witt, D. L. Altshuler, and J. V. Remsen. 2007. Phylogenetic systematics and biogeography of hummingbirds: Bayesian and Maximum Likelihood Analyses of partitioned data and selection of an appropriate partitioning strategy. *Syst. Biol.* 56:837–856.

McKenna, M. C. 1983. Holarctic landmass rearrangement, cosmic events, and Cenozoic terrestrial organisms. *Ann. Missouri Bot. Gard.* 70:459–89.

McKitrick, M. C. 1994. On homology and the ontological relationship of parts. *Syst. Biol.* 43:1–10.

McKenna, M. C. 1975. Toward a phylogenetic classification of the Mam malia. In: *Phylogeny of the Primates* (W. P. Luckett and F. S. Szalay, eds.). Plenum Press, New York: 21–46.

McLennan, D. A., D. R. Brooks, and J. D. McPhail. 1988. The benefits of communication between comparative ethology and phylogenetic systematics: A case study of gasterostid fishes. *Can. J. Zool.* 66:2177–2190.

McNamara, K. D. 1978. Paedomorphosis in Scottish olenellid trilobites (Early Cambrian). *Palaeontol.* 21:635–655.

McNyset, K. M. 2009. Ecological niche conservatism in North American freshwater fishes. *Biol. J. Linnean Soc.* 96:282–295.

Meert, J. G., and B. S. Lieberman. 2004. A palaeomagnetic and palaeobiogeographic perspective on latest Neoproterozoic and early Cambrian tectonic events. *J. Geol. Soc. London* 161:1–11.

Meier, R., and R. Willmann. 2000. The Hennigian species concept. In: *Species Concepts and Phylogenetic Theory: A Debate* (Q. D. Wheeler and R. Meier, eds.). Columbia University Press, New York: 30–42.

Merxmüller, H., P. Leins, and H. Roessler. 1977. Inuleae—systematic review. In: *The Biology and Chemistry of the Compositae* (V. H. Heywood, J. B. Harbourne, and B. L. Turner, eds.). Academic Press, San Diego: 577–601.

Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. 1953. Equations of state calculations by fast computing machines. *J. Chem. Phys.* 21:1087–1091.

Metzker, M. L., D. P. Mindell, X.-M. Liu, R. G. Ptak, R. A. Gibbs, and D. M. Hillis. 2002. Molecular evidence of HIV-1 transmission in a criminal case. *PNAS* 99:14292–14297.

Mickevich, M. F. 1978. Taxonomic congruence. *Syst. Zool.* 27:143–158.

Mickevich, M. F. 1981. Quantitative phylogenetic biogeography. In: *Advances in Cladistics: Proceedings of the First Meeting of the Willi Hennig Society* (V. A. Funk and D. R. Brooks, eds.). New York Botanical Garden, Bronx, NY: 209–222.

Mickevich, M. F. 1982. Transformation series analysis. *Syst. Zool.* 31:461–478.

Mickevich, M. F., and D. Lipscomb. 1991. Parsimony and the choice between different transformations for the same character set. *Cladistics* 7:111–139.

Mickevich, M. F., and S. J. Weller. 1990. Evolutionary character analysis: Tracing character change on a cladogram. *Cladistics* 6:137–170.

Mickevich, M. F., and M. F. Johnson. 1976. Congruence between morphological and allozyme data in evolutionary inference and character evolution. *Syst. Zool.* 25:260–270.

Mill, J. S. 1872. *A System of Logic, Ratiocinactive and Inductive: Being a Connected View of the Principles of Evidence and Methods of Scientific Investigation*. Longmans, Green and Dyer, London.

Milne, M. J., and L. J. Milne. 1939. Evolutionary trends in caddis fly case construction. *Ann. Entomol. Soc. Amer.* 32:533–542.

Minckley, W. L. 1973. *Fishes of Arizona*. Arizona Game and Fish Dept., Phoenix, AZ.

Mishler, B. D. 1985. The morphological, developmental, and phylogenetic basis of species concepts in bryophytes. *Bryologist* 88:207–214.

Mishler, B. D., and R. N. Brandon. 1987. Individuality, pluralism and the phylogenetic species concept. *Biol. and Philo.* 2:397–414.

Mishler, B. D., and M. J. Donoghue. 1982. Species concepts; a case for pluralism. *Syst. Zool.* 31:491–503.

Mishler, B. D., and E. C. Theriot. 2000. Monophyly, apomorphy, and phylogenetic species concepts. In: *Species Concepts and Phylogenetic Systematic: A Debate* (Q. D. Wheeler and R. Meier, eds.). Columbia University Press, New York: 44–54.

Miyamoto, M. M. 1985. Consensus cladograms and general classifications. *Cladistics* 1:186–189.

Miyamoto, M. M., and J. Cracraft. 1991. *Phylogenetic Anlaysis of DNA Data*. Oxford University Press, New York.

Miyamoto, M. M., and W. M. Fitch. 1995. Testing species phylogenies and phylogenetic methods with congruence. *Syst. Biol.* 44:64–76.

Mooi, R. D., and Gill, A. C. 2010. Phylogenetics without synapomorphies—A crisis in fish systematics: Time to show some character. *Zootaxa* 2450:26–40.

Morrison, D. 2005. Networks in phylogenetic analysis: New tools for population biology. *Int. J. Parasitol.* 35:567–582.

Morrone, J. J. 1994. On the identification of areas of endemism. *Syst. Biol.* 43:438–441.

Morrone, J. J. 1998. On Udvardy's Insulantarctica province: A test from the weevils (Coleoptera: Curculinoidea). *J. Biogeogr.* 25:947–955.

Morrone, J. J. 2005. Cladistic biogeography: Identity and place. *J. Biogeogr.* 32:1281–1284.

Morrone, J. J. 2008. *Evolutionary Biogeography*. Columbia University Press, New York.

Morrone, J. J., and J. M. Carpenter. 1994. In search of a method for cladistic biogeography: An empirical comparison of component analysis, Brooks Parsimony Analysis, and three-area statements. *Cladistics* 10:99–153.

Morrone, J. J., and J. V. Crisci. 1995. Historical biogeography: Introduction to methods. *Ann. Rev. Ecol. Syst.* 26:373–401.

Morrone, J. J., and T. Escalante. 2002. Parsimony analysis of endemicity (PAE) of Mexican terrestrial mammals at different units: When size matters. *J. Biogeogr.* 29: 1095–1104.

Morrone, J. J., and J. Márquez. 2001. Halffter's Mexican transition zone, beetle generalised tracks, and geographical homology. *J. Biogeogr.* 28:635–650.

Mueller, L. D., and F. J. Ayala. 1982. Estimation and interpretation of genetic distance in empirical studies. *Genetical Res.* 40:127–137.

Myers, A. A. 1991. How did Hawaii accumulate its biota? A test from the Amphipoda. *Global Ecol. Biogeogr. Lttrs.* 1:24–29.

Naef, A. 1919. *Idealistische Morphologie und Phylogenetik (zur Methodik der systematischen)*. Verlag von Gustav Fischer, Jena.

Nakhleh, L., and L. S. Wang. 2005. Phylogenetic networks: Properties and relationship to trees and clusters. *LNCS Transactions on Computational Systems Biology, II, LNBI* 3680:82–99.

Nathan, R., G. Perry, J. T. Cronin, A. E. Strand, and M. I. Cain. 2003. Methods for estimating long-distance dispersal. *Oikos* 103:261–273.

Naumann, C. M. 1977. Studies on the systematics and phylogeny of Holarctic Sesiidae (Insecta, Lepidoptera). *Amerind Publ., New Delhi*. (Trans. From Bonner Zoolog. Monogr. 1, 1971).

Navarro-Sigüenza, A. G., and A. T. Peterson. 2004. An alternative species taxonomy of Mexican birds. *Biota Neotrop.* 4(2). Online (biotaneotropica.org)

Naylor, G. J. P. 1996. Can partial warp scores be used as cladistic characters? In: *Advances in Morphometric*s (L. F. Marcus, M. Corti, A. Loy, G. Naylor, and D. Slice, eds.). Plenum Press, New York: 519–530.

Neave, S. A. (ed.). 1939–1940, 1950. *Nomenclator Zoologicus*. Zool. Soc. London, London.

Needleman, S. B., and C. D. Wunsch. 1970. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Mol. Biol.* 48:443–453.

Nei, M. 1986. Stochastic errors in DNA evolution and molecular phylogeny. In: *Evolutionary Perspectives and the New Genetics* (H. Gershowwitz, L. Rucknagel, and R. E. Tashian, eds.). Alan R. Liss, New York: 133–147.

Nelson, G. J. 1969. Gill arches and the phylogeny of fishes, with notes on the classification of vertebrates. *Bull. Amer. Mus. Nat. Hist.* 141:475–552.

Nelson, G. J. 1970. An outline of the theory of comparative biology. *Syst. Zool.* 19:373–384.

Nelson, G. J. 1971a. "Cladism" as a philosophy of classification. *Syst. Zool.* 20:373–376.

Nelson. G. J. 1971b. Paraphyly and polyphyly: Redefinitions. *Syst. Zool.* 20:471–472.

Nelson, G. J. 1972a. Phylogenetic relationship and classification. *Syst. Zool.* 21:227–231.

Nelson, G. J. 1972b. Science or politics? A reply to H. F. Howden. *Syst. Zool.* 21:341–342.

Nelson, G. J. 1972c. Review of: Die rekonstruktion der phylogenese mit Hennig's Prinzip. *Syst. Zool.* 21:350–352.

Nelson, G. J. 1972d. Comments on Hennig's "phylogenetic systematics" and its influence on Ichthyology. *Syst. Zool.* 21:364–374.

Nelson, G. J. 1973. Classification as an expression of phylogenetic relationships. *Syst. Zool.* 22:344–359.

Nelson, G. J. 1974a. Classification as an expression of phylogenetic relationships. *Syst. Zool.* 22:334–359.

Nelson, G. J. 1974b. Darwin-Hennig classification: A reply to Ernst Mayr. *Syst. Zool.* 23:452–458.

Nelson, G. 1976. Biogeography, the vicariance paradigm, and continental drift. *Syst. Zool.* 24:490–504.

Nelson, G. 1978. Fron Candolle to Croizat: Comments on the history of biogeography. *J. Hist. Biol.* 11:269–305.

Nelson, G. J. 1979. Cladistic analysis and synthesis: Principles and definitions, with a historical note on Adanson's Families des Plantes (1763–1764). *Syst. Zool.* 28:1–21.

Nelson, G. J. 1983. Reticulation in cladograms. In: *Advances in Cladistics: Proceedings of the Second Meeting of the Willi Hennig Society* (N. I. Platnick and V. A. Funk, eds.). Columbia University Press, New York: 105–111.

Nelson, G. J. 1989. Species and taxa: Systematics and evolution. In: *Speciation and Its Consequences* (D. Otte and J.A. Endler, eds.). Sinauer Associates, Sunderland, MA: 60–81.

Nelson, G. J. 1994. Homology and systematics. In: *Homology: The Hierarchical Basis of Comparative Biology* (B. K. Hall, ed.). Academic Press, San Diego: 101–149.

Nelson, G. J., and N. I. Platnick. 1980. Multiple branching in cladograms: Two interpretations. *Syst. Zool.* 29:86–91.

Nelson, G. J., and N. I. Platnick. 1981. *Systematics and Biogeography: Cladistics and Vicariance.* Columbia University Press, New York.

Neumann, D. A. Faithful consensus methods for *n*-trees. *Math. Biosci.* 63:271–287.

Newton, M. A. 1996. Bootstrapping phylogenies: Large deviations and dispersion effects. *Biometrika* 83:315–328.

Nihei, S. S. 2006. Misconceptions about parsimony analysis of endemicity. *J. Biogeogr.* 33:2099–2106.

Nix, H A. 1986. A biogeographic analysis of Australian elapid snakes. In: *Snakes: Atlas of Elapid Snakes of Australia* (R. Longmore, ed.). Australian Gov. Publ. Ser., Canberra: 4–15.

Nixon, K. C. 1999. The parsimony ratchet, a new method for rapid parsimony analysis. *Cladistics* 15:407–414.

Nixon, K. C., and J. M. Carpenter. 2000. On the other "phylogenetic systematics." *Cladistics* 16:298–318.

Nixon, K. C., and J. Davis. 1991. Polymorphic taxa, missing values and cladistic analysis. *Cladistics* 7:233–241.

Nixon, K. C., and Q. D. Wheeler. 1992. Extinction and the origin of species. In: *Extinction and Phylogeny (*M. J. Novacek and Q. D. Wheeler, eds.). Columbia University Press, New York: 119–143.

Nixon, K. C., and Q. D. Wheeler. 1990. An amplification of the phylogenetic species concept. *Cladistics* 6:211–223.

NOAA. 1999. World Ocean Atlas 1998. National Oceanographic Data Center, Silver Sping, MD. (3 CD-ROM set).

Noonan, G. R. 1988. Biogeography of North American and Mexican insects, and a critique of vicariance biogeography. *Syst. Zool.* 37:366–84.

Novacek, M. J. 1992a. Fossils as critical data for phylogeny. In: *Extinction and Phylogeny* (M. J. Novacek and Q. D. Wheeler, eds.). Columbia University Press, New York: 46–88.

Novacek, M. J. 1992b. Fossils, topologies, missing data, and the higher level phylogeny of eutherian mammals *Syst. Biol.* 41:58–73.

Ogden, T. H., and M. S. Rosenberg. 2007a. How should gaps be treated in parsimony? A comparison of approaches using simulation. *Mol. Phylogenet. Evol.* 46:807–808.

Ogden, T. H., and M. S. Rosenberg. 2007b. Alignment and topological accuracy of the direct optimization approach via POY and traditional phylogenetics via ClustalW + PAUP. *Syst. Biol.* 56:183–193.

O'Hara, R. J. 1993. Systematic generalization, historic fate, and the species problem. *Syst. Biol.* 42:231–246.

O'Grady, R. T., and G. B. Deets. 1987. Coding multistate characters, with special reference to the use of parasites as characters of their hosts. *Syst. Zool.* 36:268–279.

O'Grady, R. T., G. B. Deets, and G. W. Benz. 1989. Additional observations on nonredundant linear coding of multistate characters. *Syst. Zool.* 38:54–57.

O'Leary, M. A. 1999. Parsimony analysis of total evidence from extinct and extant taxa and the cetacean-artiodactyl question (Mammalia: Ungulata). *Cladistics* 15:315–330.

Oosterbroek, P. 1987. More appropriate definitions of paraphyly and polyphyly, with a comment on the Farris 1974 model. *Syst. Zool.* 36:103–108.

Osche, G. 1973. Das Homologisieren als eine grundlegende Methode der Phylogenetik. *Aufs. u. Reden Senckenb. naturf. Ges.* 24:155–164.

Osche, G. 1982. Rekapitulationscntwicklung und ihre Bedeutung für die Phylogenetik – Wann gilt die "Biogenetische Grundregel"? *Vern. Naturw. Ver. Hamburg (N.F.)* 25:5–31.

Otte, D., and J. A. Endler (eds.). 1989. *Speciation and its Consequences*. Sinauer Associates, Sunderland, MA.

Owen, R. 1843. *Lectures on the Comparative Anatomy and Physiology of the Invertebrate Animals*. London: Longman Brown Green and Longmans.

Owen, R. 1849. *On the Nature of Limbs*. J. Van Voor, London.

Page, R. D. M. 1990. Component analysis: A valiant failure? *Cladistics* 6:119–136.

Pagel, M. 1999. The maximum likelihood approach to reconstructing ancestral character states of discrete characters on phylogenies. *Syst. Biol.* 48:612–622.

Palmer, A. R., and L. N. Repina. 1993. Through a glass darkly: Taxonomy, phylogeny, and biostratigraphy of the Olenellina. *Univ. Kansas Paleontol. Contribs*. 3:1–35.

Pamilo, I., and M. Nei. 1988. Relationships between gene trees and species trees. *Mol. Biol. Evol.* 5:568–583.

Panchen, A. L. 1994. Richard Owen and the Concept of Homology. In: *Homology: The Hierarchical Basis of Comparative Biology* (B. K. Hall, ed.). Academic Press, San Diego: 22–62.

Paterson, H. E. H. 1978. More evidence against speciation by reinforcement. *South African J. Sci.* 74:369–371.

Paterson, H. E. H. 1984. The recognition concept of species. *South African J. Sci.* 80:312–318.

Patterson, C. 1977. The contribution of paleontology to teleostean phylogeny. In: *Major Patterns in Vertebrate Evolution* (M. K. Hecht, P. C. Goody, and B. M. Hecht, eds.). Plenum Press, New York: 579–643.

Patterson, C. 1978. Verifiability in systematics. *Syst. Zool.* 27:218–222.

Patterson, C. 1981. Significance of fossils in determining evolutionary relationships. *Ann. Rev. Ecol. Syst.* 12:195–223.

Patterson, C. 1982. Morphological characters and homology. In: *Problems of Phylogenetic Reconstruction* (K. A. Joysey and A. E. Friday, eds.). Academic Press, London: 21–74.

Patterson, C. 1983. Aims and methods in biogeography. In: *Evolution, Time and Space: The Emergence of the Biosphere* (R. W. Sims, J. H. Price, and P. E. S. Whalley, eds.). Academic Press, New York: 1–28.

Patterson, C. 1988. Homology in classical and molecular biology. *Mol. Biol. Evol.* 5:603–625.

Patterson, C., and D. E. Rosen. 1977. Review of ichthyodectiform and other Mesozoic fishes and the theory and practice of classifying fossils. *Bull. Am. Mus. Nat. Hist.* 158:81–172.

Penny, D., M. A. Steel, P. J. Lockhart, and M. D. Hendy. 1994. The role of models in reconstructing evolutionary trees. In: *Models in Phylogeny Reconstruction* (R. W. Scotland, D. J. Siebert, and D. M. Williams, eds.). Oxford University Press, Oxford: 211–230.

Peterson, A. T. 2003. Predicting the geography of species' invasions via ecological niche modeling. *Quart. Rev. Biol.* 78:419–433.

Peterson, A. T., and A. G. Navarro-Siguenza. 1999. Alternate species concepts as bases for determining priority conservation areas. *Conservation Biol.* 13:427–431.

Peterson, A. T., M. A. Ortega-Huerta, J. Bartley, V. Sánchez-Cordero, J. Soberón, R. H. Buddemeier, and D. R. B. Stockwel. 2001a. Future projections for Mexican faunas under global climate change scenarios. *Nature* 416:626–629.

Peterson, A. T., V. Sanchez-Cordero, J. Soberon, J. Bartley, R. H. Buddemeier, and A. G. Navarro-Siguenza. 2001b. Effects of global climate change on geographic distributions of Mexican Cracidae. *Ecol. Model.* 144:21–30.

Peterson, A. T., J. Soberon, and V. Sanchez-Cordero. 1999. Conservatism of ecological niches in evolutionary time. *Science* 285:1265–1267.

Peterson, A. T., D. A. Vieglais, A. G. Navarro-Sigüenza, and M. Silva. 2003. A global distributed biodiversity information network: Building the world museum. *Bull. British Ornithologists' Club* 123A:186–196.

Phillips, A., D. Janies, and W. Wheeler. 2000. Multiple sequence alignment in phylogenetic analysis. *Mol. Phylogen. Evol.* 16:317–330.

Pimentel, R. A., and R. Riggins. 1987. The nature of cladistic data. *Cladistics* 3:201–209.

Platnick, N. I. 1976. Concepts of dispersal in historical biogeography. *Syst. Zool.* 25:294–295.

Platnick, N. I. 1977. Cladograms, phylogenetic trees, and hypothesis testing. *Syst. Zool.* 26:438–442.

Platnick, N. I. 1979. Philosophy and the transformation of cladistics. *Syst. Zool.* 28:537–546.

Platnick, N. I. 1985. Philosophy and the transformation of cladistics revisited. *Cladistics* 1:87–94.

Platnick, N. I., and G. Nelson. 1978. A method of analysis for historical biogeography. *Syst. Zool.* 27:1–16.

Pleijel, F. 1995. Oncharacter coding for phylogeny reconstruction. *Cladistics* 11:309–315.

Popper, K. R. 1965. *The Logic of Scientific Discovery*. Harper Torchbooks, New York.

Porter, A. H. 1990. Testing nominal species boundaries using gene flow statistics: Taxonomy of two hybridizing admiral butterflies (*Limenitis*: Nymphalidae). *Syst. Zool.* 39:131–147.

Poss, S. G., and R. R. Miller. 1983. Taxonomic status of the plains killifish, *Fundulus zebrinus*. *Copeia* 1983:55–67.

Prager, E. M., and A. C. Wilson. 1988. Ancient origin of lactalbumin from lysozyme: Analysis of DNA and amino acid sequences. *J. Mol. Evol.* 27:326–335.

Prosada, D., and T. R. Buckley. 2004. Model selection and model averaging in phylogenetics: Advantages of Akaike Information Criterion and Bayesian approaches over likelihood ratio tests. *Syst. Biol.* 53:793–808.

Prothero, D. R., and D. B. Lazarus. 1980. Planktonic microfossils and the recognition of ancestors. *Syst. Zool.* 29:119–129.

Puorto G., M. Da Graça Salomão, R. D. G. Theakston, R. S. Thorpe, D. A. Warrell, and W. Wüster. 2001. Combining mitochondrial DNA sequences and morphological data to infer species boundaries: Phylogeography of lancehead pitvipers in the Brazilian Atlantic forest, and the status of *Bothrops pradoi* (Squamata: Serpentes: Viperidae). *J. Evol. Biol.* 14:527–538.

Quijano-Abril, M. A., R. Callejas-Posada, and D. R. Miranda-Esquivel. 2006. Areas of endemism and distribution patterns for Neotropical *Piper* species (Piperaceaea). *J. Biogeogr.* 33:1266–1278.

Quine, W. V. 1969. *Ontological Relativity and Other Essays*. Columbia University Press, New York.

Radford, A. E. 1986. *Fundamentals of Plant Systematics*. Harper & Row, New York.

Rannala, B., and Z. Yang. 1996. Probability distribution of molecular evolutionary trees: A new method of phylogenetic inference. *J. Mol. Evol.* 43:304–311.

Read, R. J. 2000. Macromolecular Crystallography Course, University of Cambridge: Likelihood: Theory and application to structure refinement. www.structmed.cimr.cam.ac.uk/Course/Likelihood/likelihood.html.

Ree, R. H. 2005. Detecting the historical signature of key innovations using stochastic models of character evolution and cladogenesis. *Evolution* 59:257–265.

Ree, R. H., and M. J. Donoghue. 1998. Step matrices and the interpretation of homoplasy. *Syst. Biol.* 47:582–588.

Ree, R. H., B. R. Moore, C. O. Webb, and M. J. Donoghue. 2005. A likelihood framework for inferring the evolution of geographic range on phylogenetic trees. *Evolution* 59:2299–2311.

Ree, R. H., and S. A. Smith. 2008. Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Syst. Biol.* 57:4–14.

Rendall, D., and A. Di Fiore. 2007. Homoplasy, homology, and the perceived special status of behavior in evolution. *J. Human Evol.* 5:504–521.

Remane, A. 1952. *Die Grundlagen des Naturlichen Systems der Vergleichenden Anatomie und der Phylogenetik*. Geest und Portig K.G.: Leipzig, Germany.

Remane, A. 1956. *Die Grundlagen des naturlichen Systems der vergleichenden Anatomie und. Phylogenetik*, 2nd edition. Geest & Portik, Leipzig.

Rice, K., M. Donoghue, and R. Olmstead. 1997. Analyzing large data sets: rbcL 500 revisited. *Syst. Biol.* 46:554–563.

Richardson, R. A. 1981. Biogeography and the genesis of Darwin's ideas on transmutation. *J. Hist. Biol.* 14:1–41.

Richter, S., and R. Meier. 1994. The development of phylogenetic concepts in Hennig's early theoretical publications (1947–1966). *Syst. Biol.* 43:212–221.

Ridley, M. 1986. *Evolution and Classification: The Reformation of Cladism*. Longman Press, London.

Riedl, R. 1978. *Order in Living Organisms*. John Wiley & Sons, New York.

Rieppel, O. 1980. Homology, a deductive concept? *Zeitschrift fur Zoologische Systematik und Evolutionsforschung* 18:315–319.

Rieseberg, L. H. 1991. Homoploid reticulate evolution in *Helianthus* (Asteraceae): Evidence from ribosomal genes. *Am. J. Bot.* 78:1218–1237.

Rieseberg, L. H., C. R. Linder, and G. J. Seiler. 1995. Chromosomal and genetic barriers to introgression in Helianthus. *Genetics* 141:1163–1171.

Rieseberg, L. H., and J. H. Willis. 2007. Plant speciation. *Science* 317:910–914.

Rindge, F. H. 1972. A revision of the moth genus *Plataea* (Lepidoptera, Geometridae). *Amer. Mus. Novit.* 2595:1–42.

Rode, A., and B. S. Lieberman. 2004. Using GIS to study the biogeography of the Late Devonian biodiversity crisis. *Palaeogeo. Palaeoclimat. Palaeoecol.* 211:345–359.

Rode, A. L., and B. S. Lieberman. 2005. Integrating biogeography and evolution using phylogenetics and PaleoGIS: A case study involving Devonian crustaceans. *J. Paleontol.* 79:267–276.

Rohlf, F. J. 1982. Consensus indices for comparing classifications. *Math. Biosci.* 59:131–144.

Rohlf, F. J. 1998. On applications of geometric morphometrics to studies of ontogeny and phylogeny. *Syst. Biol.* 47:147–158.

Rohlf, F. J., and L. F. Marcus. 1993. A revolution in morphometrics. *Trends Ecol. Evol.* 8:129–132.

Romer, A. S. 1966. *Vertebrate Paleontology*. University of Chicago Press, Chicago.

Ronquist, F. 1994. Ancestral areas and parsimony. *Syst. Biol.* 43:267–274.

Ronquist, F. 1995. Ancestral areas revisited. *Syst. Biol.* 44:572–575.

Ronquist, F. 1996. DIVA, version 1.0. Computer program for MacOS and Win 32. www.ebc.uu.se/systzoo/research/diva/diva.html.

Ronquist, F. 1997. Dispersal-vicariance analysis: A new approach to the quantification of historical biogeography. *Syst. Biol.* 46:195–203.

Ronquist, F. 1998a. Fast Fitch-parsimony algorithms for large data sets. *Cladistics* 14:387–400.

Ronquist, F. 1998b. Phylogenetic approaches in coevolution and biogeography. *Zool. Scr.* 26:313–322.

Ronquist, F. 2002. TreeFitter, version 1.3. www.ebc.uu.se/systzoo/research/treefitter.

Ronquist, F., and J. P. Huelsenbeck. 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572–1574.

Rosen, B. R. 1988. From fossils to earth history: Applied biogeography. In: *Analytical Biogeography* (A. A. Myers and P. S. Giller, eds.). Chapman and Hall, London: 437–478.

Rosen, D. E. 1978. Vicariant patterns and historical explanations in biogeography. *Syst. Zool.* 27:159–188.

Rosen, D. E. 1979. Fishes from the upland and intermontane basins of Guatemala: Revisionary studies and comparative geography. *Bull. Amer. Mus. Nat. Hist.* 162:267–376.

Rosenzweig, M. L. 1995. *Species Diversity in Space and Time*. Cambridge University Press, New York.

Ross, A., and J. R. P. Ross. 1985. Carboniferous and early Permian biogeography. *Geology* 13:27–30.

Ross, H. H. 1972. The origin of species diversity in ecological communities. *Taxon* 21:253–259.

Ross, H. H. 1986. Resource partitioning in fish assemblages: A review of field studies. *Copeia* 1986:352–388.

Roth, V. L. 1994. Within and between organisms: Replicators, lineages, and homologs. In: *Homology: The Hierarchical Basis of Comparative Anatomy* (B. K. Hall, ed.). Academic Press, San Diego: 301–337.

Rowe, T. 1988. Definition, diagnosis, and origin of Mammalia. *J. Vert. Paleontol.* 8:241–264.

Rull, V. 2004. Biogeography of the "Lost World": A palaeoecological perspective. *Earth-Sci. Rev.* 67:125–137.

Ruse, M. 1969. Definitions of species in biology. *British J. Phil. Sci.* 20:97–119.

Ruse, M. 1971. The species problem: A reply to Hull. *British J. Phil. Sci.* 22:369–371.

Ruse, M. 1987. Biological species: Natural kinds, individuals, or what? *British J. Phil. Sci.* 38:225–242.

Russell, B. 1919. *Introduction to Mathematical Philosophy*. George Allen and Unwin, London.

Ruvolo, M. 1997. Molecular phylogeny of the hominoids: Inference from multiple independent DNA data sets. *Mol. Biol. Evol.* 14(3):248–265.

Saitou, N., and M. Nei. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Evol. Biol.* 4:406–425.

Salthe, N. 1985. *Evolving Hierarchical Systems*. Columbia University Press, New York.

Sanderson, M. J., A. Purvis, and C. Henze. 1998. Phylogenetic supertrees: Assembling the tree of life. *Trends Ecol. Evol.* 13:105–109.

Sankoff, D. 1975. Minimal mutation trees of sequences. *SIAM J. Appl. Math.* 28:35–42.

Sankoff, D. 2000. The early introduction of dynamic programming into computational biology. *Bioinformatics* 16:41–47.

Sankoff, D., Y. Abel, and J. Hein. 1994. A tree, a window, a hill: Generalization of nearest neighbor interchange in phylogenetic optimization. *J. Classification* 11:209–232.

Sankoff, D., C. Morel, and R. J. Cedergren. 1973. Evolution of 5S RNA and the non-randomness of base replacement. *Nature New Biol.* 245:232–234.

Sankoff, D., and R. J. Cedergren. 1983. Simultaneous comparison of three or more sequences related by a tree. In: *Time Warps, String Edits, and Macromolecules: The Theory and Prac-*

tice of Sequence Comparison (D. Sankoff and J. B. Kruskal, eds.). Addison-Wesley, Reading, MA: 253–263.

Sankoff, D., and P. Rousseau. 1975. Locating the vertices of a Steiner tree in an arbitrary space. *Math. Program.* 9:240–246.

Sanmartin, I., and F. Ronquist. 2004. Southern hemisphere biogeography inferred by event-based models: Plant versus animal patterns. *Syst. Biol.* 53:216–243.

Sanmartin, I., H. Enghoff, and F. Ronquist. 2001. Patterns of animal dispersal, vicariance and diversification in the Holarctic. *Biol. J. Linnean Soc.* 73:345–90

Sattler, R. 1984. Homology—a continuing challenge. *Syst. Bot.* 9:382–394.

Sattler, R. 1994. Homology, homeosis and process morphology in plants. In: *Homology: The Hierarchical Basis of Comparative Biology* (B. K. Hall, ed.). Academic Press, San Diego: 423–475.

Schaeffer, B. 1967. Osteichthyan vertebrae. *J. Linn. Soc. London* 47:185–195.

Schlee, D. 1968. Hennig's principle of phylogenetic systematics, an "intuitive statistico-phentic approach"? *Syst. Zool.* 18:127–134.

Schlee, D. 1971. *Die Rekonstrucktion der Phylogenese mit Hennig's Prinzip*. Waldermar Kramer, Frankfurt am Main.

Schliewen, U. K., and B. Klee. 2004. Reticulate sympatric speciation in Cameroonian crater lake cyclids. *Frontiers Zool.* 2004:1–5.

Schuh, R. T. 1976. Pretarsal structure in the Miridae (Hemiptera) with a cladistic analysis of the relationships of the family. *Amer. Mus. Novit.* 2601:1–39.

Schultz, T. R., R. B. Cocroft, and G. A. Churchill. 1996. The reconstruction of ancestral character states. *Evolution* 50:504–511.

Schulze, F. E., K. Kukenthal, and K. Heider (eds.). 1926–1954. Nomencaltor Animalium – Genera et Subgenera. *Preussishe Akad*. Wissenscaften, Berlin.

Scotland, R. W. 1992. Character Coding. In: *Cladistics: A Practical Course in Systematics*, *Vol. 10*, *Systematics Association Publication*. (P. L. Forey, C. Humphries, J. I. J. Kitching, R. W. Scotland, D. J. Siebert, and D. M. Williams, eds.). Claredon Press, Oxford: 14–21.

Scotland, R. W. 2000. Taxic homology and three-taxon statement analysis. *Syst. Biol.* 49:480–500.

Scott, B. 1997. Biogeography of the Helicoidea (Mollusca: Gastropoda: Pulmonata): Land snails with a pangean distribution. *J. Biogeog.* 24:399–407.

Scudder, S. J. 1882. Nomencaltor Zoologicus. *U.S. Nat. Mus. Bull.* 10.

Sereno, P. C. 1997. The origin and evolution of dinosaurs. *Ann. Rev. Earth Plan. Sci.* 25:435–489.

Sereno, P. C. 1999. The evolution of dinosaurs. *Science* 284:2137–2147.

Sereno, P. C., D. B. Dutheil, M. Iarochene, H. C. E. Larsson, G. H. Lyon, P. M. Magwene, C. A. Sidor, D. J. Varricchio, and J. A. Wilson. 1996. Predatory dinosaurs from the Sahara and late Cretaceous faunal differentiation. *Science* 272:986–991.

Sherborn, C. D. 1902, 1903–1933. *Index Animalium*. C. J. Clay & Sons, London.

Shubin, N., C. Tabin, and S. Carroll. 2007. Fossils, genes and the evolution of animal limbs. *Nature* 388:639–648.

Simberloff, D., K. L. Heck, E. D. McCoy, and E. F. Connor. 1981. There have been no statistical tests of cladistic biogeographical hypotheses. In: *Vicariance biogeography: A Critique* (G. Nelson and D. E. Rosen, eds.). Columbia University Press, New York: 40–63.

Simpson, G. G. 1944. *Tempo and Mode in Evolution*. Columbia University Press, New York.

Simpson, G. G. 1945. The principles of classification and a classification of mammals. *Bull. Amer. Mus. Nature. Hist.* 85:1–350.

Simpson, G. G. 1951. *Horses: The Story of the Horse Family in the Modern World and through 60 Million Years of History*. Oxford University Press, Oxford, UK.

Simpson, G. G. 1961. *The Principles of Animal Taxonomy*. Columbia University Press, New York.

Simpson, G. G. 1975. Recent advances in methods of phylogenetic inference. In: *Phylogeny of the Primates, a Multidisciplinary Approach* (W. P. Luckett and F. S. Szalay, eds.). Plenum Press, New York: 3–19.

Simpson, M. G. 2006. *Plant Systematics*. Elsevier Academic Press, Amsterdam.

Sites, J. W., and K. A Crandall. 1997. Testing species boundaries in biodiversity studies. *Conserv. Biol.* 11(6):1289–1297.

Sites, J. W., and J. C. Marshall. 2004. Operational criteria for delimiting species. *Ann. Rev. Ecol. Syst.* 35:199–227.

Slowinski, J. B. 1993. "Unordered" versus "ordered" characters. *Syst. Biol.* 42:155–165.

Smit, H. A., T. J. Robinson, and B. J. Van Vuuren. 2007. Coalescence methods reveal the impact of vicariance on the spatial genetic structure of *Elephantulus edwardii* (Afrotheria, Macroscelidea). *Mol. Ecol.* 16:2680–2692.

Smith, G. R. 1992. Phylogeny and biogeography of the Catostomidae, freshwater fishes of North America and Asia. In: *Systematics, Historical Ecology and North American Freshwater Fishes* (R. L. Mayden, ed.). Stanford University Press, Stanford, CA: 778–826.

Smith, J. M. 1966. Sympatric speciation. *Am. Natur.* 110:637–650.

Smith, A. B., G. L. J. Patterson, and B. Lafay. 1995. Ophiuroid phylogeny and higher taxonomy: Morphological, molecular, and paleontological perspectives. *Z. J. Linnean Soc.* 114:213–243.

Smith, W. H. F., and D. T. Sandwell. 1997. Global sea floor topography from satellite alimetry and ship depth soundings. *Science* 277:1956–1962.

Sneath, P. H. A. 1976. Phenetic taxonomy at the species level and above. *Taxon* 25:437–450.

Sneath P. H. A., and R. R. Sokal. 1973. *Numerical Taxonomy*. Freeman, San Francisco.

Sober, E. 1975. *Simplicity*. Oxford University Press, New York.

Sober, E. 1983a. Parsimony in systematics: Philosophical issues. *Ann. Rev. Ecol. Syst.* 14:335–357.

Sober, E. 1983b. A likelihood justification of parsimony. *Cladistics* 1:209–233.

Sober, E. 1987. Parsimony, likelihood, and the principle of the common cause. *Philo. Sci.* 54:465–469.

Sober, E. 1988. *Reconstructing the Past: Parsimony, Evolution, and Inference*. MIT Press, Cambridge, MA.

Sober, E. 1993. *Philosophy of Biology*. Westview Press, Boulder, CO.

Sober, E. 2000. *Philosophy of Biology*, 2nd edition. Westview Press, Boulder, CO.

Sober, E. 2008. *Evidence and Evolution. The Logic Behind the Science*. Cambridge University Press, New York.

Soberón, J. 1999. Linking biodiversity information sources. *Trends Ecol. Evol.* 14:291.

Sokal, R. R., and T. Crovello. The biological species concept: A critical evaluation. *Am. Nat.* 104:127–153.

Sokal R. R., and C. D. Michener. 1958. A statistical method for evaluating systematic relationships. *Univ. Kansas Sci. Bull.* 38:1409–1438.

Sokal, R. R., and F. J. Roth. 1981. *Biometry*. W. H. Freeman, San Francisco.

Sokal, R. R., and P. H. A. Sneath. 1963. *Principles of Numerical Taxonomy*. W. H. Freeman, San Francisco.

Soltis, D., P. Soltis, M. Mort, M. Chase, V. Savolainen, S. Hoot, and C. Morton. 1998. Inferring complex phylogenies using parsimony: An empirical approach using three large DNA data sets for angiosperms. *Syst. Biol.* 47:32–42.

Sorenson, M. D., and E. A. Franzosa. 2007. *TreeRot*, version 3. Boston University, Boston, MA.

Sosa, V., and E. De Luna. 1998. Morphometric and character state recognition for cladistic analyses in the Bletia reflexa complex (Orchidaceae). *Plant Systematics and Evolution* 212:185–213.

Spemann, H. 1915. Zur Geschichte und Kritik des Begriffs der Homologie. In: *Allgemeine Biologie* (C. ChunHUN and W. Johannsen, eds.). Teubner, Berlin: 63–85. (Cited in Laubichgler, 2000; included herein for completeness.)

Stanley, S. M. 1979. *Macroevolution, Pattern and Process*. W. H. Freeman, San Francisco.

Stanley, S. M. 1990. Delayed recovery and the spacing of major extinctions. *Paleobiology* 16:401–414.

Stearn, W. T. 2004. *Botanical Latin*. Timber Press, Portland, OR.

Steel, M. A., L. Szekely, P. L. Erdös, and P. J. Waddell. 1993. A complete family of phylogenetic invariants for any number of taxa under Kimura's 3ST model. *N. Z. J. Bot.* 31:289–296.

Stevens, G. 1992. Spilling over the competitive limits to species coexistence. In: *Systematics, Ecology, and the Biodiversity Crisis* (N. Eldredge, ed.). Columbia University Press, New York: 40–58.

Stevens, N. J., and C. P. Heesy. 2006. Malagasy primate origins: Phylogenies, fossils, and biogeographic reconstructions. *Folia Primatologica* 77:419–433.

Stevens, P. F. 1984. Homology and phylogeny; morphology and systematics. *Syst. Bot.* 9:395–409.

Stevens, P. F. 1991. Character states, morphological variation, and phylogenetic analysis: A review. *Syst. Bot.* 16:553–583.

Stiassny, M. L. J., L. R. Parenti, and G. D. Johnson (eds.). 1996. *Interrelationships of Fishes*. Academic Press, San Diego.

Stiassny, M. L. J., E. O. Wiley, G. D. Johnson, and M. R. deCaravalho. 2004. Gnathostome fishes. In: *Assembling The Tree of Life* (J. Cracraft and M. J. Donoghue, eds.). Oxford University Press, Oxford, UK: 410–429.

Stigall Rode, A. L. 2005a. The application of geographic information systems to paleobiogeography: Implications for the study of invasions and mass extinctions. In: *Paleobiogeography: Generating New Insights into the Coevolution of the Earth and Its Biota* (B. S. Lieberman and A. L. Stigall Rode, eds.). *Paleontol. Soc. Pap.* 11:77–88.

Stigall Rode, A. L. 2005b. Systematic revision of the Devonian brachiopods *Schizophoria (Schizophoria)* and "*Schuchertella*" from North America. *J. Syst. Palaeontol.* 3:133–167.

Stigall, A. L. 2008. Tracking species in space and time: Assessing the relationships between paleobiogeography, paleoecology, and macroevolution. In: *From Evolution to Geobiology: Research Questions Driving Paleontology at the Atart of a New Century* (P. H. Kelly and R. K. Bambach, eds.). *Paleontol. Soc. Pap*. 14:227–242.

Stigall, A. L., and B. S. Lieberman. 2006a. Quantitative Paleobiogeography: GIS, Phylogenetic biogeographic analysis, and conservation insights. *J. Biogeogr.* 33:2051–2060.

Stigall Rode, A. L., and B. S. Lieberman. 2006b. Using environmental niche modelling to study the Late Devonian biodiversity crisis. In: *Understanding Late Devonian and Permian–Triassic Biotic and Climatic Events: Towards an Integrated Approach* (D. J. Over, J. R. Morrow, and P. B. Wignall, eds.). Developments in Palaeontology and Stratigraphy. Amsterdam: Elsevier: 93–180.

Stockwell, D. R. B. 1999. Genetic algorithms II. In: *Machine Learning Methods for Ecological Applications* (A. H. Fielding, ed.). Kluwer Academic, Boston: 123–144.

Stockwell, D. R. B., and I. R. Noble. 1992. Induction of sets of rules from animal distribution data: A robust and informative method of analysis. *Mathematics and Computers in Simulation* 33:385–390.

Stockwell, D. R. B., and D. P. Peters. 1999. The GARP modeling system: Problems and solutions to automated spatial prediction. *Interna. J. Geographic Inform. Sci.* 13:143–158.

Strait, D., M. Moniz, and P. Strait. 1996. Finite mixture coding: A new approach to coding continuous characters. *Syst. Biol*. 45:67–78.

Strickland, H. E. 1842. Rules for zoological nomenclature. Rept. 12th meeting, British Assoc. *Adv. Sci. Rpt.* 1842:105–121.

Striedter, G. F., and R. G. Northcutt. 1989. Two distinct visual pathways through the superficial pretectum in a percomorph teleost. *J. Comp. Neurol.* 283:342–354.

Studier J. A., and K. L. Keppler. 1988. A note on the neighbor–joining algorithm of Saitou and Nei. *Mol. Biol. Evol.* 5:729–731.

Style Manual Committee, Council of Science Editors. 2006. *Scientific Style and Format: The CSE Manual for Authors, Editors, and Publishers*. Council of Science Editors in cooperation with the Rockefeller University Press, Reston, VA.

Stussey, T. F. 2009. *Plant Taxonomy*. Columbia University Press, New York.

Sullivan, J., and P. Joyce. 2005. Model selection in phylogenetics. *Ann. Rev. Evol. Syst.* 36:445–466.

Sullivan, J., and D. L. Swofford. 2001. Should we use model-based methods for phylogenetic inference when we know that assumptions about among-sites variation and nucleotide substitution patterns are violated. *Syst. Biol.* 50:723–729.

Sullivan, J., J. A. Markert, and C. W. Kilpatrick. 1997. Phylogeography and molecular systematics of the *Peromyscus aztecus* group (Rodentia: Muridae) inferred using parsimony and likelihood. *Syst. Biol.* 46:426–440.

Sulloway, F. J. 1979. Geographic isolation in Darwin's thinking: The vicissitudes of a crucial idea. In: *Studies in the History of Biology* (W. Coleman and C. Limoges, eds.). Johns Hopkins University Press, Baltimore, MD: 23–65.

Sun, G., Q. Ji, D. L. Dilcher, S. Zheng, K. C. Nixon, and X. Wang. 2002. Archaefructaceae, a new basal angiosperm family. *Science* 296:899–904.

Swiderski, D. L., M. L. Zelditch, and W. L. Fink. 2002. Comparability, morphometrics and phylogenetic systematics. In: *Morphometrics, Shape and Phylogenetics* (N. MacLeod and P. Forey, eds.). Taylor and Francis, London: 67–99.

Swofford, D. L. 1991. When are phylogeny estimates from molecular and morphological data incongruent? In: *Phylogenetic Analysis of DNA Sequences* (M. M. Miyamoto and J. Cracraft, eds.). Oxford University Press, New York: 295–333.

Swofford, D. L. 2001. PAUP*: Phylogenetic Analysis Using Parsimony* (* and Other Methods), Version 4.0b8. Sinauer, Sunderland, MA. (PAUP first appears in 1990.)

Swofford, D. L., and S. H. Berlocher. 1987. Inferring evolutionary trees from gene frequency data under the principle of maximum parsimony. *Syst. Zool.* 36:293–325.

Swofford, D. L., and W. P. Maddison 1987 Reconstructing ancestral character states under Wagner parsimony. *Math. Biosci.* 87:199–229.

Swofford, D., and D. Maddison.1992. Parsimony, character-state reconstructions, and evolutionary inferences. In: *Systematics, Historical Ecology, and North American Freshwater Fishes* (R. Mayden. ed.). Stanford University Press, Stanford, CA: 186–223.

Swofford, D. L., and G. J. Olsen. 1990. Phylogeny reconstruction. In: *Molecular Systematics* (D. M. Hillis and C. Moritz, eds.). Sinauer Associates, Sunderland, MA: 411–501.

Swofford, D. L., G. J. Olsen, P. J. Waddell, and D. M. Hillis. 1996. Phylogenetic inference. In: *Molecular Systematics*, 2nd edition (D. M. Hillis, C. Moritz, and B. Mable, eds.). Sinauer Associates, Sunderland, MA: 407–514.

Taberlet, P., and J. Bouvet. 1994. Mitochondrial DNA polymorphism, phylogeography, and conservation genetics of the brown bear (*Ursos arctos*) in Europe. *Proc. R. Soc. London, Ser. B* 255:195–200.

Takhtajan, A. 1969. *Flowering Plants: Origin and Dispersal* (translated by C. Jeffrey). Oliver and Boyd, Edinburgh.

Tanaka, M., Münsterberg, A., Anderson, W. G., Prescott, A. R., Hazon, N., and C. Tickle. 2002. Fin development in a cartilaginous fish and the origin of vertebrate limbs. *Nature* 416:527–531.

Tattersall, I., and N. Eldredge. 1977. Fact, theory and fantasy in human paleontology. *Amer. Sci.* 65:204–211.

Tavaré, S. 1986. Some probabilistic and statistical problems in the analysis of DNA sequences. *Lectures on Mathematics in the Life Sciences* 17:57–86.

Templeton, A. R. 1989. The meaning of species and speciation: A genetic perspective. In: *Speciation and Its Consequences* (D. Otte and J. A. Endler, eds.). Sinauer Associates, Sunderland, MA: 3–27.

Templeton, A. R. 1983a. Phylogenetic inference from restriction endonuclease cleavage site maps with particular reference to the evolution of humans and the apes. *Evolution* 37:221–244.

Templeton, A. R. 1983b. Convergent evolution and non-parametric inferences from restriction fragment and DNA sequence data. In: *Statistical Analysis of DNA Sequence Data* (B. Weir, ed.). Marcel Dekker, New York: 151–179.

Templeton, A. R. 2001. Using phylogeographic analyses of gene trees to test species status and boundaries. *Mol. Ecol.* 10:779–791.

Templeton, A. R. 2004. Statistical phylogeography: Methods of evaluating and minimizing inference errors. *Mol. Ecol.* 13:789–809.

Templeton A. R., E. Routman, and C. A. Phillips. 1995. Separating population structure from history: A cladistic analysis of the geographical distribution of mitochondrial DNA haplotype in the tiger salamander, *Ambystoma tigrinum*. *Genetics* 140:767–782.

Teranishi, T., and H. Ohhama. 2004. The first capture record on histological observation of largemouth bass, *Micropterus slamoides*, ovary in Hokkaido. *Sci. Rep. Hokkaido Hatchery* 58:33–39.

Theriot, E. 1992. Clusters, species concepts and morphological evolution of diatoms. *Syst. Biol.* 41:141–157.

Thiele, K. 1993. The Holy Grail of the perfect character: The cladistic treatment of morphometric data. *Cladistics* 9:275–304.

Thiele, K., and P. Y. Ladiges. 1988. A cladistics analysis of Angophora Cav. (Myrtaceae). *Cladistics* 4:23–42.

Thomas, C. D., A. Cameron, R. E. Green, M. Bakkenes, L. J. Beaumont, Y. C. Collingham, B. F. N. Erasmus, M. Ferreira de Siqueira, A. Grainger, L. Hannah, L. Hughes, B. Huntley, A. S. van Jaarsveld, G. F. Midgley, L. Miles, M. A. Ortega-Huerta1, A. T. Peterson, O. L. Phillips, and S. E. Williams. 2004. Extinction risk from global climate change. *Nature* 427:145–148.

Thompson, D. W. 1917. *On Growth and Form*. Cambridge University Press, Cambridge. (Several later editions in various forms have been published.)

Thorpe, R. S. 1984. Coding morphometric characters for constructing distance Wagner networks. *Evolution* 38:244–355.

Tilley, S. G., and M. J. Mahoney. 1996. Patterns of genetic differentiation in salamanders of the *Desmognathus ochrophaeus* complex (Amphibia: Plethodontidae). *Herp. Mono.* 10:1–42.

Tuffey, C., and M. Steel. 1997. Links between maximum likelihood and maximum parsimony under a simple model of site substitution. *Bull. Math. Bio.* 59:581–607.

Tuomikoski, R. 1967. Notes on some principles of phylogenetic systematics. *Ann. Entomol. Fenn.* 33:137–147.

Turner, A. H., N. D. Smith, and J. A. Callery. 2009. Gauging the effects of sampling failure in biogeographic analysis. *J. Biogeogr.* 36:612–625.

Turner, H., and R. Zandee. 1995. The behaviour of Goloboffís tree fitness measure F. *Cladistics* 11:57–72.

Upchurch, P., C. A. Hunn, and D. B. Norman. 2002. An analysis of dinosaurian biogeography: Evidence for the existence of vicariance and dispersal patterns caused by geological events. *Proc. R. Soc. London, Ser. B* 269:613–621.

Valentine, J. W., and E. M. Moores. 1970. Plate tectonic regulation of faunal diversity and sea level: A model. *Nature* 228:657–659.

Valentine, J. W., and E. M. Moores. 1972. Global tectonics and the fossil record. *J. Geol.* 80:167–184.

Valentine, J. W., T. C. Foin, and D. Peart. 1978. A provincial model of Phanerozoic marine diversity. *Paleobiology* 4:55–66.

Van Dam, J. A., H. A. Azia, M. A. A. Sierra, F. J. Hilgen, L. W. van den Hoek Ostende, L. J. Lurens, P. Mein, A. J. van der Meulen, and P. Pelaez-Campomanes. 2006. Long-period astronomical forcing of mammal turnover. *Nature* 443:687–691.

Vander Zanden, M. J., J. D. Olden, J. H. Thorne, and N. E. Mandrak. 2004. Predicting occurrences and impacts of smallmouth bass introductions in north temperate lakes. *Ecol. Applications* 14:132–148.

Van Valen, L. 1976. Ecological species, multispecies, and oaks. *Taxon* 25:233–239.

Van Valen, L. M. 1982. Homology and causes. *J. Morphol.* 173:305–312.

Van Veller, M. G. P., and D. R. Brooks. 2001. When simplicity is not parsimonious: *A priori* and *a posteriori* methods in historical biogeography. *J. Biogeogr.* 28:1–11.

Van Veller, M. G. P., D. J. Kornet, and M. Zandee. 2000. Methods in vicariance biogeography: Assessment of the implementations of assumptions zero, 1 and 2. *Cladistics* 16:319–345.

Van Veller, M. G. P., D. J. Kornet, and M. Zandee. 2002. *A posteriori* and *a priori* methodologies for testing hypotheses of causal processes in vicariance biogeography. *Cladistics* 18:207–217.

Van Veller, M. G. P., M. Zandee, and D. J. Kornet. 1999. Two requirements for obtaining valid common patterns under different assumptions in vicariance biogeography. *Cladistics* 15:393–406.

Van Veller, M. G. P., M. Zandee, and D. J. Kornet. 2001. Measures for obtaining inclusive solution sets under assumptions zero, 1 and 2 with different methods for vicariance biogeography. *Cladistics* 17:248–259.

Vari, R. P. 1978. The terapon perches (Percoidei, Teraponidae). A cladistic analysis and taxonomic revision. *Bull. Amer. Mus. Nat. Hist.* 159(5):175–340.

Varon, A., L. S. Vinh, and W. C. Wheeler. 2010. POY version 4: Phylogenetic analysis using dynamic homologies. *Cladistics* 26:72–85.

Vermeij, G. 1978. *Biogeography and Adaptation*. Harvard University Press, Cambridge, MA.

Vieglais, D., E. O. Wiley, C. R. Robins, and A. T. Peterson. 2000. Harnessing museum resources for the Census of Marine Life: The FISHNET project. *Oceanography* 13(3):10–13.

Voigt, O., D. Erpenbeck, and G. Wörheide. 2008. Molecular evolution of rDNA in early diverging Metazoa: First comparative analysis and phylogenetic application of complete SSU rRNA secondary structures in Porifera. *BMC Evol. Biol.* 8:69.

Vrba, E. S. 1980. Evolution, species, and fossils: How does life evolve? *South African J. Sci.* 76:61–84.

Vrba, E. S. 1985. Environment and evolution: Alternative causes of the temporal distribution of evolutionary events. *South African J. Sci.* 81:229–236.

Vrba, E. S. 1992. Mammals as a key to evolutionary theory. *J. Mammal.* 73:1–28.

Vrba, E. S. 1993. Turnover-pulses, the Red Queen, and related topics. *Amer. J. Sci.* 293:418–452.

Waggoner, B. M. 1999. Biogeographic analyses of the Ediacara biota: A conflict with paleotectonic reconstructions. *Paleobiology* 25:440–458.

Wagner, G. P. 1996. Homologs, natural kinds and the evolution of modularity. *Am. Zool.* 36:4–13.

Wagner, G. P. 1999. A research program for testing the biological homology concept. In: *Homology. Novartis Symposium* No. 222 (G. R. Bock and G. Cardew, eds.). John Wiley & Sons, New York: 125–134.

Wagner, G. P., and C.-H. Chiu. 2001. The tetrapod limb: A hypothesis on its origin. *J. Exper. Zool. (Mol. Dev. Evol.)* 291:226–240.

Wagner, G. P., and P. F. Stadler. 2003. Quasi-independence, homology and the unity of type: A topological theory of characters. *J. Theor. Biol.* 220:505–527.

Wagner, M. 1868. *Die Darwinische Theorie und das Migrationgesetz der Organismen.* Duncker & Humbolt, Leipzig, Germany.

Wagner, M. 1889. *Die Entstehung der Arten durch raumliche Sonderung: Gesammelte Aufsatze.* Benno Schwabe, Basel, Switzerland.

Wagner, W. H., Jr. 1961. Problems in the classification of ferns. In: *Recent Advances in Botany* (International Botanical Congress 1959, Montreal). University of Toronto Press, Toronto: 841–844.

Wagner, W. H., Jr. 1983. Reticulistics: The recognition of hybrids and their role in cladistics and classification. In: *Advances in Cladistics: Proceedings of the Second Meeting of the Willi Hennig Society.* (N. I. Platnick and V. A. Funk, eds.). Columbia University Press, New York: 63–79.

Wagner, W. L., and V. A. Funk. (eds.). 1995. *Hawaiian Biogeography: Evolution on a Hot Spot Archipelago.* Smithsonian Institution Press.

Wake, M. H. 1993. Non-traditional characters in the assessment of caecilian relationships. *Herpetol. Monogr.* 7:42–55.

Wakeley, J. 2008. *Coalescent Theory: An Introduction.* Roberts and Company, Greenwood Village, CO.

Wallace, A. R. 1855. On the law which has regulated the introduction of new species. *Ann. Mag. Nat. Hist.* 16:184–196.

Wallace, A. R. 1860. On the zoological geography of the Malay Archipelago. *J. Proc. Linn. Soc., Zool.* 4:172–184.

Wallace, A. R. 1876. *The Geographical Distributions of Animals.* Harpers, New York.

Wanntorp, H.-E. 1983. Reticulated cladograms and the identification of hybrid taxa. In: *Advances in Cladistics: Proceedings of the Second Meeting of the Willi Hennig Society.* (N. I. Platnick and V. A. Funk, eds.). Columbia University Press, New York: 81–88.

Waterhouse, C. O. 1902, 1912. *Index Zoologicus.* Zool. Soc. London.

Watrous, J. E., and Q. D. Wheeler. 1981. The out-group comparison method of character analysis. *Syst. Zool.* 30:1–11.

Webb, S. D. 1978. A history of savanna vertebrates in the new world, part II: South America and the Great Interchange. *Ann. Revs. Ecol. Syst.* 9:393–426.

Weisstein, E. W. 1998. *CRC Concise Encyclopedia of Mathematics.* CRC Press, Boca Raton, FL.

Wenzel, J. W. Behavioral homology and phylogeny. *Ann. Rev. Ecol. Syst.* 23:361–381.

Weston, P. 2000. Process morphology from a cladistic perspective. In: *Homology and Systematics: Coding Characters for Phylogenetic Analysis*. (R. W. Scotland and T. Pennington, eds.). Taylor and Francis, London: 124–144.

Wetmore, A. 1960. A classification of the birds of the world. *Smithsonian Miscell. Collections* 139:1–37.

Wettstein, H. 1999. The Causal Theory of Names. In: *The Cambridge Dictionary of Philosophy* (R. Audi, ed.). Cambridge University Press, New York:124–125.

Wheeler, W. C. 1992. Extinction, Sampling, and molecular phylogenetics. In: *Extinction and Phylogeny* (M. J. Novacek and Q. D. Wheeler, eds.). Columbia University Press, New York: 205–215.

Wheeler, W. C. 1996. Optimization alignment: The end of multiple sequence alignment in phylogenetics? *Cladistics* 12:1–9.

Wheeler, W. C. 2001. Homology and the optimization of DNA sequence data. *Cladistics* 17:3–11.

Wheeler, W. C. 2003a. Implied alignment: A synapomorphy-based multiple-sequence alignment method and its use in cladogram search. *Cladistics* 19:261–268.

Wheeler, W. C. 2003b. Iterative pass optimization of sequence data. *Cladistics* 19:254–260.

Wheeler, W. C. 2006. Dynamic homology and the likelihood criterion. *Cladistics* 22:157–170.

Wheeler, Q. D., and R. Meier (eds.). 2000. *Species Concepts and Phylogenetic Theory. A Debate*. Columbia University Press, New York.

Wheeler, Q. D., and N. I. Platnick. 2000. The phylogenetic species concept (*sensu* Wheeler and Platnick). In: *Species Concepts and Phylogenetic Theory: A Debate*. (Q. D. Wheeler and R. Meier, eds.). Columbia University Press, New York: 55–69.

Wheeler, W. C., P. Cartwright, and C. Y. Hayashi. 1993. Arthropod phylogeny: A combined approach. *Cladistics* 9:1–39.

Whewell, W. 1847. *History of the Inductive Sciences, from the Earliest to the Present Time*, 2 vols. John Parker, London.

Whitcher, I. N., and J. Wen. 2001. Phylogeny and biogeography of Corylus (Betulaceae): Inferences from ITS sequences. *Syst. Bot.* 26:283–298.

White, M. J. D. 1978. *Modes of Speciation*. W. H. Freeman, San Francisco.

Whitney, G. P. 1940. The nomenclator zoologicus and some fish names. *Aust. Nat. Sydney Mag.* 1940:241–243.

Wiens, J. J. 1998a. The accuracy of methods for coding and sampling higher-level taxa for phylogenetic analysis: A simulation study. *Syst. Biol.* 47:381–397.

Wiens, J. J. 1998b. Does adding characters with missing data increase or decrease phylogenetic accuracy? *Syst. Biol.* 47:625–640.

Wiens, J. J. 1998c. Testing phylogenetic methods with tree-congruence: Phylogenetic analysis of polymorphic morphological characters in phrynosomatid lizards. *Syst. Biol.* 47:411–428.

Wiens, J. J. 1999. Polymorphism in systematics and comparative biology. *Ann. Rev. Ecol. Syst.* 30:327–362.

Wiens, J. J. 2001. Character analysis in morphological phylogenetics: Problems and solutions. *Syst. Biol.* 50:689–699.

Wiens, J. J. 2003a. Incomplete taxa, incompletel characters and phylogenetic accuracy: Is there a missing data problem? *J. Vert. Paleontol.* 23:297–310.

Wiens, J. J. 2003b. Missing data, incomplete taxa, and phylogenetic accuracy. Systematic.

Wiens, J. J., and T. L. Penkrot. 2002. Delimiting species based on DNA and morphological variation and discordant species limits in spiny lizards (Sceloporus). *Syst. Biol.* 51:69–91.

Wiens, J. J., and T. W. Reeder. 1995. Combining data sets with different numbers of taxa for phylogenetic analysis. *Syst. Biol.* 44:548–558.

Wiens, J. J., and M. R. Servedio. 1998. Phylogenetic analysis and intraspecific variation: Performance of parsimony, likelihood, and distance methods. *Syst. Biol.* 47:228–253.

Wiens, J. J., and M. R. Servedio. 2000. Species delimitation in systematics: Inferring diagnostic differences between species. *Proc. R. Soc. London, Ser. B* 267:631–636.

Wijk, R., van der (chief editor), W. D. Morgadant, and P. A. Florschütz. 1959–1969 (5 vol.). *Index Muscorum*. International Bureau of Plant Taxonomy and Nomenclature of the International Association for Plant Taxonomy, Utrecht.

Wiley, E. O. 1975. Karl R. Popper, systematics, and classification: A reply to Walter Bock and other evolutionary taxonomists. *Syst. Zool.* 24:233–243.

Wiley, E. O. 1976. The systematics and biogeography of fossil and Recent gars (Acintopterygii: Lepisosteidae). *Misc. Publ. Mus. Nat. Hist. Univ. Kansas* 64:1–111.

Wiley, E. O. 1977a. Are monotypic genera paraphyletic? A response to Norman Platnick. *Syst. Zool.* 26:352–355.

Wiley, E. O. 1978. The evolutionary species concept reconsidered. *Syst. Zool.* 27:17–26.

Wiley, E. O. 1979a. Cladograms and phylogenetic trees. *Syst. Biol.* 28:88–92.

Wiley, E. O. 1979b. Ancestors, species, and cladograms. Remarks on the symposium. In: *Phylogenetic Analysis and Paleontology* (J. Cracraft and N. Eldredge, eds.). Columbia University Press, New York: 211–225.

Wiley, E. O. 1979c. An annotated Linnean Hierarchy, with comments on natural taxa and competing systems. *Syst. Zool.* 28:308–337.

Wiley, E. O. 1979d. Ventral gill arch muscles and gnathostome phylogeny, with a new classification of vertebrates. *Zool. J. Linnean Soc.* 67:149–179.

Wiley, E. O. 1980. Is the evolutionary species fiction? A consideration of classes, individuals and historical entities. *Syst. Zool.* 29:76–80.

Wiley, E. O. 1981a. *Phylogenetics. The Theory and Practice of Phylogenetic Systematics*. Wiley-Interscience, New York.

Wiley, E. O. 1981b. Convex groups and consistent classifications. *Syst. Bot.* 6:346–358.

Wiley, E. O. 1987. Methods in vicariance biogeography. In: *Systematics and Evolution: A Matter of Diversity* (P. Hovenkamp, ed.). Utrecht University, Utrecht, the Netherlands: 283–306.

Wiley, E. O. 1988a. Vicariance biogeography. *Ann. Rev. Ecol. Syst.* 19:513–542.

Wiley, E. O. 1988b. Parsimony analysis and vicariance biogeography. *Syst. Zool.* 37:271–290.

Wiley, E. O. 1989. Kinds, individuals, and theories. In: *What the Philosophy of Biology Is* (M. Ruse, ed.). Kluwer Academic, Dordrecht: 289–300.

Wiley, E. O. 2002. On species and speciation with reference to the fishes. *Fish and Fisheries* 3:1–10.

Wiley, E. O. 2007. Species concepts and their importance in fisheries management and research. *Trans. Amer. Fisheries Soc.* 136(4):1126–1135.

Wiley, E. O. 2008. Homology. Identity and transformation. In: *Mesozoic Fishes 4: Homology and Phylogeny* (G. Arratia, H.-P. Schultze, and M. V. H. Wilson, eds.). Verlag Dr. Pfiel, Munich: 9–21.

Wiley, E. O. 2010. Why trees are important. *Evol. Edu. Outr.* 3:499–505.

Wiley, E. O. and J. G. Johnson. 2010. A teleost classification based on monophyletic groups. In: *Origin and Phylogenetic Interrelationships of Teleosts, Honoring Gloria Arratia* (J. S. Nelson, H.-P. Schultze, and M. V. H. Wilson, eds.). Verlag Dr. Friedrich Pfeil, Munich: 123–182.

Wiley, E. O., and R. L. Mayden. 1985. Species and speciation in phylogenetic systematics, with examples from the North American fish fauna. *Ann. Missouri Bot. Gard.* 72:596–635.

Wiley, E. O., and R. L. Mayden. 2000a. The evolutionary species concept. In: *Species Concepts and Phylogenetic Systematics. A Debate* (Q. D. Wheeler and R. Meier, eds.). Columbia University Press, New York: 70–89.

Wiley, E. O., and R. L. Mayden. 2000b. A critique from the Evolutionary Species Concept perspective. In: *Species Concepts and Phylogenetic Systematics. A Debate* (Q. D. Wheeler and R. Meier, eds.). Columbia University Press, New York: 146–158.

Wiley, E. O., and R. L. Mayden.2000c. A defense of the Evolutionary Species Concept. In: *Species Concepts and Phylogenetic Systematics. A Debate* (Q. D. Wheeler and R. Meier, eds.). Columbia University Press, New York: 198–208.

Wiley, E. O., and A. T. Peterson. 2003. Distributed information systems and predictive biogeography: Putting natural history collections to work in the 21st century. In: *The New Panorama of Animal Evolution. Proc. 18th Int. Congress of Zoology* (A. Legakiws et al., eds.). Pensoft Publ., Sofia, Bulgaria.:619–624.

Wiley, E. O., K. M. McNyset, A. T. Peterson, C. R. Robins, and A. M. Stewart. 2003. Niche modeling and geographic range predictions in the marine environment using a machine-learning algorithm. *Oceanography* 16(3):120–127.

Wiley, E. O., D. Siegel-Causey, D. R. Brooks, and V. A. Funk. 1991. *The Compleat Cladist, A Primer of Phylogenetic Systematics*. Spec. Publ., Mus. Nat. Hist., Univ. Kansas.

Wilkinson, M. 1992. Ordered versus unordered characters. *Cladistics* 8:375–385.

Wilkinson, M. 1995a. A comparison of two methods of character construction. *Cladistics* 11:297–308.

Wilkinson, M. 1995b. Coping with abundant missing entries in phylogenetic inference using parsimony. *Syst. Biol.* 44:501–514.

Wilkinson, M., and M. J. Benton. 1995. Missing data and rhynchosaur phylogeny. *Histor. Biol.* 10:137–150.

Willmann, R. 1986. Reproductive isolation and the limits of species in time. *Cladistics* 2:356–358.

Willmann, R. 2003. From Haeckel to Hennig: The early development of phylogenetics in German-speaking Europe. *Cladistics* 19:449–479.

Williams, D. M., and M. C. Ebach. 2008. *Foundations of Systematics and Biogeography*. Springer Science and Business Media, New York.

Williams, D. M., and D. J. Siebert. 2000. Characters, homology and three-item analysis. In: *Homology and Systematics: Coding Characters for Phylogenetic Analysis* (R. W. Scotland and T. Pennington, eds.). Taylor and Francis, London: 183–208.

Wilson, E. O. 1993. *The Diversity of Life*. The Belknap Press of Harvard University Press, Cambridge, MA.

Wilson, E. O. 1994. *Naturalist*. Island Press, Washington, DC.

Wilson, E. O., and F. M. Peter (eds.). 1988. *Biodiversity*. National Academy Press, Washington, DC.

Wilson, R. A. 1999a. Realism, essence, and kind: Resuscitating species essentialism? In: *Species: New Interdisciplinary Essays* (R. A. Wilson, ed.). MIT Press, Cambridge, MA: 187–207.

Wilson, R. A. (ed.). 1999b. *Species: New Interdisciplinary Essays*. MIT Press, Cambridge, MA.

Winsor, M. P. 2003. Non-Essentialist Methods in Pre-Darwinian Taxonomy. *Biol. Phil.* 18:387–400.

Winsor, M. P. 2006. The creation of the essentialism story: An exercise in metahistory. *Hist. Phil. Life Sci.* 28:149–174.

Wojcicki, M., and D. R. Brooks. 2005. PACT: An efficient and powerful algorithm for generating area cladograms. *J. Biogeogr.* 32:755–774.

Wood, P. 1994. *Scientific Illustration*. John Wiley & Sons, New York.

Woodger, J. H. 1952. From biology to mathematics. *British J. Philos. Sci.* 3:1–21.

Wortman, J. L. 1903. Studies of Eocene Mammalia in the Marsh collection, Peabody Museum: Primates. *Amer. J. Sci.* 15:163–176, 399–414, 419–436.

Wright S. 1931. Evolution in Mendelian populations. *Genetics* 16:97–159.

Wu, C. F. J. 1986. Jackknife, bootstrap and other resampling methods in regression analysis. *Annals of Statistics* 14:1261–1295.

Wu, C.-I. 1991. Inferences of species phylogeny in relations to segregation of ancestral polymorphisms. *Genetics* 127:429–435.

Wu, C.-I. 2001a. Genes and speciation. *J. Evol. Biol.* 14:889–891.

Wu, C.-I. 2001b. The genec view of the process of speciation. *J. Evol. Biol.* 14:851–865.

Xia, X., Z. Xie, and K. M. Kjer. 2003. 18S Ribosomal RNA amd tetrapod phylogeny. *Syst. Biol.* 52(3):283–295.

Xiang, Q.-Y., and D. E. Soltis. 2001. Dispersal-vicariance analyses of intercontinental disjuncts: Historical biogeographical implications for angiosperms in the Northern Hemisphere. *Intl. J. Plant Sci.* 162:S29–39.

Yang, Z. 1993. Maximum likelihood estimation of phylogeny from DNA sequences when substitution rates differ over sites. *Mol. Biol. Evol.* 10:1396–1401.

Yang, Z. 1994. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: Approximate methods. *J. Mol. Evol.* 39:306–314.

Yang, Z. 2006. *Computational Molecular Evolution*. Oxford Series in Ecology and Evolution, Oxford University Press, New York.

Yang, Z., N. Goldman, and A. Friday. 1995. Maximum likelihood trees from DNA sequences: A particular statistical estimation problem. *Syst. Biol.* 44:384–399.

Zachos, F. E., and U. Hoßfeld. 2006. Adolf Remane (1898–1976) and his views on systematics, homology and the modern synthesis. *Theory in Biosci.* 124:335–348.

Zandee, M., and M. C. Roos. 1987. Component-compatibility in historical biogeography. *Cladistics* 3:305–323.

Zangrel, R., and G. R. Case. 1976. Cobelodus aculeatus (Cope), an acanthodian shark from the Pennsylvanian black shales of North America. *Paleontogr., Abt. A, Bd.* 154:107–157.

Zelditch, M. L., W. L. Fink, and D. L. Swiderski. 1995. Morphometrics, homology, and phylogenetics: Quantified characters as synapomorphies. *Syst. Biol.* 44:179–189.

Zelditch, M. L., D. L. Swiderski, D. H. Sheets, and W. L. Fink. 2004. *Geometric Morphometrics for Biologists*. Elsevier, San Diego.

Zimmermann, W. 1937. Arbeitsweise der bontanischen Phylogenetik und anderer Gruppierungswissenscaften. In: *Handbunch der biologischen Arbeotsmethoden*, Abt. 3.2, Teil 9 (E. Abderhalden, ed.). Urban und Schwarzenberg, Berlin: 941-1-53.

Zimmermann, W. 1943. Die Methoden der Phylogenetik. In: *Dei Evolution der Organismsn 1, Aulf G* (G. Henberer, ed.). Justav Fisher, Jena: 20–56.

Zink, R. M. 1991. The geography of mitochondrial DNA variation in two sympatric sparrows. *Evolution* 45:329–339.

Zink, R. M. 1996. Comparative phylogeography in North American birds. *Evolution* 50:308–317.

Zink, R. M., and M. C. McKitrick 1995. The debate over species concepts and its implications for ornithology *Auk* 112:701–719.

Zink, R. M., R. C. Blackwell-Rago, and F. Ronquist. 2000. The shifting roles of dispersal and vicariance in biogeography. *Proc. R. Soc. London, Ser. B.* 267:497–503.

Zweifel, F. W. 1988. *Handbook of Biological Illustration*. University of Chicago Press, Chicago.
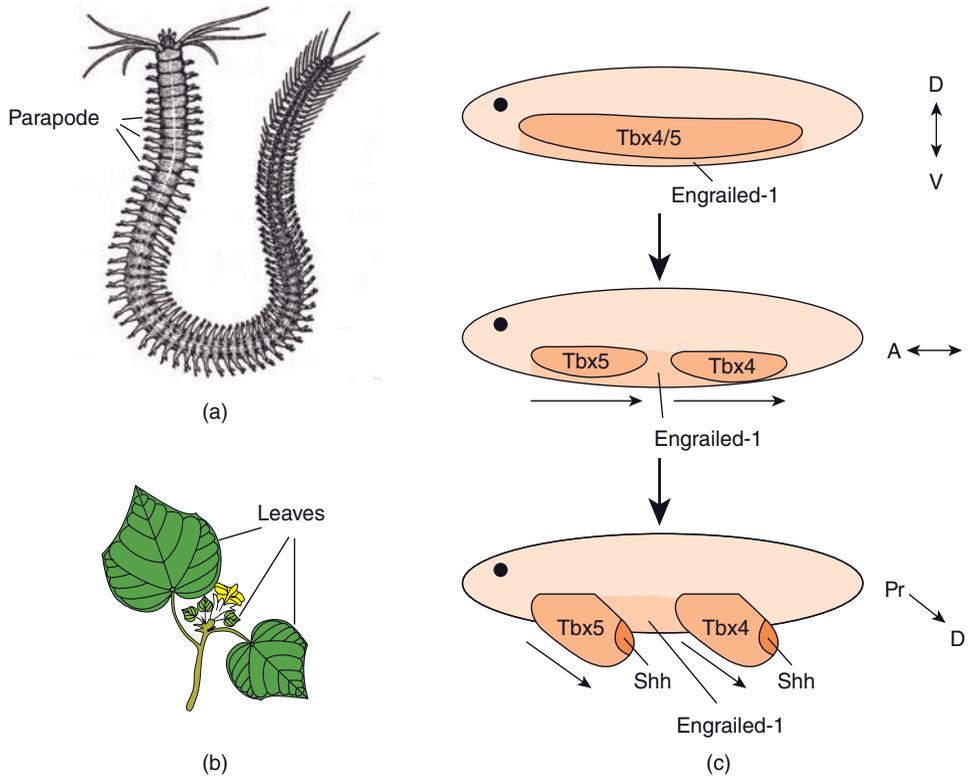
INDEX

Figure 5.1. Iterative and ontogenetic homology. Parapods of different segment (a) and leaves on a vascular plant (b) illustrate iterative homology (homonomy). (c) Ontogenetic homology: Tanaka et al. (2002) present a hypothesis of ontogenetic homology between the undifferentiated body wall cells and the paired appendages of cartilaginous vertebrates (and by inference gnathostomes). Differentiation during development is marked by expression of various signaling proteins.
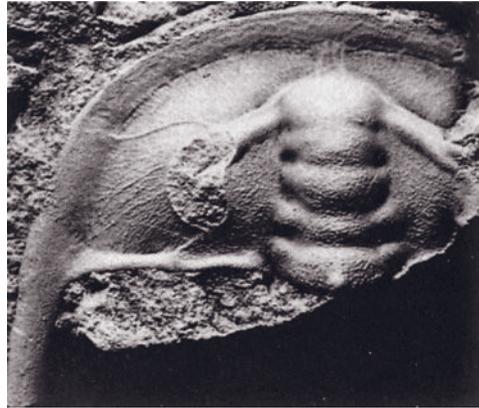
Figure 5.14. Different types of cephalic spines in trilobites illustrated using four species of Early Cambrian olenelline taxa. (a) *Bristolia harringtoni*, (b) *Olenelloides armatus*, (c) *Holmiella preancora*, (d) *Fallotaspidella musatovi*. From Palmer and Repina (1993), used with permission of the Paleontological Institute, University of Kansas.
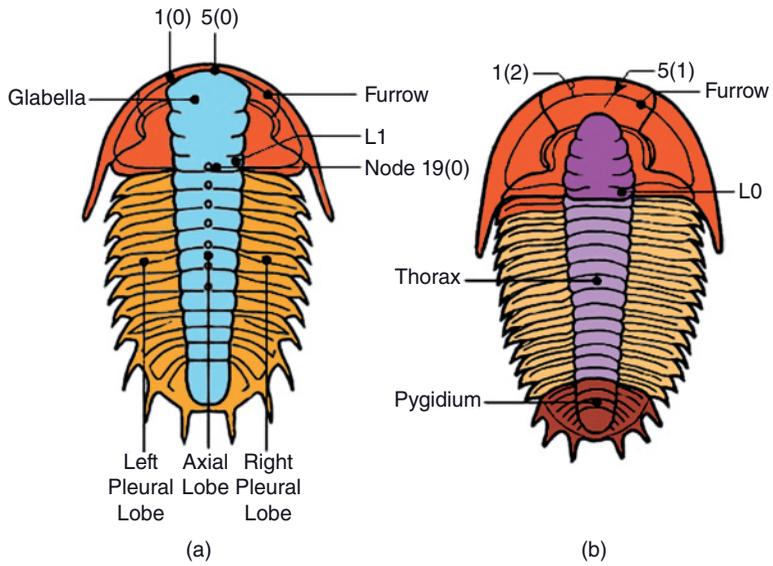
Figure 6.19. Two diagrammatic trilobites illustrating some of the characters used in the Lieberman (2002a) analysis. Key to characters: Anterior border of head shield narrow [1(0)] versus broad [1(2)]; glabella contacts furrow [5(0)] or not [5(1)]; a faint node on the occipital ring [19(0)] versus a spine [19(1)]. Used with permission of the Paleontological Institute, University of Kansas.
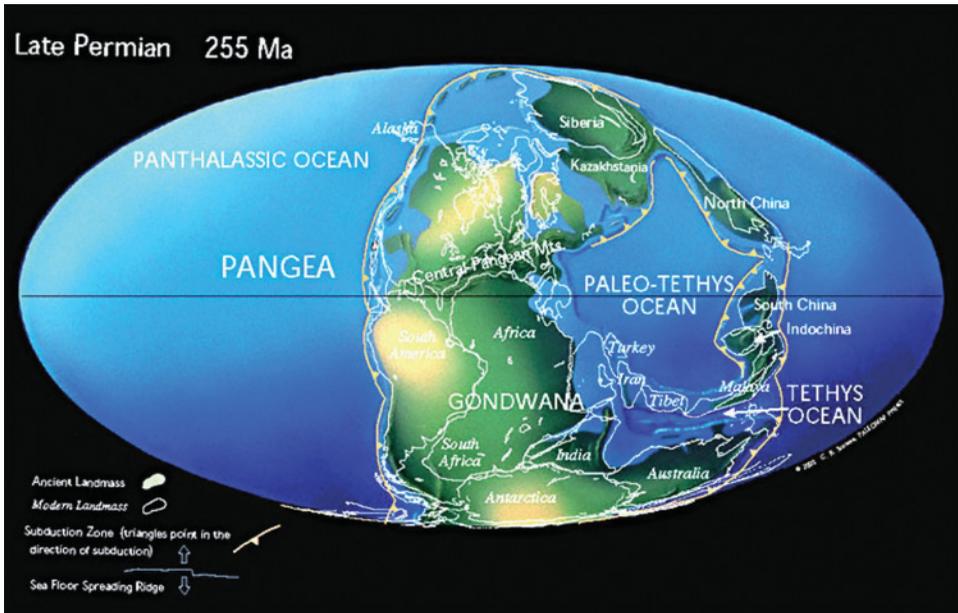
Figure 9.2. Paleogeographic reconstructions showing the approximate position of what were then the Earth's major continent blocs roughly (a) 750 and (b) 580 million years ago. These were part of a supercontinent that included Laurentia, primeval North America. The rifting that split up this supercontinent proceeded first on present day Laurentia's western margin, roughly 750 million years ago, and 150–200 million years later on its present day eastern margin. Major continental blocs are abbreviated in (b): Lau, Laurentia, North America, plus Greenland; Ama, Amazonia; Bal, Baltica; Ind, India; Aus, Australia; Sib, Siberia; Ant, Antarctica; Ara, Arabia; Arm, Armorica; Ava Avalonia. Major oceans are labeled in bold. Parts of present-day South America and Africa, which were also once distinct continental blocs, are also abbreviated including Rio, Sao, in the case of South America, and, in the case of Africa: Waf, West Africa; Con, Congo; Kal, Kalahari. Images courtesy of J. Meert, University of Florida.

Figure 9.3. The supercontinent Pangaea, in existence from roughly the end of the Paleozoic Era to the middle part of the Mesozoic Era, roughly 250–160 million years ago. Image courtesy of C. Scotese, University of Texas at Arlington, Paleomap Project.
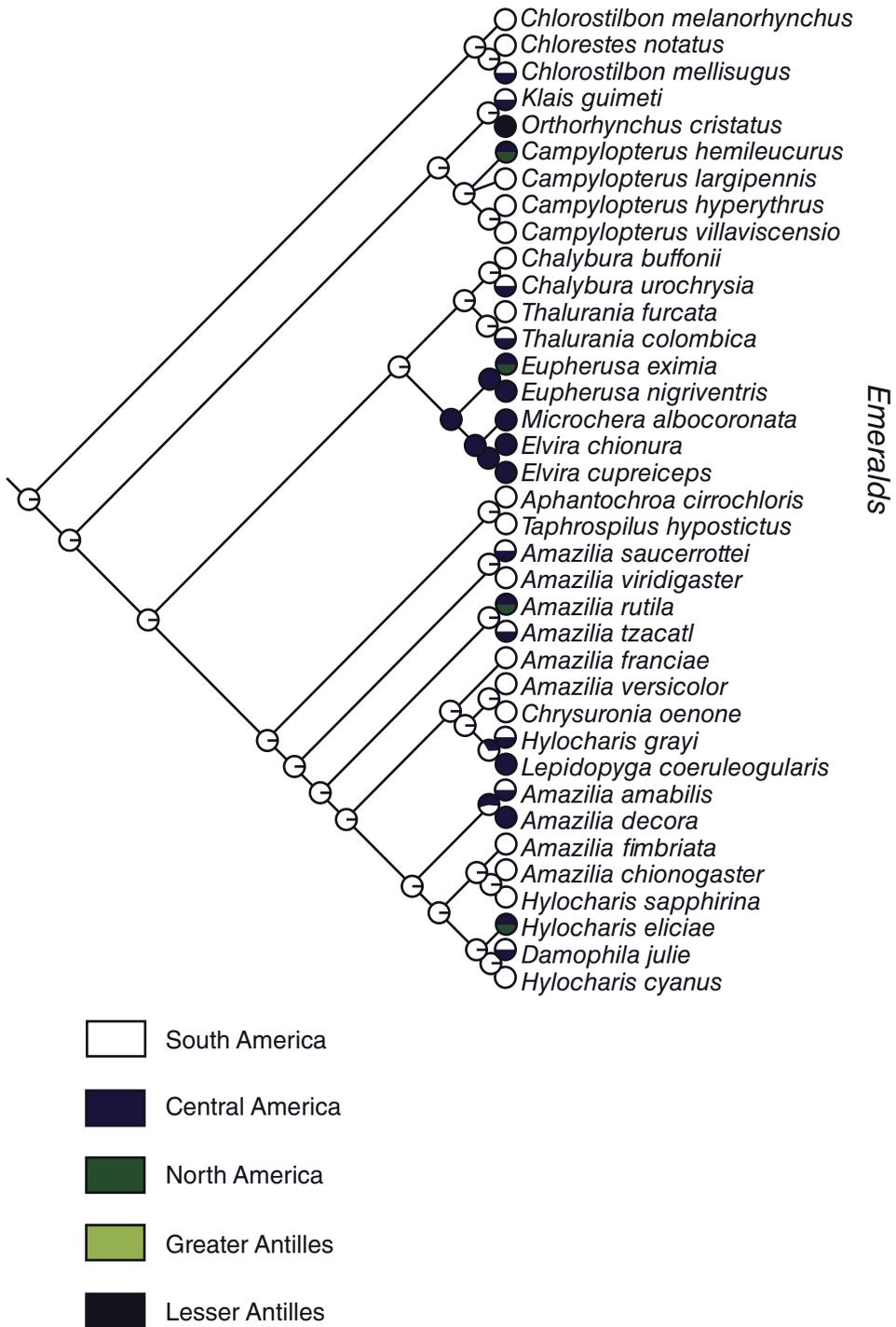
Figure 9.14. An example from McGuire et al.'s (2007) work on South American humming-birds showing the application of likelihood methods to the reconstruction and interpretation of biogeographic history. Used with permission of *Systematic Biology*, the Society of Systematic Biologists, Oxford University Press, and J. McGuire, University of California, Berkeley.
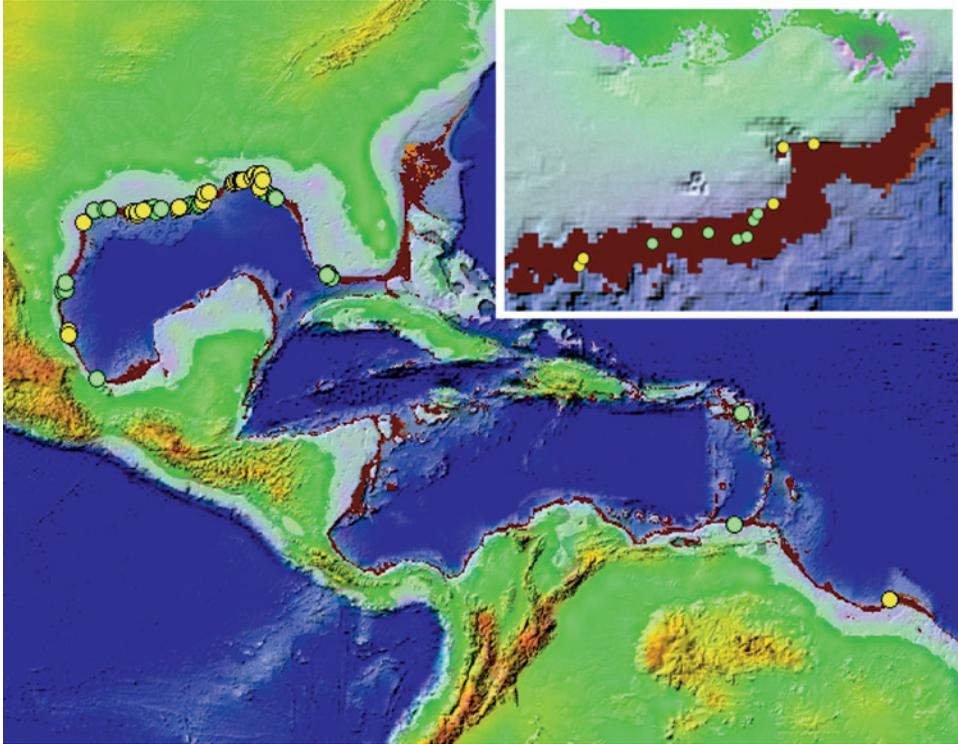
Figure 10.1. Prediction of geographic distribution of the shark *Etmopterus schultzi* in the Central Atlantic, Caribbean and Gulf of Mexico using GARP. Some point localities are used by GARP in concert with 9 WOA 98 environmental surface coverages and bathymetry. Other point localities are withheld from modeling and used to test the prediction. Blue denotes bottom depth, with lighter blue indicating relatively shallow waters. Pink to rust brown shading denotes number of model intersections: pink, 5–6; red, 7–9; rust brown, 10 intersections respectively. The inset shows details from off Louisiana. From Wiley et al. (2003), Oceanography, volume 16, number 3, Figure 2: 124, used with permission.
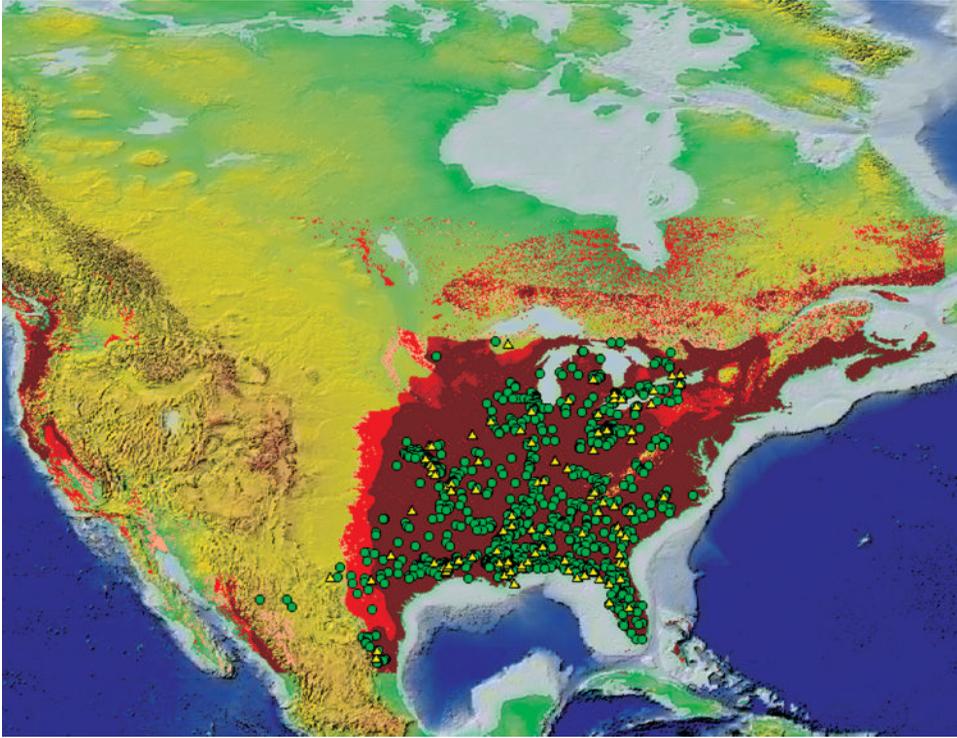
Figure 11.1. Distribution of the largemouth bass, *Micropterus salmoides*, in its presumed native range, with an ecological forecast for North America using GARP from Iguchi et al. (2004). Point data (dots and triangles) gathered from 11 museum databases via FishNet and direct access. Dots and triangles are locality data, with dots used for niche modeling and triangles for testing the resulting models. Dark red represents the joint predictions of potential range from 10 models, light red represents the joint prediction of 7–9 models. The 10 models visualized were determined by objective criteria, as discussed in Iguchi et al. (2004). From Wiley (2007). Transactions of the American Fisheries Society 136, Fig. 1, p. 1131. Used with permission of the American Fisheries Society.