

Michael J. Donoghue
Joel Cracraft

Introduction

Charting the Tree of Life

Many, perhaps even most, people today are comfortable with the image of a tree as a representation of how species are related to one another. The Tree of Life has become, we think, one of the central images associated with life and with science in general, alongside the complementary metaphor of the ecological Web of Life. But this was not always the case. Before Darwin, the reigning view was perhaps that life was organized like a ladder or “chain of being,” with slimy “primitive” creatures at the bottom and people (what else!) at the very top. Darwin (1859) solidified in our minds the radically new image of a tree (fig. 1.1), within which humans are but one of many (as we now know, millions) of other species situated at the tips of the branches. The tree, it turns out, is the natural image to convey ancestry and the splitting of lineages through time, and therefore is the natural framework for “telling” the genealogical history of life on Earth.

Very soon after Darwin, interest in piecing together the entire Tree of Life began to flourish. Ernest Haeckel’s (1866) trees beautifully symbolize this very active period and also, through their artistry, highlight the comparison between real botanical trees and branching diagrams representing phylogenetic relationships (fig. 1.2).

However, during this period, and indeed until the 1930s, rather little attention was paid to the logic of inferring how species (or the major branches of the Tree of Life) are related to one another. In part, the lack of a rigorous methodology (especially compared with the newly developing fields of genetics and experimental embryology) was responsible for

a noticeable lull in activity in this area during the first several decades of the 1900s. But, beginning in the 1930s, with such pioneers as the German botanist Walter Zimmermann (1931), we begin to see the emergence of the basic concepts that underlie current phylogenetic research. For example, the central notion of “phylogenetic relationship” was clearly defined in terms of recency of common ancestry—we say that two species are more closely related to one another than either is to a third species if and only if they share a more recent common ancestor (fig. 1.3).

This period in the development of phylogenetic theory culminated in the foundational work of the German entomologist Willi Hennig. Many of his central ideas were put forward in German in the 1950s (Hennig 1950), but worldwide attention was drawn to his work after the publication of *Phylogenetic Systematics* in English (Hennig 1966). Hennig emphasized, among many other things, the desirability of recognizing only monophyletic groups (or clades—single branches of the Tree of Life) in classification systems, and the idea that shared derived characteristics (what he called synapomorphies) provided critical evidence for the existence of clades (fig. 1.4).

Around this same time, in other circles, algorithms were being developed to try to compute the relatedness of species. Soon, a variety of computational methods were implemented and were applied to real data sets. Invariably, given the tools available in those early days, these were what would now be viewed as extremely small problems.

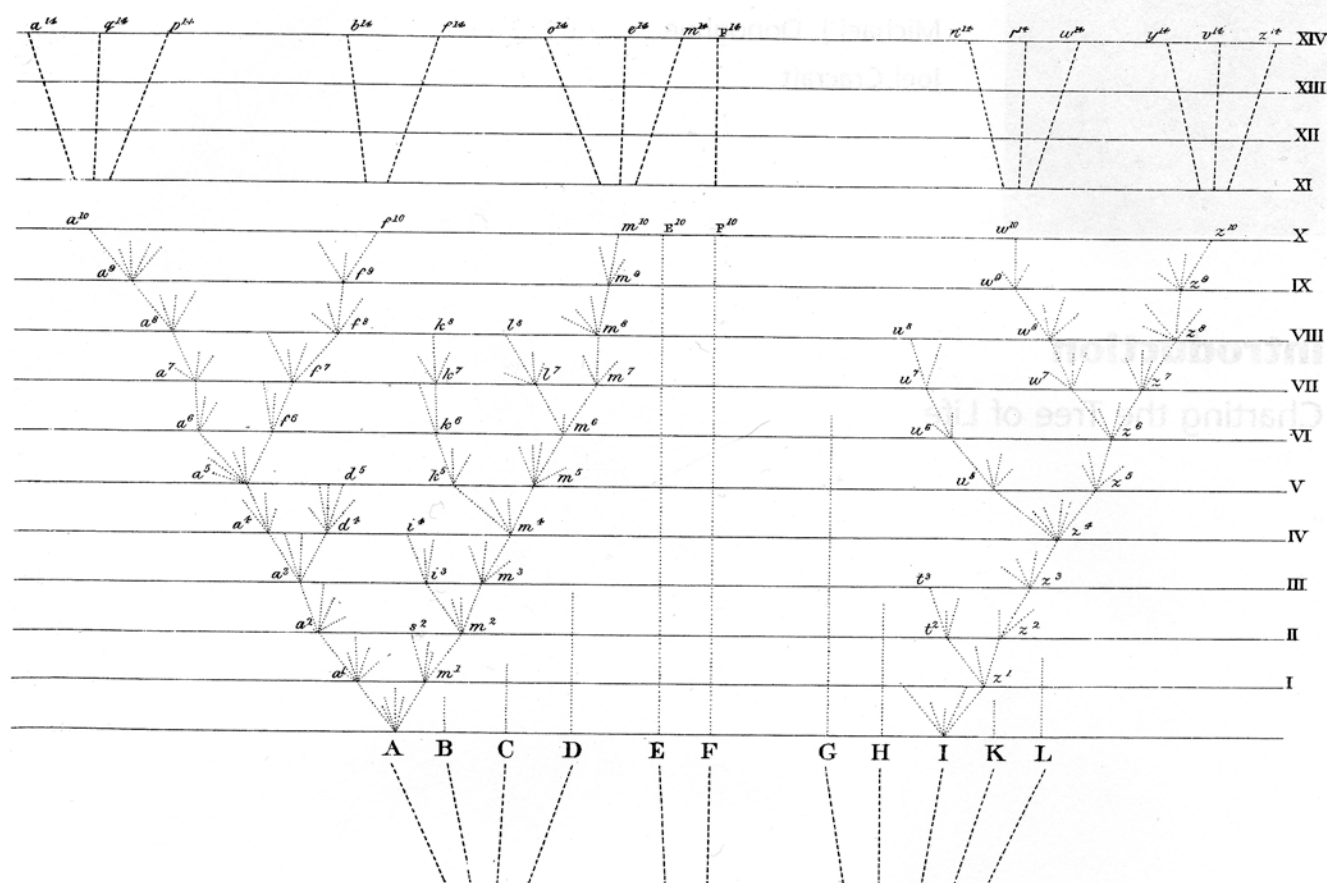


Figure I.1. The only illustration in Darwin's *Origin of Species* (1859), which can be taken to be the beginning of "tree thinking."

Since that time major developments have occurred along several lines. First, although morphological characters were at first the sole source of evidence for phylogenetic analyses, molecular data, especially DNA sequences, have become available at an exponential rate. Today, many phylogenetic analyses are carried out using molecular data alone. However, morphological evidence is crucial in many cases, but especially when the object is to include extinct species preserved as fossils. Ultimately, of course, there are advantages in analyzing all of the evidence deemed relevant to a particular phylogenetic problem—morphological and molecular. And many of our most robust conclusions about phylogeny, highlighted in this volume, are based on a combination of data from a variety of sources.

A second major development has been increasing computational power, and the ease with which we can now manipulate and analyze extremely large phylogenetic data sets. Initially, such analyses were extremely cumbersome and time-consuming. Today, we can deal effectively and simultaneously with vast quantities of data from thousands of species.

Beginning in the 1990s these developments all came together—the image and meaning of a tree, the underlying

conceptual and methodological developments, the ability to assemble massive quantities of data, and the ability to quantitatively evaluate alternative phylogenetic hypotheses using a variety of optimality criteria. Not surprisingly, the number of published phylogenetic analysis skyrocketed (Hillis, ch. 32 in this vol.). Although it is difficult to make an accurate assessment, in recent years phylogenetic studies have been published at a rate of nearly 15 a day.

Where has this monumental increase in activity really gotten us in terms of understanding the Tree of Life? That was the question that motivated the symposium that we organized in 2002 at the American Museum of Natural History in New York, and which yielded the book you have in front of you. Although it may be apparent that there has been a lot of activity, and that a lot can now be written about the phylogeny of all the major lineages of life, it is difficult to convey a sense of just how rapidly these findings have been accumulating. Previously, there was a similar attempt to provide a summary statement across all of life—a Nobel symposium in Sweden in 1988, which culminated in a book titled *The Hierarchy of Life* (Fernholm et al. 1989). That was an exciting time, and the enthusiasm and potential of this en-

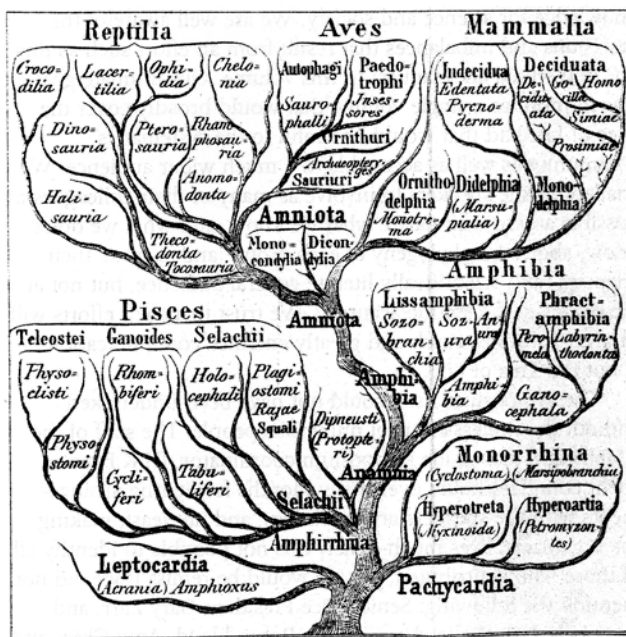


Figure 1.2. A phylogenetic tree realized by Haeckel (1866), soon after Darwin's *Origin*.

deavor were expressed in the chapters of that book. But, in looking back at those pages we are struck by the paucity of data and the minuscule size of the analyses that were being performed at what was surely the cutting edge of research at the time.

It is also clear that so much more of the Tree of Life is being explored today than only a decade ago. Now we can honestly present a picture of the relationships among all of the major branches of the Tree of Life, and within at least some of these major branches we are now able to provide

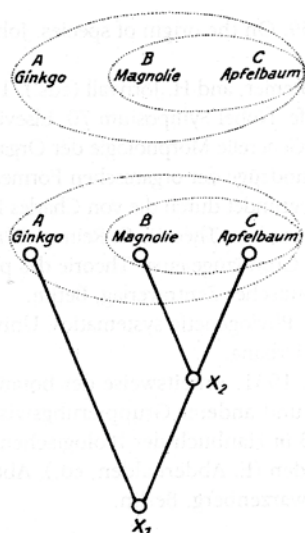


Figure 1.3. Zimmermann's (1931) tree, illustrating the concept of "phylogenetic relationship."

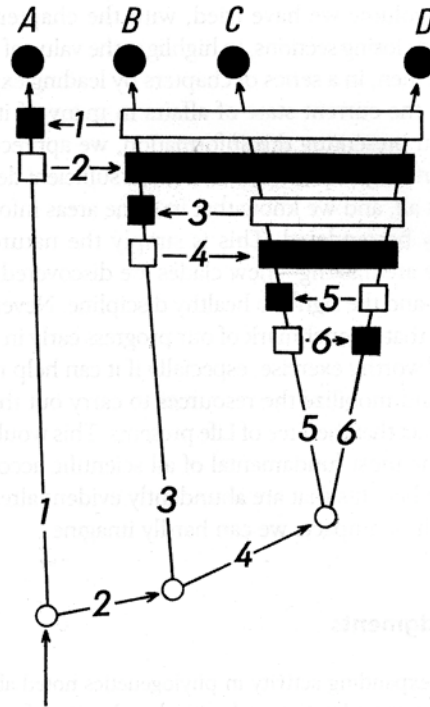


Figure 1.4. The conceptual phylogenetic argumentation scheme of Hennig (1966: 91), with solid boxes representing derived (apomorphic) and open boxes representing primitive (plesiomorphic) characters.

considerable detail. A decade ago the holes in our knowledge were ridiculously obvious—we were really just getting started on the project. There are giant holes today, which will become increasingly obvious in the years to come (as we learn more about species diversity, and database phylogenetic knowledge), but we believe that it is now realistic to conceive of reconstructing the entire Tree of Life—eventually to include all of the living and extinct species. A decade ago, we could hardly conjure up such a dream. Today we not only can imagine what the results will look like, but we now believe it is attainable.

It also has become increasingly obvious to us just how important it is to understand the structure of the Tree of Life in detail. With the availability of better and better estimates of phylogeny, awareness has rapidly grown outside of systematic biology that phylogenetic knowledge is essential for understanding the history of character change and for interpreting comparative data of all sorts within a historical context. At the same time, phylogeny and the algorithms used to build trees have taken on increasing importance within applied biology, especially in managing our natural resources and in improving our own health and well-being. Phylogenetic trees now commonly appear in journals that had not previously devoted much space to trees or to "tree thinking," and many new tools have been developed to leverage this new information on relationships.

In this volume we have tried, with the chapters in the opening and closing sections, to highlight the value of the Tree of Life, and then, in a series of chapters by leading experts, to summarize the current state of affairs in many of its major branches. In presenting this information, we appreciate that many important groups are not covered in sufficient detail, and a few not at all, and we know that in some areas information will already be outdated. This is simply the nature of the progress we are making—new clades are discovered literally every day—and the sign of a healthy discipline. Nevertheless, our sense is that a benchmark of our progress early in the 21st century is a worthy exercise, especially if it can help motivate the vision and mobilize the resources to carry out the megascience project that the Tree of Life presents. This would surely be one of the most fundamental of all scientific accomplishments, with benefits that are abundantly evident already and surprises whose impacts we can hardly imagine.

Acknowledgments

The rapidly expanding activity in phylogenetics noted above set the stage for a consideration and critical evaluation of our current understanding of the Tree of Life. This juncture in time also coincided with the inception of the International Biodiversity Observation Year (IBOY; available at <http://www.nrel.colostate.edu/projects/iboy>) by the international biodiversity science program DIVERSITAS (<http://www.diversitas-international.org>) and its partners. Assembling the Tree of Life (ATOL) was accepted as a key project of IBOY, and a symposium and publication were planned. This volume is the outgrowth of that process.

The ATOL symposium would not have been possible without the participation of many institutions and individuals. Key, of course, was the financial commitment received from the host institutions, the American Museum of Natural History (AMNH) and Yale University, and from the International Union of Biological Sciences (IUBS), a lead partner of DIVERSITAS and convenor of Systematics Agenda 2000 International. Assembling the Tree of Life (ATOL) was accepted as a core project of the DIVERSITAS program, International Biodiversity Observation Year (IBOY). We especially acknowledge the leadership of Ellen Futter (president) and Michael Novacek (senior vice president and provost) of the AMNH and of Alison Richard (provost) of Yale University for making the symposium possible. In addition, a financial contribution from IUBS facilitated international attendance, and we are grateful to Marvalee Wake (president), Talal Younes (executive director), and Diana Wall (director, IBOY) for their support.

The scientific program of the symposium was planned with the critical input of Michael Novacek and many other colleagues, and we are grateful for their suggestions. Ultimately, we tried to cover as much of the Tree of Life as possible in three days and at the same time to include plenary speakers whose charge was to summarize the importance of phylogenetic

knowledge for science and society. We are well aware of the omissions and imbalances that result from an effort such as this one and which are manifest in this volume. Our ultimate goal was to produce a single volume that would broadly cover the Tree of Life and that would be useful to the systematics community as well as accessible to a much wider audience. We challenged the speakers to involve as many of their colleagues as possible and to summarize what we know, and what we don't know, about the phylogeny of each group, and to write their chapters for a scientifically literate general audience, but not at the expense of scientific accuracy. We trust that their efforts will catalyze future research and greatly enhance communication about the Tree of Life.

The symposium itself could not have been undertaken without the tireless effort of numerous people. The staff of the AMNH and its outside symposium coordinator, DBK Events, spent countless hours over many months facilitating arrangements with the speakers and attendees, and not least, making the organizers' lives much easier. It is not possible to identify all of those who contributed, but we would be remiss if we did not mention the following: Senior Vice President Gary Zarr, and especially Ann Walle, Anne Canty, Robin Lloyd, Amy Chiu, and Rose Ann Fiorenzo of the AMNH Department of Communications; Joanna Dales of Events and Conference Services; Mike Benedetto of IT-Network Systems; Frank Rasor and Larry Van Praag of the Audio-Visual Department; and Jennifer Kunin of DBK Events.

Finally, many colleagues helped with production of this volume. Many referees, both inside and outside of our institutions, contributed their time to improve the chapters. Merle Okada and Christine Blake, AMNH Department of Ornithology, helped in many ways with editorial tasks, and Susan Donoghue assisted with the index. Most important, we are grateful to Kirk Jensen of Oxford University Press for believing in the project and facilitating its publication, and to Peter Prescott for seeing it through.

Literature Cited

- Darwin, C. R. 1859. *On the origin of species*. John Murray, London.
- Fernholm, B., K. Bremer, and H. Jönvall (eds.). 1989. *The hierarchy of life*. Nobel Symposium 70. Elsevier, Amsterdam.
- Haeckel, E. 1866. *Generelle Morphologie der Organismen: allgemeine Grundzüge der organischen Formen-Wissenschaft, mechanisch begründet durch die von Charles Darwin reformirte Descendenz-Theorie*. G. Reimer, Berlin.
- Hennig, W. 1950. *Grundzüge einer Theorie des phylogenetischen Systematik*. Deutscher Zentralverlag, Berlin.
- Hennig, W. 1966. *Phylogenetic systematics*. University of Illinois Press, Urbana.
- Zimmermann, W. 1931. *Arbeitsweise der botanischen Phylogenetik und anderer Gruppierungswissenschaften*. Pp. 941–1053 in *Hanbuch der biologischen Arbeitsmethoden* (E. Abderhalden, ed.), Abt. 3, 2, Teil 9. Urban & Schwarzenberg, Berlin.