

## Assembling the Tree of Life

### Where We Stand at the Beginning of the 21st Century

Few endeavors in biology, or in all the sciences, can match our quest to understand the course of life's history on Earth, which stretches across billions of years and captures the descent of untold millions of species. The notion that scientific inquiry might achieve that goal is little more than a century and a half old, and yet surprisingly, most of the species that have appeared on the twigs of the Tree of Life (TOL) have been put there only in the last decade. The systematists who have contributed to the chapters in this volume have collectively contributed a significant step toward a grand vision of systematic biology: achieving a comprehensive picture of the TOL is finally within our grasp. Darwin, Haeckel, Huxley, and the other giants who convinced the world of life's long history of change, and built the first scaffold of that history, might very well say "finally . . . it's about time"!

That it has taken so long to get to this point is testimony to the fundamental conceptual and technical challenges that have faced systematic biologists over the years. For many decades systematists had no clear theoretical or methodological idea how to recover life's history in an objective way. That challenge, as many of the greatest in the sciences, was met by deceptively simple logic. Willi Hennig, and the phylogenetic principles he developed (1950, 1966), quickly formed the foundation for quantitative, objective methodologies for comparing the characters of organisms. The technical challenges, in turn, were met when it became easier to collect new kinds of data, primarily molecular, and as computational

software and hardware improved to make these comparisons faster and more efficient.

The last major summary of our knowledge of the TOL—compiled from the 1988 Nobel symposium titled "The Hierarchy of Life" (Fernholm et al. 1989)—establishes a point of comparison with which to understand the intense work of the past decade. The phylogenetic trees presented in that volume rarely included more than 15–20 taxa, and data sets hardly exceeded 100 or 200 characters, most far fewer than that. Perusal of the journals of that time paints a similar story.

The scientific work summarized here, in contrast, manifests a huge growth in phylogenetics research. Virtually all the chapters include taxon and character samples that were unheard of a mere 10 years ago. Yet, because the focus of the chapters in this volume is the relationships among the higher taxa, even these summaries cannot convey the vast increase in our knowledge that has taken place at all hierarchical levels. For that, the reader will have to go to specialized volumes—Benton (1988), Stiassny et al. (1996), Fortey and Thomas (1998), Littlewood and Bray (2001), and Judd et al. (2002) are but five examples that have been published in recent years—as well as to the numerous journals publishing phylogenetic results in every issue.

Having knowledge of the phylogenetic relationships of life is crucial if we are to advance societal well-being, including, importantly, building a sustainable world. In this volume, the chapters by Yates et al. (ch. 1), Colwell (ch. 2), and

Futuyma (ch. 3) describe numerous examples of the contributions that phylogenetic understanding has already made to science and society. Phylogenetic relationships establish the framework for all comparative analyses of biological data, and this hierarchical structure is also a predictive tool that leads us from those characteristics we now know about organisms to those we might expect to find in those less known or newly discovered. Such logic, whether expressed explicitly or not, underlies the expectation that certain organisms might harbor pharmacologically important compounds, might be pathogenetic or toxic, might express agriculturally important gene products, and so on. Indeed, the use of phylogenetic knowledge, including analytical methods that have been developed to solve phylogenetic problems, has grown so rapidly in recent years that even a single volume devoted to the subject could not be comprehensive.

The practical outcomes and applications of TOL research are certainly a clear reason why society should continue to support a better understanding of phylogenetic relationships (see ch. 1–3; see also Cracraft 2002). Yet, what drives many scientists engaged in this effort is the sheer wonder associated with knowing a chunk of life's history. To step back and attempt to grasp the entire history of life on Earth is itself an almost unimaginable task. Here we are, one species out of hundreds of millions that have existed since the diversification of life began several billion years ago, and we are attempting to see how that history has unfolded. It is difficult enough to see how we will build the TOL for the living species, let alone for all those that vanished over the course of time, but it is an exciting prospect. All people on the planet understand something about their "genealogical roots," and that serves as a crucial metaphor for seeing how human existence and origins fit into the bigger picture of life's diversity. This is a nontrivial exercise, for truly understanding that history is bound to influence the ethical picture people develop about the importance of life forms other than our own and how these have been inextricably linked to our own well-being over time. Obviously, it is not easy for us to step back from an anthropocentric view of the world, but a TOL can facilitate such a perspective.

Darwin's vision had a profound effect on people's understanding of themselves. Yet the understanding that went along with this change in thought is not universally appreciated even today, despite 150 years of evolutionary thought and science. The TOL will be a key element in advancing an expanded vision of life's history.

### The Tree of Life: An Ongoing Synthesis

The chapters in this volume summarize our current understanding of the phylogenetic history of the major groups of organisms. It is time to stand back and see the big picture. Figure 34.1 presents a summary TOL that attempts to provide an estimate of the interrelationships among the extant

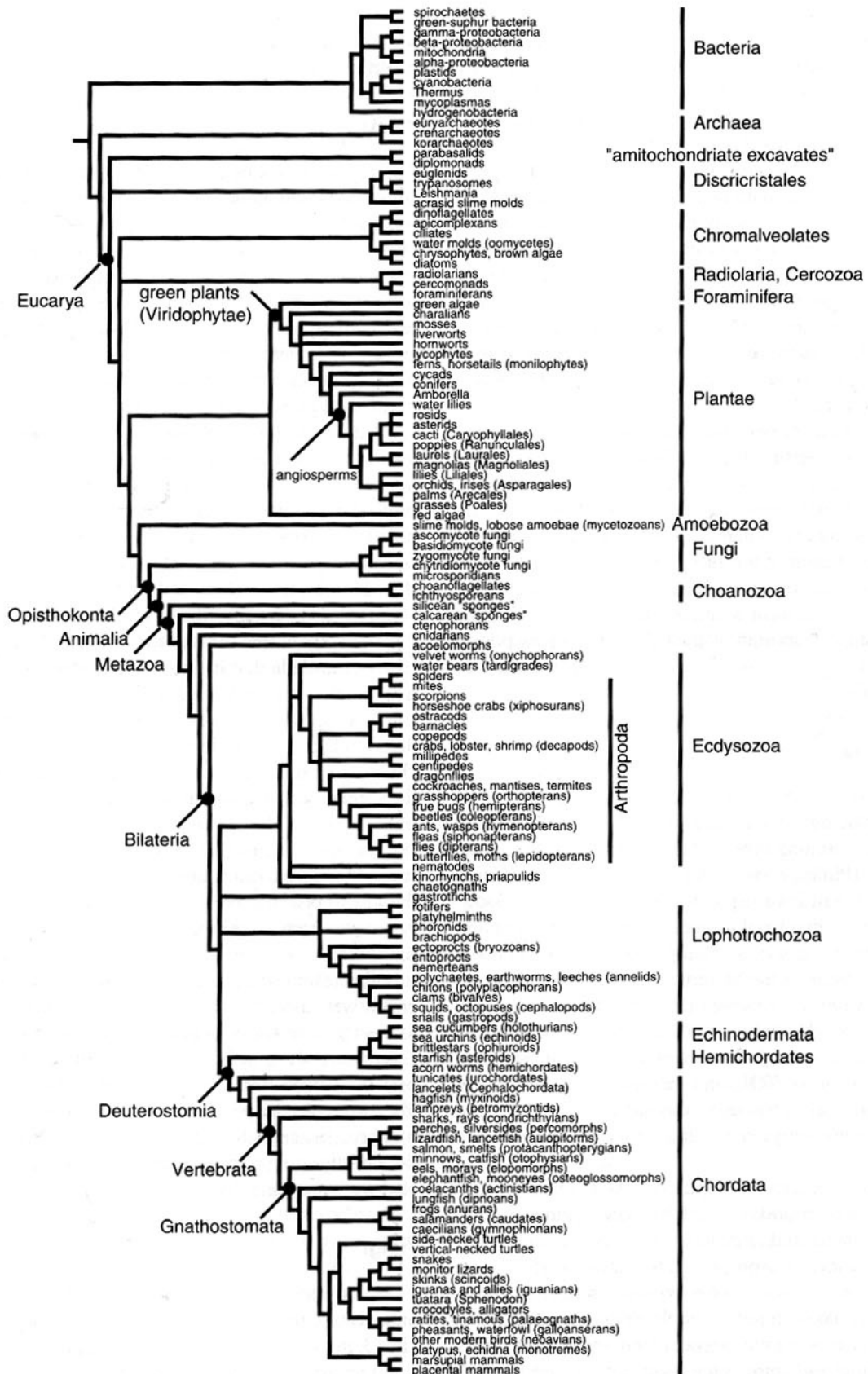
clades of life. Its scope and depth, which is skewed toward the "higher" eucaryotes, is primarily a function of the coverage of the chapters in this book, which, in turn, generally reflect known, described taxonomic diversity. Clearly, many more groups could have been added to this tree, and numerous friends working on megadiverse taxa have suggested how their favorite groups could be expanded. Yet, the best way for this tree to serve an educational purpose is to limit detail and to include groups that are familiar to a wide audience.

Conceptually, the tree is constructed as a composite—constructed by piecing together the trees presented for the different groups. It is not derived from an analysis of a "supermatrix." It attempts to represent relationships that are moderately to well supported, yet there are unresolved nodes. Some will see the tree as too conservative and would recommend resolving certain nodes; others would prefer that more nodes be depicted as ambiguous. Because the tree is not built from a data matrix, it is not a rigorous phylogenetic hypothesis in a traditional sense. Rather, it is a summary of where we are now and a step in the continuous process of building a TOL. Importantly, it also stands as a framework for discussing some of the key problems and controversies raised in the individual chapters of this volume.

### The Basal Clades of Life

It has been standard for a number of years now to recognize three major basal branches ("domains") of the TOL, the Bacteria, the Archaea, and the Eucarya (see Baldauf et al., ch. 4, and Pace, ch. 5, in this vol.), all of which are generally treated as monophyletic. A major impediment for understanding the nature of that monophyly and the relationships among these groups is, of course, the problem of where to place the root of the TOL. The present conventional wisdom is that the root lies along the branch between the Bacteria and the other two on the basis of evidence presented by duplicated genes (ch. 4). Some workers, on the other hand, have raised the issue of lateral gene transfer as possibly confounding the placement of the root (Doolittle, ch. 6), or that analytical artifacts such as long-branch attraction can lead to misleading relationships, which also could affect the placement of the root (Philippe, ch. 7). Philippe also argues that we have seen the evolutionary world as proceeding from the simple to the complex and thus have potentially overlooked the possibility that "prokaryotic"-like organisms could have been derived from eucaryotes by simplification. A major concern for all these scenarios, however, is that given the monophyly of these three groups, the placement of the root may be unsolvable because it remains a three-taxon problem.

The trailblazing work of Carl Woese, Norman Pace, and others to use the small subunit ribosomal RNA (rRNA) gene to reconstruct life's earliest branches can truly be said to have revolutionized our view of the TOL, and at the same time those data have shaped how the question of basal relationships has been approached. It is now clear that rRNA se-



**Figure 34.1.** A Tree of Life for the major groups of organisms. The relationships shown attempt to summarize those discussed in the chapters of this book. See discussion in text.

quences alone cannot resolve the branching order among bacterial lineages to a convincing degree (Pace, ch. 5).

Bacterial relationships have been strongly influenced by decades of attempts to classify using phenetic data sets of a small number of key "characters" (e.g., gram-positive vs. gram-negative staining). This approach is bound to create some nonmonophyletic taxa. Bacterial systematists have also classified these taxa at high taxonomic rank (subkingdoms, divisions, phyla) on the basis of distinctiveness, and that tradition has continued as genetically distinct forms have been discovered from environmental samples. As Pace (ch. 5) describes, there are two main groups of Archaea, the crenarchaeotes and the euryarchaeotes. The third group shown on figure 34.1, the korarchaeotes, is only represented by environmental rRNA gene sequences and is of uncertain status (see also Baldauf et al., ch. 4).

Are viruses life, or not, and what has been their history? These are the subjects of chapter 8 by David Mindell and his colleagues. Although the topic of viral phylogeny was not the subject of a talk in the New York symposium *Assembling the Tree of Life*, its inclusion in this volume was deemed important for understanding the full panoply of biotic history. Mindell et al. show that viruses have arisen multiple times, and they summarize what we understand of their evolutionary relationships. Importantly, they also discuss how phylogenetics and its methodology can be applied to issues of human health.

### Basal Eucarya

The base of the eucaryotic tree is very uncertain, with candidate groups being the parabasalid + diplomonad clade or discricristates, among others (Baldauf et al., ch. 4). Some would argue (Philippe, ch. 7) that the basal position of such taxa as parabasalids or diplomonads is probably a long-branch artifact. Their basal position seems reasonable at first glance, because it has been thought they branched off before the acquisition of the bacterial precursors of mitochondria. It is now known, however, that these "amitochondriate excavates" have some mitochondrial genes in their nuclear genomes. "Basal" eucaryotes remain one of the most unexplored regions of the TOL, and inasmuch as some groups are apparently very diverse, numerous candidates for the basal eucaryotic divergences are likely to emerge as new data are acquired.

There are three large monophyletic clades of eucaryotes, the green plants (upwards of 500,000 species), fungi (around 60,000 described), and animals (more than one million described). It is widely accepted that millions of species of fungi and animals remain to be discovered and described, whereas plant diversity has been more completely characterized. One of the more interesting phylogenetic findings of recent years is that the fungi and animals are sister taxa relative to other organisms (Opisthokonta; see Baldauf et al., ch. 4). It is important to note, however, that there are numerous single-

celled taxa whose relationships to these three clades are still unresolved; therefore, the tripartite division discussed here is certainly simplistic.

### Plants

The overall backbone of plant phylogeny is moderately well supported (Donoghue, ch. 33, and Delwiche et al., ch. 9, in this vol.). The term "algae" has been applied to a diverse array of unrelated taxa possessing plastids, some of which lie at the base of the land plants, and indeed from the perspective of Delwiche et al., the land plants simply comprise a terrestrial lineage of green algae. Although the relationships among these algal groups need much further study, current molecular evidence identifies the Charales as the sister group of the land plants (embryophytes).

Within the embryophytes, the interrelationships among the three major groups of nonvascular plants—the liverworts, hornworts, and mosses—and the vascular plants (tracheophytes) are still a matter of controversy (Delwiche et al., ch. 9). The base of the tracheophyte tree is less controversial, with lycophytes being the sister group of the rest and then monilophytes (horsetails and various "fern" groups) being the sister group of the seed plants (Pryer et al., ch. 10). Relationships within the monilophytes, and especially at the base of the clade that includes the modern seed plants, are not entirely resolved. Within the latter group, which contains some 300,000 species, the angiosperms comprise the most diverse clade. The phylogenetic unity of the clade that includes the extant "gymnosperms" is still questionable, and the sister group of all the angiosperms has not yet been identified with confidence.

The angiosperms (flowering plants) are by far the dominant group of land plants, and their interrelationships have been the subject of a large number of morphological and molecular systematic studies over the last decade. Soltis and colleagues (ch. 11) have been important contributors to this effort. They note that relationships at the base of the angiosperms are moderately well understood. One of the more remarkable findings to emerge in recent years is that *Amborella trichopoda* of New Caledonia is the only living representative of the sister group of all other angiosperms, and the next branch contains the water lilies. The three largest clades within the core angiosperms—monocots, magnoliids, and the eudicots—are well defined, but their relationships to one another and to several other smaller clades remain unresolved (ch. 11).

### Fungi

In recent years fungi have emerged as the sister group to the animals (see Baldauf et al., ch. 4, and Eernisse and Peterson, ch. 13, in this vol.). It is also becoming increasingly apparent that they will eventually be seen as one of the most diverse groups on Earth. The large-scale phylogenetic structure of the fungi has become clearer with the addition of sequence data,



and it is now accepted that the two great groups of terrestrial fungi, the ascomycetes and basidiomycetes, are monophyletic and sister taxa (Taylor et al., ch. 12). As Taylor and colleagues note, relationships within these two diverse groups are still in need of considerable study. The base of the fungal tree is also poorly understood and is occupied by lineages usually assigned to two more obscure groups, the zygomycotans and chytridiomycotans, both of which may be nonmonophyletic.

### Basal Animals

Animals are taken here to include the choanoflagellates and their sister group, the metazoans (see Eernisse and Peterson, ch. 13 in this vol.). Eernisse and Peterson review the evidence showing that animal and metazoan monophyly has become increasingly well established in recent years, but that relationships at the base of the Metazoa have been in a state of flux, particularly when it comes to those organisms typically called "sponges." Traditional classifications using morphological data recognized a monophyletic Porifera, but molecular data have led to the conclusion that siliceous sponges branched off first, followed by the calcareous sponges, the latter of which are the sister group to the eumetazoans (ch. 13). Relationships among the major clades of metazoans—ctenophorans, cniderians, placozoans, and eumetazoans—also remain uncertain because of conflicts among data sets (see ch. 13 for details).

### Bilaterians

\* The monophyletic bilaterians are composed of three main groups, the ecdysozoans, lophotrochozoans, and deuterostomes, and more and more evidence is pointing to the conclusion that acoelomorph flatworms are their sister group (see Eernisse and Peterson, ch. 13, and Littlewood et al., ch. 14, in this vol.). Intense examination of the monophyly of these groups and the interrelationships of their included taxa has essentially revolutionized our view of bilaterian evolution over the last decade by eliminating the simplistic acoelomate to pseudocoelomate to coelomate description of phylogenetic history. Although the monophyly of ecdysozoans, lophotrochozoans, and deuterostomes—particularly the latter—is increasingly accepted (at least for the "core" taxa of the first two), their interrelationships are controversial, as is the placement of a number of small, morphologically disparate metazoan groups often classified at the phylum level (Littlewood et al., ch. 14, discuss no less than 15 "phyla"). Therefore, a major question is whether there exists an ecdysozoan + lophotrochozoan clade—thus implying the classical protostome–deuterostome dichotomy.

### Lophotrochozoans

As reviewed by Eernisse and Peterson (ch. 13 in this vol.), the interrelationships among lophotrochozoan taxa are ex-

ceedingly complex and contentious due to conflicts in data, especially morphological versus molecular. Several groups are regularly recognized: (1) the lophophorates, encompassing brachiopods and phoronids; (2) the trochozoans, including the annelids and mollusks, and their allies (see fig. 34.1); and (3) the platyzoans (rotifers, platyhelminths, and others). The latter two groups have traditionally been clustered in the Spiralia on the basis of possessing spiral cleavage and a trochophore larva, although it is entirely possible that lophophorates are within the trochozoans.

The two great groups of lophotrochozoans are sister taxa, the annelids (Siddall et al., ch. 15) and the mollusks (Lindberg et al., ch. 16). Within the former, leeches and earthworms are related, but the sister group of leeches within the earthworms is still uncertain. Morphological and molecular data conflict on annelid relationships, along with those of sipunculans, relative to the diverse marine polychaete worms (ch. 15). Clearly much more work will be required to resolve the history of these groups.

The interrelationships of the major clades of mollusks are moderately well accepted (Lindberg et al., ch. 16; see also fig. 34.1), with cephalopods and gastropods being sister taxa and related to bivalves and chitons at the base of the tree. All these groups have a deep evolutionary history, with considerable fossil diversity, and an integrated picture of their phylogeny will significantly advance paleontology. Not unexpectedly, the interrelationships of the recent molluscan biota are comparatively poorly understood given their extensive diversity.

### Ecdysozoans

Different lines of evidence point to the ecdysozoans being a natural group (summarized in Eernisse and Peterson, ch. 13 in this vol.), yet many questions remain about their interrelationships, reflected in the unresolved tree in figure 34.1. Four ecdysozoan clades are now generally accepted (ch. 13 and 14): (1) the panarthropods; (2) nematodes and nematomorphs; (3) the kinorhynchs, priapulids, and loriciferans; and (4) chaetognaths. The latter two groups have low diversity, but the nematodes are thought to be the most numerically abundant metazoans on Earth, and they undoubtedly have a tremendous undescribed diversity greatly exceeding the 25,000 or so species already named. Littlewood and colleagues (ch. 14) briefly note recent progress on the phylogenetics of this group.

The arthropods—insects (Hexapoda); centipedes and millipedes (Myriapoda); crabs, crayfish, and their allies (Crustaceans); and the spiders and allies (Chelicerata)—include a number of megadiverse clades, especially the mites, spiders, and insects, and together they represent roughly 60% of the known species diversity on Earth. Wheeler and colleagues (ch. 17 in this vol.) describe the complex problem of deciphering relationships among the major groups of arthropods, the conflicting topologies implied by different data sets, and the fact that inclusion of fossil taxa in total evidence analy-

ses often has dramatic effects on phylogenetic inferences. Although most of the evidence clusters crustaceans, myriapods, and hexapods together (as the Mandibulata because they possess mandibles) to the exclusion of the chelicerates, resolving relationships among the mandibulates has not been straightforward (ch. 17).

The higher level relationships of the chelicerates are moderately well supported, with mites and spiders being sister taxa and related to scorpions and their allies, and those three, in turn, are the sister group of the horseshoe crabs (fig. 34.1; Coddington et al., ch. 18). Over the past decade, relationships among the spiders have received considerable attention, and they are the best understood of the chelicerates, whereas relationships among the diverse clades of mites remain very poorly resolved (ch. 18).

As reviewed by Schram and Koenemann (ch. 19 in this vol.), the monophyly of the crustaceans has been contentious, with morphological data tending to support monophyly and some molecular data sets denying it. Even in this volume, differences of interpretation exist: Schram and Koenemann (ch. 19) question monophyly, whereas the analyses of Wheeler and colleagues (ch. 17) generally find a monophyletic Crustacea. Many of these differences, and those in the literature, come down to *apparent* conflicts between molecules and morphology, to alternative interpretations of morphological characters, especially those of fossils, and to which clade is to be called Crustacea. There is relatively little argument (see fig. 34.1; see also ch. 19), however, that the core crustacean clades are monophyletic and related to one another, especially the maxillopods (copepods, barnacles, ostracods) and the malacostracans (crabs, shrimps, and allies).

Arguably, the greatest challenge to the TOL—as we currently understand organic diversity—is the relationships within the hexapods, or insects and their allies. The vast diversity of forms creates multiple challenges for understanding insect history. Willmann (ch. 20 in this vol.) presents a summary of the complexities of hexapod phylogeny and how viewpoints have shifted over time, and Whiting (ch. 21) discusses phylogenetic relationships within the most diverse clade of hexapods, the holometabolic insects. Arguments over insect relationships exemplify the debates in other groups—molecules versus morphology, fossil versus extant taxa. The overall structure of the insect tree, however, is remarkably consistent from one study to the next (ch. 20): aside from a number of basal groups, most insects can be clustered into the Pterygota (those with wings), at the base of which are the mayflies and dragonflies (whether sister taxa or not is in dispute) and their great sister group, the Neoptera.

No less than 80% of the insects are found in the neopteran group, the Holometabola—those insects with complete metamorphosis. This generally accepted monophyletic group contains the most familiar of the insects—beetles, butterflies, wasps, flies, and so forth—yet the interrelationships of these well-defined groups have engen-

dered considerable debate (Whiting, ch. 21 in this vol.). The tree shown in figure 34.1 includes only five holometabolic clades whose relationships appear with regularity on trees using both morphology and/or molecules (ch. 21), but many other smaller clades are omitted. Clearly, the complexity of the vast taxonomic and morphological diversity of this group will feed controversy for many years; and it seems that considerable data will be required to resolve these long-standing phylogenetic questions.

## Deuterostomia

The third great group of the bilaterians is the Deuterostomia (or Deuterostomata), which includes the echinoderms and hemichordates (ambulacrarians), on the one hand, and their sister group, the chordates. Until recently, the boundaries of the deuterostomes were ambiguous, but morphological and molecular work has clearly eliminated lophophorates, ectoprocts, and chaetognaths from the clade and established the remainder as a monophyletic group (see Eernisse and Peterson, ch. 13 in this vol.).

## Ambulacraria

Smith and colleagues (ch. 22 in this vol.) review recent advances in hemichordate and echinoderm phylogenetics. The former group is small in terms of diversity, and relationships within the group have still not been deciphered satisfactorily. Echinoderms are also not especially diverse, having only about 6000 extant species, but they possess an extensive fossil record and are among the best known marine organisms. As Smith and colleagues detail, relationships among the major monophyletic groups are moderately well supported on both molecular and morphological grounds (fig. 34.1).

## Chordata and Vertebrata

The overall pattern of chordate phylogeny is moderately well corroborated by both morphological and molecular data (fig. 34.1; see Rowe, ch. 23 in this vol.). The tunicates and lancelets are the successive sister taxa to the craniates (hagfish + vertebrates). For many years the hagfish and lampreys (fig. 34.1) were grouped together as the agnathans, but the preponderance of evidence does not favor this, especially the morphological and developmental data. Rowe notes in his review, however, that some molecular data find a monophyletic Agnatha; therefore, the problem needs further attention using combined data sets, and fossils as well as extant taxa.

Moving up the vertebrate tree, the next node subtends the sharks and allies (Chondrichthyes) and all other vertebrates (Osteichthyes), which are together termed the Gnathostomata. The Osteichthyes, in turn, are subdivided into the sarcopterygians (coelacanths, lungfish, and tetrapods) and the actinopterygian fishes. Stiassny and colleagues (ch. 24) lead us through the world of things called “fishes,” in their

case, chondrichthyans, actinopterygians, and the “fishlike” sarcopterygians. Chondrichthyans are easily divided into elasmobranchs (sharks, rays) and chimaeras, but relationships within the former clade are still uncertain. Morphological data recognize two basal sister taxa (galeomorphs and squalomorphs) and support a moderately resolved phylogeny within the latter; the conflict comes with some emerging molecular data that is said to question the monophyly of the rays and sharks (ch. 24). Within sarcopterygians, resolution of the coelacanth–lungfish–tetrapod trichotomy has been contentious. Stiassny et al. remain agnostic on this issue, whereas Rowe (ch. 23) resolves this in favor of lungfish + tetrapods while noting that the debate continues.

The actinopterygian fishes are the most diverse group of vertebrates and have a huge diversity of forms, so relationships have generally been difficult to resolve. Most of the actinopterygian nodes on figure 34.1 are based on morphological data, as Stiassny and colleagues (ch. 24) note, but new molecular data are being generated at a rapid rate. Although the interrelationships of these major groups might be generally accepted, phylogenetic understanding within most of them has a long way to go, especially given their high diversity.

It has long been accepted that amphibians are at the base of the tetrapod tree and are the sister group to all other vertebrates, which are grouped together as the Amniota (Rowe, ch. 23 in this vol.). Living amphibians are clearly monophyletic, and the relationships among the three clades have long been accepted (Cannatella and Hillis, ch. 25). Thus, caecilians are the sister group of the salamanders and frogs. Relationships within the three living taxa, especially within salamanders and frogs, are greatly unsettled.

The amniotes, so named because they share an amniote egg, are divided into two major clades, the Reptilia—including turtles, lepidosaurs (snakes, lizards, tuataras), and archosaurs (crocodiles and birds)—and the Mammalia (Rowe, ch. 23, and Lee et al., ch. 26, in this vol.). Higher level relationships within the reptiles have been particularly contentious. Crocodiles and birds go together on all trees, but the turtles, tuataras, and snakes and lizards sort out in different ways depending on the data set. There are significant conflicts across and within data sets that leave these relationships unresolved. In contrast to this somewhat dismal situation, Lee and colleagues (ch. 26) show that higher level relationships within turtles and within the lizards and snakes, for example, are becoming better understood (fig. 34.1), although at lower taxonomic levels many gaps in our knowledge still exist.

Higher level relationships within living birds remain perhaps the least understood of all the major groups of tetrapods (Cracraft et al., ch. 27). The basal split between the tinamous and ratites (paleognaths) and all other birds (neognaths), and then within the neognaths between the galliforms–anseriforms and all others (Neoaves), are well supported by various data. Phylogenetic pattern among the traditional neoavian “orders,” on the other hand, are largely

unresolved. The reason for this is pretty simple—lack of adequate character and taxon sampling—but that is rapidly changing.

Our understanding of mammalian interrelationships has made great strides in recent years because of the addition of very large morphological and molecular data sets. Yet, at the same time, as discussed by O’Leary and colleagues (ch. 28 in this vol.), there exists a great deal of conflict among data sets, and over their interpretation. All agree that monotremes are the sister group of the marsupials and placentals, but within the latter group there is considerable debate about how the traditional orders are related. The increasingly large molecular data sets appear to be converging on an answer, but morphological (and paleontological) data often conflict.

Finally, the symposium included a discussion of our current picture of hominid phylogenetics (Wood and Constantino, ch. 29 in this vol.)—for after all, it is a subject that generates great scientific and public interest and controversy. In contrast to other contributors, Wood and Constantino focus attention on the basal taxa of *Homo*—what systematists generally call species—because it is difficult to understand human evolution without delimiting those units. These authors come down on the “many species” side of the debate, as opposed to “just a few,” and they argue that deciphering relationships among these taxa is challenging because so much of the fossil material is fragmentary and difficult to compare. They also demonstrate that debates over human origins—in the sense of which species is related to which—are likely to continue for quite some time.

## Perspectives on the Tree of Life

A volume like this was not possible a decade or so ago, as a comparison with *The Hierarchy of Life* (Fernholm et al. 1989) makes clear. New analytical methods and new and more abundant data have transformed the field. But there has also been a sea change in biology’s attitude toward systematics and TOL research. Our interest in life on Earth has accelerated, not only because it is rapidly disappearing, or is in our self-interest to find new ways to make money from it, or because increased understanding will contribute to the well-being of humanity, or it is intrinsically interesting. It is all these reasons. In his perspective, E. O. Wilson (ch. 30) makes the case for “a complete account of Earth’s biodiversity, pole to pole, bacteria to whales, at every level of organization from genome to ecosystem, yielding as complete as possible a cause-and-effect explanation of the biosphere, and a correct and verifiable family tree for all the millions of species—in short a unified biology.” Amen to that. Indeed, discovering and describing biodiversity and understanding the TOL go hand in hand, and both are increasingly seen as a foundation for all of biology. Importantly, TOL research has moved into mainstream experimental and molecular biology.



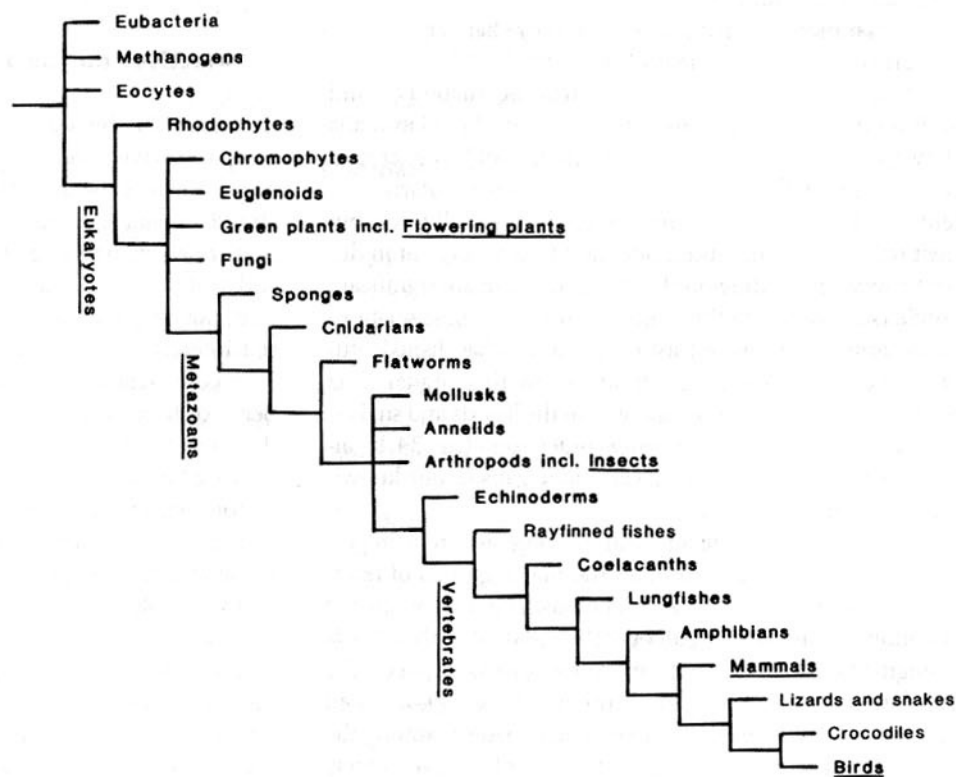
Growth in TOL research over the past decade, as David Wake (ch. 31) and David Hillis (ch. 32) observe, is readily apparent. Hillis also makes the important point that as TOL research expands, so do its applications to science and society. We are certainly on a roll, but how we might measure progress is not so straightforward, as Michael Donoghue notes (ch. 33). His tentative conclusion, discussed more below, is that it is the recognition and abandonment of paraphyletic groups that is perhaps the best measure of progress. Although it seems that some new paraphyletic groups will inevitably be created as more taxa are investigated, a successful war against paraphyly is the surest measure of success.

### The Tree of Life: Progress Against Paraphyly

A survey of previous literature leads one to the conclusion that assembling the TOL must be an exceedingly complex problem because very few have attempted to resolve the whole tree (one of the few attempts has been in the popular literature; Tudge 2000). The present volume signals that we have entered a new era of research in phylogenetics. If we look back more than a decade ago, the overall state of knowledge discussed at the 1988 Nobel symposium might appear disappointing, and in today's terms, it was. If we compare, for example, the "summary tree" from that symposium (fig. 34.2) with the one discussed here (fig. 34.1), the con-

trast is striking. As noted above, it reflects a change not only in data and data analysis but also in attitude toward what we now know we can accomplish. The latter is not to be dismissed: a decade ago, not everyone was convinced a universal tree was at hand, or possible (even for relatively small chunks of the tree). Today, the attitude of systematists has changed. We will have a universal tree, and the operative questions are when, how well supported it will be, and how we are going to create a new field of phyloinformatics to tap the tree's benefits.

Concepts of monophyly, paraphyly, and polyphyly are not really associated with phylogeny per se but how relationships map to classification. When we say that the goal of TOL research is to discover and eliminate paraphyly, we mean eliminate named groups that are not natural groups. The practical manifestation of the chapters in this book is to rid systematic biology of nonmonophyletic groups, but this activity will be resisted by some. TOL research is caught to some extent in the language of the past, in which groups are ranked on the basis of distinctness. In the past, it was morphological distinctness, but today "genetic distinctness"—however that might be measured objectively—is increasingly an important criterion. The notion that distinctness should enter into hierarchical classifications through ranking has created paraphyletic groups in its wake and hindered phylogenetic progress. The plethora of high taxonomic ranks, such as domains, kingdoms, phyla, and the like, does nothing to clarify the phylogenetic history of life.



**Figure 34.2.** A "summary" tree (hierarchy) of life of selected major groups of organisms in which those taxa underlined were the subject of discussion at the 1988 Nobel symposium (Fernholm et al. 1989).



Although we can be “immeasurably” optimistic that progress on the TOL will continue unabated, those involved in research know the task is a difficult one. A theme of the 1988 Nobel symposium was molecules versus morphology. In the early years of molecular systematics, there was an abundance of exuberance that molecules were going to sweep away morphology in reconstructing the TOL. That has not exactly happened, if one is to judge by the myriad molecular data sets that conflict with one another. Indeed, as this volume attests, more and more workers are seeking to combine molecular and morphological data, and there is a growing realization that if we are truly to have a TOL, extinct life—at least 90% of all of it—must be included. The view here is that most of the conflicts we see among different data sets are more a matter of the selection of data, method of analysis, and lack of sufficient data than they are anything substantially “wrong” with a particular kind of data. Evidence is evidence, and we should, as scientists, bring all that is relevant to bear on a problem that we can. This view echoes Colin Patterson’s (1989) closing remarks for the 1988 Nobel symposium: molecules allow us to gather large amounts of data quickly, but morphological data give us access to other dimensions of life—ontological, paleontological, temporal, and of form and function. Systematics needs all this. Biology needs all this.

### Acknowledgments

All the participants and coauthors of the chapters in this book, especially Sandie Baldauf, Mark Siddall, Ward Wheeler, Pam Soltis, Kathleen Pryer, Timothy Rowe, Douglas Eernisse, Tim Littlewood, Max Telford, and John Taylor, provided expert advice and help in constructing figure 34.1.

### Literature Cited

- Benton, M. (ed.). 1988. The phylogeny and classification of the tetrapods, vols. 1 and 2. Clarendon Press, Oxford.
- Cracraft, J. 2002. The seven great questions of systematic biology: an essential foundation for conservation and the sustainable use of biodiversity. *Ann. Missouri Bot. Garden.* 89:127–144.
- Fernholm, B., K. Bremer, and H. Jönvall (eds.). 1989. The hierarchy of life: molecules and morphology in phylogenetic analysis. Elsevier, Amsterdam.
- Forsey, R. A., and R. H. Thomas. 1998. Arthropod relationships. Chapman and Hall, London.
- Hennig, W. 1950. Grundzüge einer Theorie des phylogenetischen Systematik. Deutscher Zentralverlag, Berlin.
- Hennig, W. 1966. Phylogenetic systematics. University of Illinois Press, Urbana.
- Judd, W. S., C. S. Campbell, E. A. Kellogg, P. F. Stevens, and M. J. Donoghue. 2002. Plant systematics: a phylogenetic approach. 2nd ed. Sinauer, Sunderland, MA.
- Kenrick, P., and P. R. Crane. 1997. The origin and early diversification of land plants: a cladistic study. Smithsonian Institution Press, Washington, DC.
- Littlewood, D. T. J., and R. A. Bray (eds.). 2001. Interrelationships of the platyhelminthes. Taylor and Francis, London.
- Patterson, C. 1989. Phylogenetic relations of major groups: conclusions and prospects. Pp. 471–488 in *The hierarchy of life: molecules and morphology in phylogenetic analysis* (B. Fernholm, K. Bremer, and H. Jönvall, eds.). Elsevier, Amsterdam.
- Stiassny, M. L. J., L. R. Parenti, and G. D. Johnson (eds.). 1996. Interrelationships of fishes. Academic Press, San Diego.
- Tudge, C. 2000. The variety of life. Oxford University Press, Oxford.