

# Understanding Evolutionary Trees

T. Ryan Gregory

Published online: 12 February 2008  
© Springer Science + Business Media, LLC 2008

**Abstract** Charles Darwin sketched his first evolutionary tree in 1837, and trees have remained a central metaphor in evolutionary biology up to the present. Today, phylogenetics—the science of constructing and evaluating hypotheses about historical patterns of descent in the form of evolutionary trees—has become pervasive within and increasingly outside evolutionary biology. Fostering skills in “tree thinking” is therefore a critical component of biological education. Conversely, misconceptions about evolutionary trees can be very detrimental to one’s understanding of the patterns and processes that have occurred in the history of life. This paper provides a basic introduction to evolutionary trees, including some guidelines for how and how not to read them. Ten of the most common misconceptions about evolutionary trees and their implications for understanding evolution are addressed.

**Keywords** Branch · Clade · Common ancestor · Evolution · Node · Phylogeny · Sister taxa · Topology · Trend

## Introduction: The Importance of Tree Thinking

In a flourish indicative of both his literary style and perceptive understanding of nature, Darwin (1859) offered the following arboreal metaphor to describe the diversification and extinction of species:

As buds give rise by growth to fresh buds, and these, if vigorous, branch out and overtop on all sides many a feebler branch, so by generation I believe it has been

with the great Tree of Life, which fills with its dead and broken branches the crust of the earth, and covers the surface with its ever-branching and beautiful ramifications.

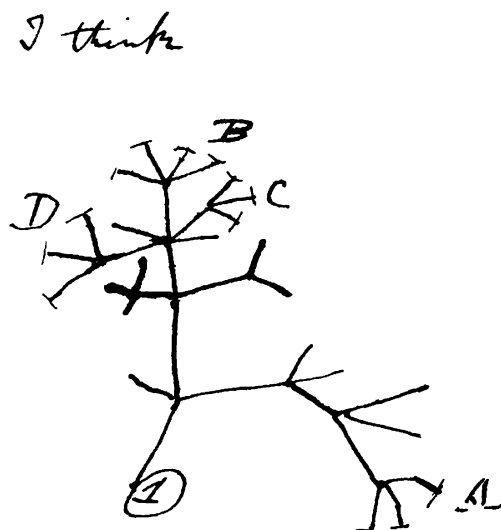
Darwin clearly considered this Tree of Life as an important organizing principle in understanding the concept of “descent with modification” (what we now call evolution), having used a branching diagram of relatedness early in his exploration of the question (Fig. 1) and including a tree-like diagram as the only illustration in *On the Origin of Species* (Darwin 1859). Indeed, the depiction of historical relationships among living groups as a pattern of branching predates Darwin; Lamarck (1809), for example, used a similar type of illustration (see Gould 1999).

Today, evolutionary trees are the subject of detailed, rigorous analysis that seeks to reconstruct the patterns of branching that have led to the diversity of life as we know it (e.g., Cracraft and Donoghue 2004; Hodkinson and Parnell 2007; Lecointre and Le Guyader 2007; Maddison and Schultze 2007). An entire discipline known as phylogenetics (Gr. *phyle*, tribe + *genesis*, birth) has emerged, complete with professional societies, dedicated scientific journals, and a complex technical literature that can be impenetrable to many nonspecialists. The output of this profession has become prodigious: It has been suggested that phylogeneticists as a group publish an average of 15 new evolutionary trees per day (Rokas 2006). Little surprise, then, that it has been argued that evolutionary biology as a whole has undergone a shift to “tree thinking” (O’Hara 1988), akin to the earlier movement toward “population thinking” that helped to shape the Neo-Darwinian synthesis around the mid-twentieth century (Mayr and Provine 1980).

Whereas tree thinking has permeated much of professional evolutionary biology, it has yet to exert its full

---

T. R. Gregory (✉)  
Department of Integrative Biology, University of Guelph,  
Guelph, ON N1G 2W1, Canada  
e-mail: rgregory@uoguelph.ca



**Fig. 1** The first evolutionary tree sketched by Darwin (1837) in one of his notebooks. It is also of note that the only illustration in *On the Origin of Species* (Darwin 1859) was an evolutionary tree. Other early evolutionists before and after Darwin, including Lamarck (1809), also drew branching diagrams to indicate relatedness (see Gould 1999)

influence among nonscientists. As Baum et al. (2005) recently pointed out, “Phylogenetic trees are the most direct representation of the principle of common ancestry—the very core of evolutionary theory—and thus they must find a more prominent place in the general public’s understanding of evolution.” In this regard, it is not so much the technical aspects of phylogenetic analysis<sup>1</sup> that are of interest but a more practical understanding of what evolutionary trees represent and, at least as important, what they do not represent. As Baum et al. (2005) continued,

Tree thinking does not necessarily entail knowing how phylogenies are inferred by practicing systematists. Anyone who has looked into phylogenetics from outside the field of evolutionary biology knows that it is complex and rapidly changing, replete with a dense statistical literature, impassioned philosophical debates, and an abundance of highly technical computer programs. Fortunately, one can interpret trees and use them for organizing knowledge of biodiversity without knowing the details of phylogenetic inference.

Unfortunately, it is becoming clear that many readers lack a sufficient level of phylogenetic literacy to properly interpret evolutionary patterns and processes. For example, a recent study of undergraduate students who had received at least introductory instruction in evolutionary science

revealed a range of common misconceptions about phylogenetic trees that represent “fundamental barriers to understanding how evolution operates” (Meir et al. 2007).<sup>2</sup> Early correction of these misconceptions would be of obvious benefit, and it has been suggested that the importance for biology students of learning how to interpret evolutionary trees is on par with that of geography students being taught how to read maps (O’Hara 1997). Given the growing significance of phylogenetic analyses in forensic, medical, and other applications (e.g., Vogel 1997; Rambaut et al. 2001; Mace et al. 2003; Mace and Holden 2005) in addition to their pervasive influence in evolutionary studies, this claim does not appear to be overstated.

This paper aims to provide a brief introduction to evolutionary trees and some basic details on how they should and should not be read and interpreted. This is followed by a discussion of ten of the most common misconceptions about evolutionary trees, many of which are held simultaneously and any of which can severely impede one’s understanding of evolution.

## The Basics of Phylogenetic Literacy

### What is an Evolutionary Tree?

In the most general terms, an evolutionary tree—also known as a *phylogeny*<sup>3</sup>—is a diagrammatic depiction of biological entities that are connected through common descent, such as species or higher-level taxonomic groupings. An overwhelming body of evidence supports the conclusion that every organism alive today and all those who have ever lived are members of a shared heritage that extends back to the origin of life some 3.8 billion years ago. One might therefore expect it to be possible, at least in principle, to reconstruct the Tree of Life, branch by branch and bough by bough, from the current diversity residing at the outermost twigs to a universally shared root. However, this proposition remains controversial—not because there is any scientific doubt about the historical relatedness of species (i.e., the fact of evolution; Gregory 2008) but because of the complex nature of evolutionary processes.

For a start, relatedness among species is a concept that depends on genetics as well as history, and there is ample evidence that even distantly diverged lineages have, at times, experienced significant gene sharing (a process known as lateral or horizontal gene transfer, in contrast to

<sup>1</sup> A discussion of phylogenetic methods is well beyond the scope of this article. Introductions to the technical aspects of phylogenetic analysis are provided by Hillis et al. (1996), Page and Holmes (1998), Nei and Kumar (2000), Felsenstein (2003), Salemi and Vandamme (2003), and Hall (2007).

<sup>2</sup> The quiz used by Meir et al. (2007) is available to instructors upon request by email (info@simbio.com). See also the “Tree Thinking Challenge” supplemental quiz by Baum et al. (2005).

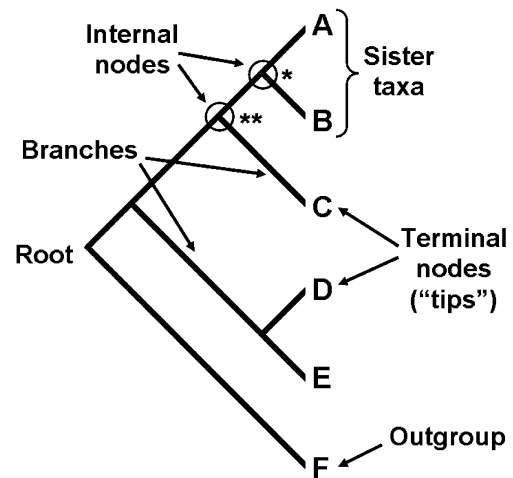
<sup>3</sup> For the purposes of this discussion and regardless of whether this will annoy some specialists, “evolutionary tree,” “phylogenetic tree,” and “phylogeny” are used interchangeably.

the more typical “vertical” transmission of genes from parent down to offspring). Some authors argue that this was sufficiently rampant in the earliest period of life’s history, and has been common enough throughout the more recent past, to create a “Web of Life” lacking any single root, rather than a strictly bifurcating tree in which branches, once split, remain separate forever (e.g., Doolittle 2000; Doolittle and Baptiste 2007). At the very least, it must be noted that in light of processes such as lateral gene transfer and gene duplication, the history of individual genes may not follow the same historical paths as those of the species in which they reside. In many cases, “gene trees” and “species trees” may not be equivalent, a fact that complicates (but does not preclude) the reconstruction of phylogenies using molecular information (e.g., Wolf et al. 2002; Rokas 2006).

These issues aside, living organisms do have a history, and this does include universal relatedness of one sort or another, be it analogous to a simple tree, a more complex web, or something else. Moreover, there is no fundamental principle that prevents the pattern of ancestry from differing both temporally and taxonomically: it is possible (but by no means confirmed) that a straightforward tree metaphor is inappropriate for, say, ancient (or perhaps even modern) bacteria but is accurate when applied to eukaryotes. In the case of the latter, at least, there may be a “true” phylogeny that accurately depicts the historical patterns of ancestry connecting eukaryote branches to their common root, but the shape of the tree is far from resolved (Baldauf 2003). In fact, except in rare instances where the pattern of evolutionary branching is created in the laboratory and observed directly as it occurs (e.g., Hillis et al. 1992; Sanson et al. 2002), it is impossible to know with certainty that any given phylogeny is historically accurate. As a result, any reconstructed phylogenetic tree is a hypothesis about relationships and patterns of branching and thus is subject to further testing and revision with the analysis of additional data. Fully resolved and uncontroversial phylogenies are rare, and as such, the generation, testing, and updating of phylogenetic hypotheses remain an active and sometimes hotly debated area of research.

### Anatomy of a Phylogeny

The old cliché contends that an undue focus on individual trees can prevent one from appreciating the grandeur of a forest. The reverse applies with regard to evolutionary trees, in that their collective importance is obvious, but many people are unfamiliar with the basic features of individual phylogenies. Whether they illustrate relationships among a few species or thousands (e.g., Bininda-Emonds et al. 2007) or of larger groupings of species (genera, families,



**Fig. 2** The anatomy of a phylogeny. An evolutionary tree includes several components. At the *right* (in this case; see Figs. 4 and 5 for alternatives) are the terminal nodes or “tips” of the tree. These typically represent individual species or larger taxonomic groups, and all are contemporaries of one another (for example, all may be living at present, in which case *A* through *F* would all represent modern species). The terminal nodes are connected to one another through branches that join at “internal nodes.” Internal nodes represent inferred lineage splitting (speciation) events that give rise to descendant sister groups—in other words, they represent the common ancestors from which two or more related lineages are descended. In this figure, the *node marked with an asterisk* represents the most recent common ancestor of species *A* and *B*, and the *one marked with two asterisks* is the most recent common ancestor of species *A*, *B*, and *C*. The pattern of branching—known as the “topology” of the tree—indicates evolutionary relatedness. For example, species *A* and *B* share a recent common ancestor that was not shared by the other species and are therefore called “sister taxa.” Similarly, species *D* and *E* are sister taxa. Species *F* is the most distantly related of the sample of species and is known as the “outgroup.” Outgroup species are necessary to “root” an evolutionary tree—that is, to indicate the last common ancestor (i.e., the deepest internal node) shared by the entire group of species. The term “basal lineage” is sometimes used to describe the branch leading to the outgroup, but this is not recommended as it is often incorrectly taken to imply that it has undergone less change and is therefore more “primitive” or “ancestral” than the other lineages (Fig. 15; see also Crisp and Cook 2005)

phyla,<sup>4</sup> etc.), all evolutionary trees provide the same basic information: a historical pattern of ancestry, divergence, and descent. They do so by depicting a series of branches that merge at points representing common ancestors, which themselves are connected through more distant ancestors.

The general anatomy of a phylogeny is summarized in Fig. 2. This tree shows the relationships among six species

<sup>4</sup> Students (including many graduate students) sometimes exhibit confusion regarding the singular and plural forms of terms such as these. “Species” is both the singular and the plural (“specie” is not a biological term—it refers to coins). “Genus” is the singular, whereas “genera” is the plural. “Phylum” is the singular and “phyla” is the plural. Other terms of interest include “taxon” (singular) and “taxa” (plural) and the widely misused “data,” which is the plural form of “datum.” While on the topic, it bears mentioning that one human is still referred to as *Homo sapiens*, which means “wise man” and does not represent the plural of “Homo sapien.”

(labeled A through F), all of which are alive at present. Each species represents one “terminal node” or “tip” on the tree. In this tree, the far right where these terminal nodes are located represents the present day, with evolutionary time extending deeper into the past as one moves from right to left on the phylogeny. These living species are located at the ends of “branches” that join one another in the past at “internal nodes.” Each internal node is taken to represent an ancestor shared by two lineages, and each branch reflects the independent evolution of the lineages that has occurred after their divergence from a given common ancestor. Ultimately, all five species share a single common ancestor at the deepest internal node, also known as the “root” of the tree. Overall, the shape of the tree and therefore the pattern of branching that it hypothesizes are known as its “topology.”

By definition, the more common ancestors that two species share to the exclusion of other species, the more closely related they are. For example, in Fig. 2, from the terminal nodes to the root, species A and B share four common ancestors, species A and D share two common ancestors, and species F shares only one ancestor (the root itself) with any of the other five species. Species A and B are linked through a recent common ancestor that is not shared by any other taxa on the tree and are therefore known as “sister taxa.” The next closest relative of species A and B is species C, with whom they share an ancestor to the exclusion of species D, E, and F. Species D and E are sister taxa and are the next closest relatives of A+B+C. Species F, by contrast, is not linked to any of the other species beyond a single distant ancestor and is known as the “outgroup.” An outgroup is necessary to root a tree (unrooted trees also can be drawn, but these are less informative and are not covered here).

## How to Read Evolutionary Trees

### Phylogenies as Family Trees

Although the technical jargon of phylogenetics may be confusing on first pass, achieving a basic understanding of evolutionary trees need not be daunting. Notably, humans in all cultures are skilled at recognizing and understanding relatedness in other contexts, and many of these abilities apply equally well to phylogenies. There are some similarities between species phylogenies and human family pedigrees, and thinking of an evolutionary tree as a “family tree” can be helpful.<sup>5</sup> This

analogy is made explicit in Fig. 3a, which shows some easily recognized relationships among human family members. In this case, the reader is depicted as one of four terminal nodes, each of which represents a contemporary person. The reader’s closest cogenerational relative is a sibling, with whom he or she shares a common ancestor (the parent) to the exclusion of other individuals. The parent herself has a sibling, both of whom are descended from the reader’s grandparent; if this tree had been drawn one generation earlier, the grandparent would have been labeled as parent, and the aunt and parent would have been labeled as siblings. Descendants of the reader’s aunt (i.e., of the parent’s sister) are the reader’s cousins. The very same basic relationships apply to the species shown in Fig. 3b.

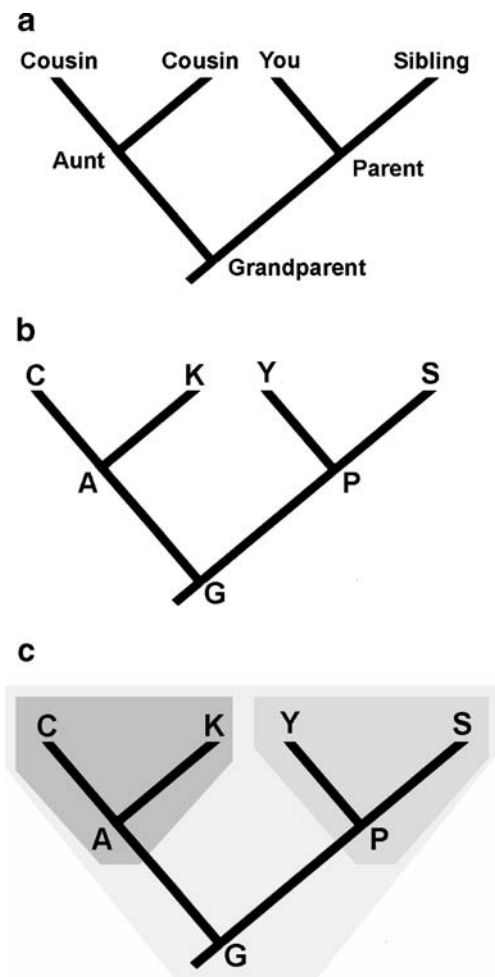
This simple comparison between phylogenies and family pedigrees highlights some other important points regarding the interpretation of evolutionary trees. First, contemporary entities (whether individual family members, species, or larger groupings) are related through common ancestors—they are not themselves ancestors of one another. Thus, the reader is not descended from a sibling; rather, both are descended from a shared parent. Likewise, the reader is not descended from a cousin, but they share a more distant common ancestor, namely their grandparent. Second, not only individual relatedness but also the relatedness of nested and increasingly inclusive groups is indicated on a tree. The reader, his or her sibling, and their shared parent represent an “immediate family,” whereas adding the cousins, the aunt, and the grandparent would also produce a coherent grouping that could be labeled more generally as a “family.” The analogous groups in phylogenetic terms, ones that include an ancestor and all of its descendants, are called “clades” (Fig. 3c). As O’Hara (1994) explained, “If you were to grab hold of the tree at any point, and cut immediately below your grip—below in the sense of toward the root—the chunk of the tree in your hand would by definition be a clade.” In other words, clades are branches that include all the twigs that have sprouted from them. Third, all members of an immediate family are equally related to individuals outside of their immediate family but with whom they share a more distant ancestor. For example, in Fig. 3a, both the reader and his or her sibling are equally related to both cousins. In like fashion, species Y and S in Fig. 3b are equally related to species C and to species K. Indeed, no matter how many descendants a parent and an aunt have, all siblings will be equally related to all of their first cousins. The same is true of species.

### Types of Trees

In the most general terms, tree diagrams are known as “dendrograms” (after the Greek for tree), whereas phylo-

<sup>5</sup> Of course, one must not take this analogy too far. Human offspring have two parents, four grandparents, and so on, whereas each species in a phylogenetic tree is usually considered to have descended from a single parental species through a branching event (speciation). In this way, a more appropriate analogy would be to a pedigree showing only the males or only the females of a family or to the family tree of individual organisms that reproduce either through asexual fission or budding.





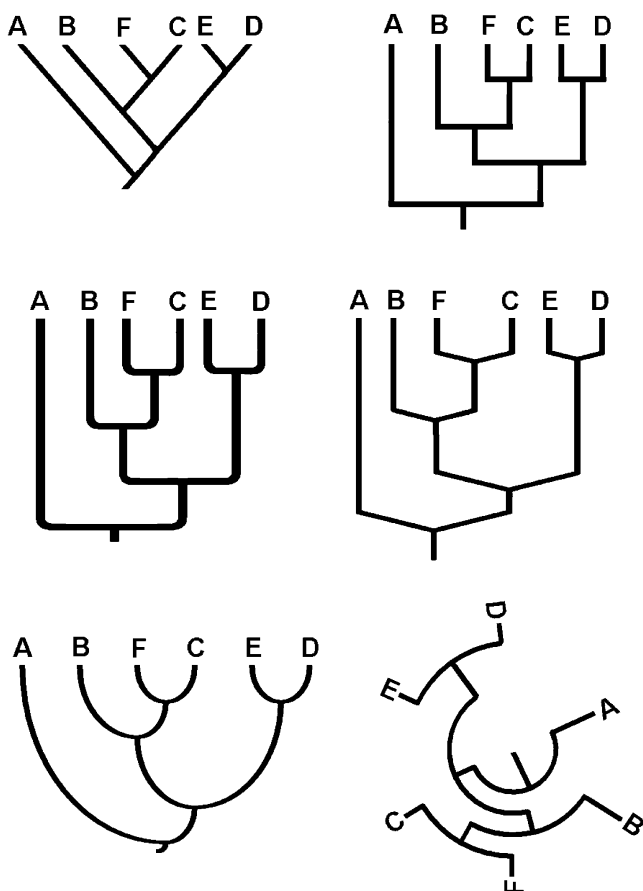
**Fig. 3** Phylogenies indicate both relatedness and historical descent. As a rough analogy, phylogenies can be compared to pedigrees of human families. **a** shows a simple “phylogeny” of human relationships. You, the reader, are indicated as one terminal node, along with a sibling with whom you share a recent common ancestor (your parent). Also depicted are two of your cousins, each of whom is equally related to you and to your sibling and with whom you share a more distant common ancestor (your grandparent). You are not descended from your sibling or from your cousin (nor are they descended from you)—rather, your relatedness is determined by the pattern of descent and ancestry going back in time. **b** shows a phylogeny of the contemporary species *C*, *K*, *Y*, and *S*. *Y* is not descended from *S* (nor vice versa), but rather these are sister taxa that are both descended from their common ancestor, *P*. Similarly, *A* is the common ancestor of the sister taxa *C* and *K*. All of these species share a more distant common ancestor, *G*. The ancestors *A*, *P*, and *G* are extinct and may be known only from fossils or may be inferred to have existed using genetic or morphological data. **c** shows the groupings of related species and their shared ancestors, which are known as “clades.” The clades in this tree include the groupings [*C*+*K*+*A*] and [*Y*+*S*+*P*], which are nested within a larger clade [*C*+*K*+*A*+*Y*+*S*+*P*+*G*]. Both *A* and *P* are descended from their common ancestor *G*, and their descendant clades (i.e., [*C*+*K*+*A*] and [*Y*+*S*+*P*], respectively) are “sister groups” of one another (*A* and *P* were themselves sister taxa). Groups that do not include the most recent shared ancestor or that include only some of an ancestor’s descendants are not clades

genies that depict only branching order are known more specifically as “cladograms.” Cladograms can be drawn in many ways. Which one is used is largely insignificant: what matters is the order of branching. Figure 4 provides examples of six common ways of drawing cladograms, each of which is exactly equivalent in terms of topology. The overall orientation of cladograms is similarly irrelevant to their accuracy, although most conventional representations place the root either at the bottom or to the left (Fig. 5). The arrangement of the tips is unimportant as well, so long as the branching patterns are maintained. In this respect, phylogenies are like a baby’s mobile: every internal node can be rotated without any implications for the pattern in which the branches are connected (Fig. 6). As will be seen, grasping the concept of freely rotating nodes can, by itself, help to correct several major misinterpretations of phylogenetic relationships and thus of the evolutionary process that generates them.

As an example of the importance of rotating nodes, consider the relationships shown in Fig. 7. In this case, the tree is “unbalanced,” meaning that one of the branches

arising from a deep common ancestor contains many more species than the other. This can reflect either a real difference in “species richness” between two lineages, or it may simply result from sampling bias in which the full diversity of one lineage (or both) is not included in the tree (e.g., Harcourt-Brown et al. 2001). Panels a and b of this figure show a “ladderized” cladogram, in which the most diverse branches are consistently positioned to the right (panel a) or left (panel b) of each internal node. However, because each node can be rotated, the tree in panel c, although perhaps somewhat more cluttered, is equally accurate.

Branching pattern is the only piece of information that can be reliably gleaned from a cladogram, regardless of how it is presented. As with the order of the terminal nodes, the lengths of individual branches on a cladogram do not convey any information whatsoever (Fig. 8a). By contrast, trees known as “phylograms” present branch lengths as being proportional to some measure of divergence between species, for example based on comparisons of deoxyribonucleic acid sequences or morphological features. These trees typically include a scale bar to indicate the degree of divergence represented by a given length of branch (Fig. 8b). Topology remains important, but in these trees, the tips are not aligned at one end of the tree, although the species they represent are no less contemporary than in a cladogram. To both align the tips and present branch lengths as being proportional to divergence or time, one of several transformation algorithms can be used to “ultrametricize” a phylogram (Fig. 8c).



**Fig. 4** Evolutionary trees can be presented in a variety of ways. This figure shows six common approaches to depicting evolutionary relationships using rooted trees. The two most common types, diagonal and rectangular, are shown at the *top*, but any of these may be encountered in the scientific literature or textbooks. In all six trees, the pattern of branching and relatedness is identical—for example,  $F+C$  and  $E+D$  are pairs of sister taxa, whereas  $A$  is the outgroup—and therefore all six are exactly equivalent. The reader is encouraged to confirm that the pattern of branching is the same in all six trees

Finally, some evolutionary trees are designed to provide information not only on branching order and time but also regarding features such as relative species diversity, geographical distribution, or ecological characteristics (Fig. 9). These are neither cladograms nor phylograms, and they may not provide any explicit information on the branching order of individual species within larger groups. Instead, they usually represent large-scale evolutionary trees that are drawn using fossil data and other information to provide a general overview of the history of a lineage.

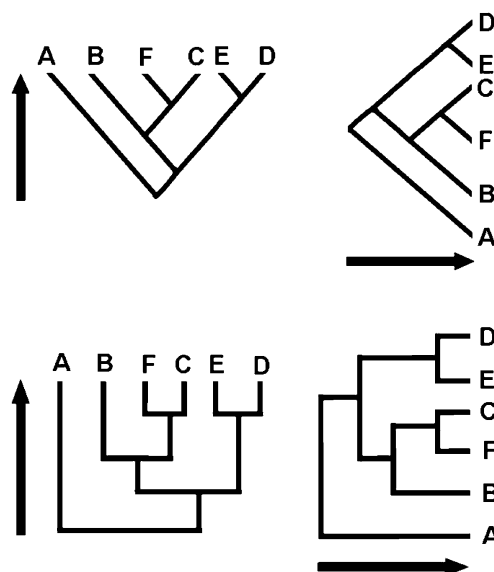
### How Not to Read Evolutionary Trees

Misunderstandings of evolutionary trees are pervasive among students, in the media, and among other non-specialists. Even more alarming, they also surface frequently

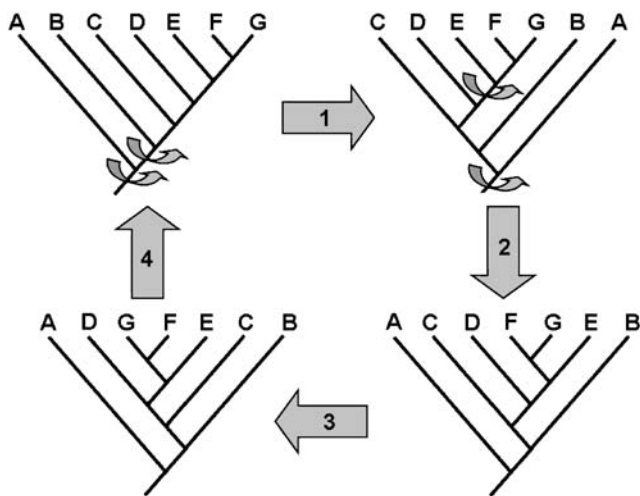
in the peer-reviewed scientific literature, often with significant implications for the conclusions drawn from comparative analyses (see Crisp and Cook 2005 for several examples). The following sections describe and seek to correct ten of the most commonly encountered misconceptions about evolutionary trees. Several of these are interrelated and therefore overlap to an extent, but each can be illustrated using distinct examples. Learning (and teaching students) to avoid these misunderstandings represents a key step toward the development of adequate tree thinking skills.

#### Misconception #1: Higher and Lower

Notions of a “Great Chain of Being” or *scala naturae* (scales of nature), in which living species (and, in some cases, nonliving matter and/or the divine) are ranked from lowest to highest and extend back at least as far as Aristotle. Although Darwin (1837) himself noted early on



**Fig. 5** Temporal directionality on a phylogeny. Regardless of the type of tree depicted (Fig. 4), a rooted tree can be read as indicating the earliest ancestor at the root, from which are descended the internal nodes and, more recently, the terminal nodes. That is to say, evolutionary trees indicate the passage of time beginning from the root (oldest) to the terminal nodes (youngest). Time cannot be read in any other direction on the tree (for example, across the tips), because all terminal nodes represent contemporary species (see Figs. 11 and 16). On all four trees shown here, the *arrow* indicates the direction of time from earliest ancestor (at the *root*) to modern species (at the *tips*). Trees are most commonly oriented to face up or right, but this is convention only, and downward or leftward trees would be equally accurate. Note that trees such as these do not imply specific amounts of time per branch, nor do they indicate when particular branching events occurred; they merely indicate the historical order of branchings within lineages. For example, this tree indicates that the split between the lineage leading to species  $D$  and  $E$  occurred sometime after the split of lineages from the common ancestor of  $D+E+C+F+B$ . By contrast, it does not indicate that the  $D+E$  and  $C+F$  splits occurred at the same time



**Fig. 6** All internal nodes can be rotated without changing the topology. Although they may look quite different, the four trees shown here are exactly equivalent to one another. This is because it is the order of branching events—the topology—that is relevant. Each internal node can be rotated without affecting the topology because this does not alter the groupings of species. For example, all four trees retain  $F+G$  as sister taxa, with the next closest relative of these species being  $E$ , then  $D$ , and so on. The reader is encouraged to confirm that the topologies of these trees are identical. To convert from the tree at the *top left* into the one at the *top right* (arrow 1), one need only rotate the node joining  $B$  to  $[C+D+E+F+G]$  and the node joining  $A$  to  $[B+C+D+E+F+G]$ , as indicated in the figure. To convert this second tree to the one at the *bottom right* (arrow 2), one must rotate the node joining  $[E+F+G]$  and the node joining  $A$  to  $[C+D+E+F+G+B]$ . The reader is invited to identify the node rotations necessary to convert this third tree to the one at the *bottom left* (arrow 3) and thence back to the original tree at the *top left* (arrow 4)

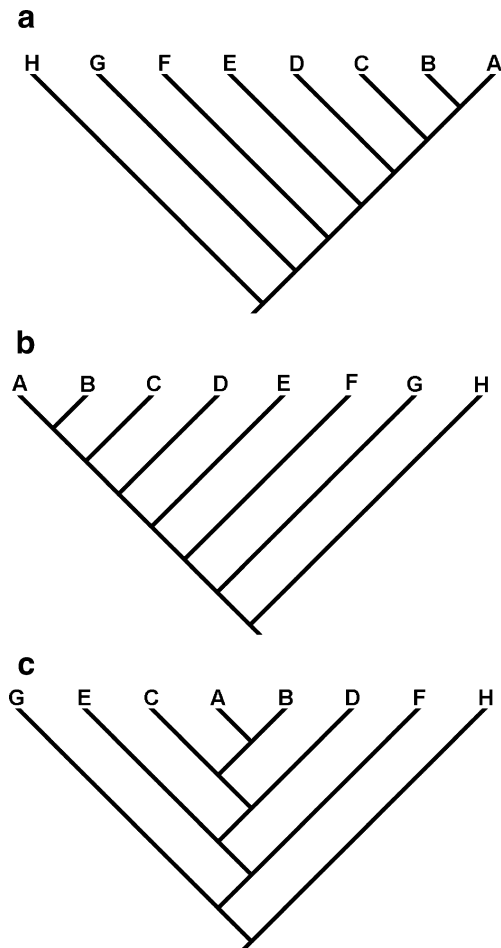
that “It is absurd to talk of one animal being higher than another,” in many respects, his contribution merely shifted the explanation for the perceived rankings, replacing the scales of nature with an “evolutionary scale” or “evolutionary ladder” (Ruse 1996). Talk of “higher” and “lower” organisms, made in reference to contemporaneous species, persists in both public and professional scientific discourse. Not surprisingly, humans typically are (self-)designated as the “highest” organisms, with other living species ranked as higher or lower on the “evolutionary scale” according to how similar they are to this particular terminal node on the phylogeny of animals.

As many prominent authors have noted, there is no scientifically defensible basis on which to rank living species in this way, regardless of how interesting or unique some aspect of their biology may be to human observers (e.g., Dawkins 1992; Gould 1994, 1996). This error does not so much reflect a specific misunderstanding of phylogenetic diagrams per se but a failure to grasp the very concept of common descent. Therefore, the adjustment to be made in this case is from imagining evolution as a linear, progressive process that generates ladder-like ranks

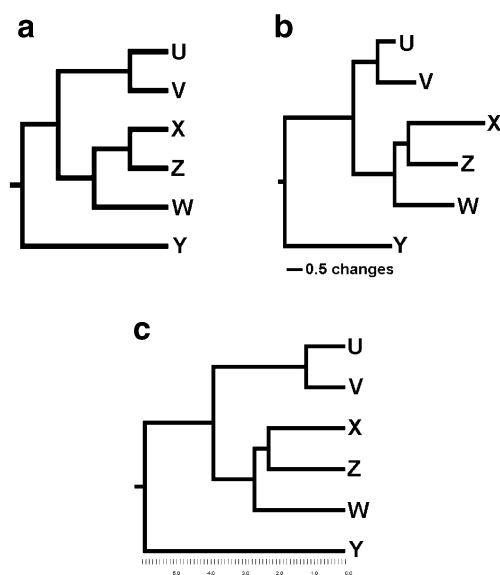
to one of branching and diversification of which trees are the result (e.g., O’Hara 1992, 1997; Nee 2005).

#### Misconception #2: Main Line and Side Tracks

Although it is clearly a critical first step, recognizing evolution as tree-like does not in itself eliminate progres-



**Fig. 7** Unbalanced trees can be ladderized or nonladderized without changing the topology. Many evolutionary trees are “unbalanced,” meaning that not all of the sister groups contain the same number of species. This can be due to real differences in diversity among groups or to incomplete sampling in which not all contemporary species are included in the tree. Figure 3b shows a balanced tree, but the trees depicted here are unbalanced because the major branches do not contain an equal number of species (i.e., one branch from the root contains only one species,  $H$ , whereas the other branch includes species  $G, F, E, D, C, B$ , and  $A$ ). In **a** and **b**, the trees are “ladderized,” which means that they are presented with the most diverse sister group on the same side of every internal node. **c** shows the same tree, with an identical topology, in a nonladderized format. This was done simply by rotating several nodes (Fig. 6) so that more diverse groups do not always appear to the right (**a**) or left (**b**) of the internal nodes. Although ladderized trees appear less cluttered, they are no more accurate than nonladderized ones, and in fact they may cause readers to falsely interpret the information provided in the tree (Figs. 11 and 14; see also Crisp and Cook 2005)



**Fig. 8** Information other than topology requires different trees. The tree shown in **a** is known as a “cladogram” and is the same as the others in the paper in that the information contained within it is limited to branching order; the lengths of the branches in such trees do not convey any information. However, other types of trees can be drawn to indicate additional information, such as the degree of genetic or morphological divergence between species. **b** shows a special kind of phylogeny known as a “phylogram,” in which branch length is proportional to some measure of divergence. Phylograms typically include a scale bar to indicate how much change is reflected in the lengths of the branches. Total divergence between two species is determined on a phylogram by adding up the entire length of branches separating them: from one species to the common ancestor and then from the ancestor to the second species. In this tree, the lineage leading to species *U* has undergone less change than the lineage leading to species *V* since these lineages split from a common ancestor. Conversely, the lineage leading to *X* has undergone more change than the lineage leading to *W* since these lineages diverged from their most recent shared ancestor (which includes another split that led to *Z*—recall that neither *Z* nor *W* is an ancestor of *X*). It is important to note that, as with the cladogram in **a**, all species *U*–*Z* in **b** are contemporary species. To make this clearer, trees such as those in **b** are sometimes “ultrametricized” as in **c**, meaning that the terminal nodes are aligned with each other and the internal branch lengths are scaled to show the degree of divergence among sister groups rather than among individual species. Alternatively, a tree like that in **c** can be scaled to time (e.g., in millions of years before the present), for example if fossil or “molecular clock” data are available for calibrating the specific timing of branching events (Benton and Ayala 2003; see also Fig. 9)

sionist interpretations of life’s history. Even those who acknowledge the branching nature of evolutionary change may continue to interpret it as a progressive process in which a “main line” has led to a distinct endpoint (namely *Homo sapiens*). In this narrative, all other modern species are derivatives of “side tracks,” anomalous offshoots of the main line to humans that all went astray for one reason or another. Even Huxley (1880) fell prey to this line of thinking when he suggested that the teleost fishes “appear to me to be off the main line of evolution—to

represent, as it were, side tracks starting from certain points of that line.”

Figure 10a provides an illustration of how Huxley could reach such a false conclusion while still accepting the basic concept of tree-like branching. This represents an unbalanced, right-ladderized tree with representatives of several vertebrate lineages, including the cartilaginous fishes, teleost fishes, amphibians, birds, and the mammalian lineage as represented by humans. For many, an intuitive interpretation of this tree is that humans represent an endpoint of a “main line,” with all other lineages branching off from this line at some time in the past.

Two points can help to correct this misconception. First, recall from Fig. 3 that all siblings are equally related to their shared cousins. In this tree, all members of the clade that includes frogs, birds, and humans (tetrapods) are equally related to all members of the clade that includes goldfish and trout (teleost fishes). Second, a simple rotation of a few internal nodes or adding a better representation of some of the most diverse groups, as reflected in Fig. 10b, completely changes this perception without any effect on the accuracy of the tree. Few readers would interpret Fig. 10b as implying that a main line exists with perch as the endpoint and humans and other mammals as a minor, early branching offshoot, even given the fact that teleost fishes make up roughly 50% of all vertebrate diversity, whereas mammals represent about 10%. Still, the logic behind such an obviously silly reading of this tree is no more fallacious than of the intuitive “main line” interpretation of Fig. 10a. To quote O’Hara (1992),

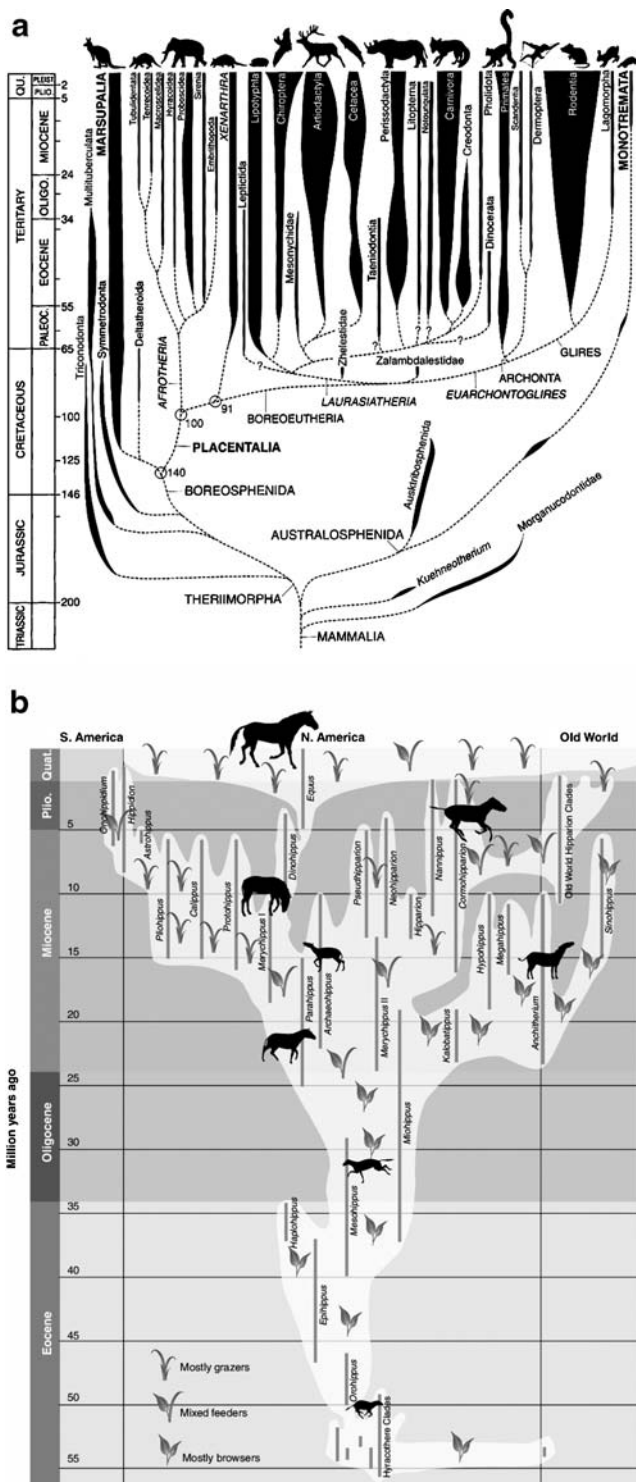
When we come to realize that even among the vertebrates there are 50,000 different ‘vertebrate stories’, each one with a different ending and each one with a different narrative landscape; when we truly think in terms of the diverging tree, instead of the line; when we understand that it is absurd to talk of one animal being higher than another; only then will we see the full grandeur of the historical view of life.

As a matter of fact, it is most likely that evolutionary history will be misconstrued as representing a progressive “main line” when there is only one obvious endpoint available. In what he called “life’s little joke,” Gould (1991) noted that only *unsuccessful* lineages with very few living representatives are taken as endpoints of a supposed main line.

### Misconception #3: Reading across the Tips

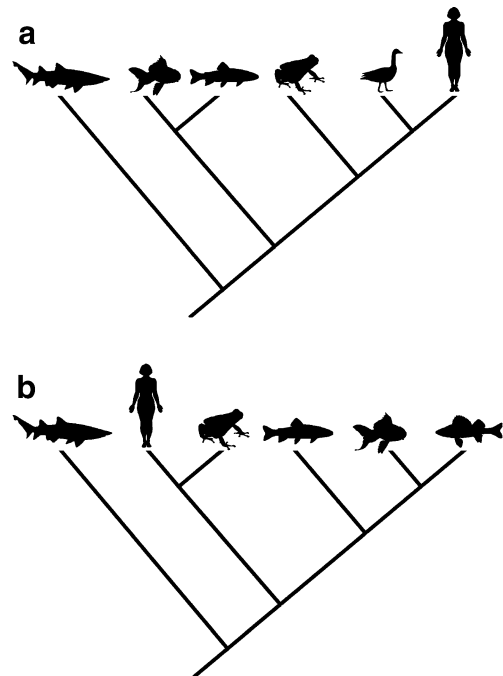
Referring to a cladogram similar to the one shown in Fig. 11a, Baum et al. (2005) asked readers to consider the following question on the basis of their reading of the tree: “Is the frog more closely related to the fish or the human?” The expectation, which has been reinforced by additional studies





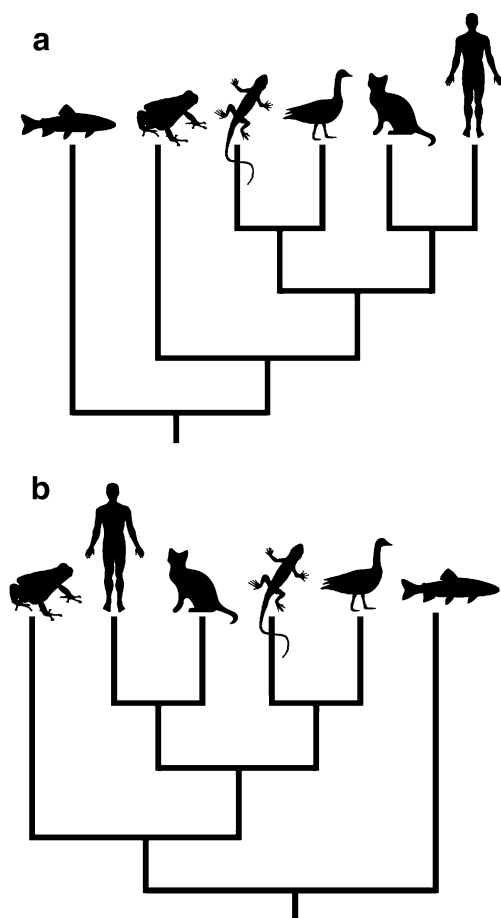
**Fig. 9** Some evolutionary trees include information about time and diversity. **a** shows an example of an evolutionary tree that includes not only information about topology but also time as given in the axis at the left and relative species diversity as indicated by the width of branches. Note that not all branches are of equal length because the tree is scaled to geological time periods and includes lineages that are extinct and therefore do not extend to the present. From Benton (2005), reproduced by permission of Blackwell. **b** shows an example of an evolutionary tree that provides information about geographical distribution and feeding ecology but that provides only a general indication of evolutionary relationships rather than explicit links between individual species (but see MacFadden 1992). From MacFadden (2005), illustration by P. Huey. Reprinted by permission of the American Association for the Advancement of Science

As a means of correcting this misinterpretation, one may take the time to identify the clades depicted in the tree (Baum et al. 2005). Humans, cats, and their common mammalian ancestor represent one clade, as do birds, lizards, and their common ancestor. These lineages together



**Fig. 10** There are no “main lines” or “side tracks” in evolution. Undoubtedly, many readers will consider the tree depicted in **a** to reflect a main line of evolutionary progress from a primitive ancestor to an “advanced” species like humans, with other groups such as cartilaginous or bony fishes appearing as side tracks off that line, despite the fact that roughly half of all vertebrate species are teleost fishes (and only 10% are mammals). Notably, the tree in **b** is equally valid and by the same false logic would have perch as the endpoint of an assumed main line and all terrestrial vertebrates, including humans, as an apparent side track. It is important that the positions of terminal nodes, all of which represent contemporary species, not be mistaken as having some significance, because they do not (see also Fig. 11). Note also that humans are more closely related to bony fishes than either is to sharks. Phylogenetically speaking, “fish” is an invalid category resulting from different rates of morphological change among lineages and does not reflect real relationships

(Meir et al. 2007), is that many people intuitively interpret the tree as indicating a closer relatedness between frog and fish than between frog and human. They do so because they incorrectly read meaning into the order of the terminal nodes, rather than assessing the pattern of branching that links these contemporary tips to one another historically.

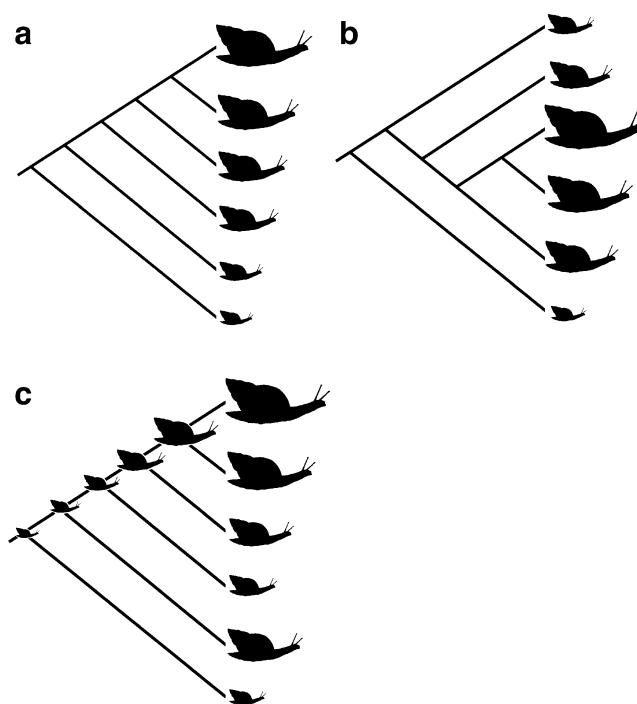


**Fig. 11** The order of terminal nodes is meaningless. One of the most common misconceptions about evolutionary trees is that the order of the terminal nodes provides information about their relatedness. Only branching order (i.e., the sequence of internal nodes) provides this information; because all internal nodes can be rotated without affecting the topology (Fig. 6), the order of the tips is meaningless. Nevertheless, there is a strong tendency for readers to take the tree in **a** as indicating that frogs are more closely related to fishes than humans are. They are not: both frogs and humans (and birds and lizards and cats) are equally closely related to fishes because as tetrapods they share a common ancestor to the exclusion of bony fishes. On the other hand, humans and cats are more closely related to each other than either is to any of the other species depicted because they share a recent common ancestor to the exclusion of the other species. The tree in **b** exhibits an identical topology to the one in **a** and is therefore equally valid. In this case, the same misinterpretation of “reading across the tips” would lead to the erroneous conclusion that birds are more closely related to fishes than cats are or that humans are more closely related to frogs than to lizards and birds. Because they share a common ancestor as amniotes, birds, cats, lizards, and humans are all equally related to frogs. It is good practice to rotate a few internal nodes mentally when first examining a tree to dispel misinterpretations based on reading the order of tips

with their shared ancestor represent a clade (amniotes) in which the first two clades are nested. Adding frogs and the ancestor linking them to the aforementioned species creates a yet larger clade (tetrapods). Adding fishes and the common ancestor of all species on this tree creates the final and largest clade (vertebrates). Because frogs can be

included in a clade with humans before fishes can—in other words, because frogs and humans share a common ancestor that is not shared with fishes—frogs are more closely related to humans than to fishes. Indeed, frogs and humans are exactly equally related to fishes through this common ancestor (recall that two cousins are equally related to a third, more distant relative).

A more rapid approach is to mentally rotate a few internal nodes with no effect on the topology of the tree, as shown in Fig. 11b. In this modified tree, humans are still sister to cats and birds are sister to lizards, frogs are then sister to amniotes, and fishes are the outgroup to the tetrapods. This second tree is identical in topology and is therefore equally accurate as the first tree. However, it



**Fig. 12** Evolutionary trends cannot be identified by reading across the tips. In addition to resulting in incorrect interpretations of relatedness (Fig. 11), reading across the tips can engender a false impression of evolutionary trends. For example, many readers confronted with the tree in **a** might be tempted to infer an evolutionary trend toward increased body size in snail species over time (or, in Fig. 11a, an increase in complexity or intelligence over time). Unfortunately, misinterpretations such as this can be found even in the primary scientific literature. Once again, this can be corrected simply by rotating a few internal nodes, as has been done in **b**, in which the topology is the same but where the supposed trend is no longer apparent. **c** shows evidence of a real evolutionary trend toward increased body size. The important consideration is internal branching: In this case, there is information about ancestral states (e.g., from fossils), and it is evident that in every branching event, the two descendant species have been larger than their shared ancestor. Despite this being a clear evolutionary trend, there is no pattern evident across the terminal nodes. Thus, reading across the tips can create apparent trends where there are none and can mask real trends that are strongly supported by historical information

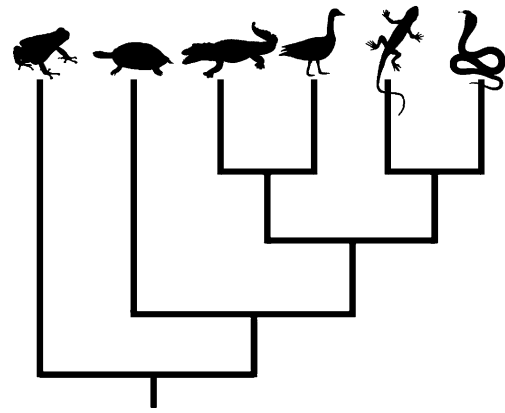
should be obvious that humans are not suddenly more closely related to frogs than to reptiles and birds.

Reading across the tips is not just problematic when interpreting relatedness. It can also lead readers and even authors of scientific publications to incorrectly intuit the existence of evolutionary trends where none exist or to overlook them where they do. For example, the phylogeny depicted in Fig. 12a may seem to show a trend toward increased body size in this snail clade. However, a simple rotation of a few internal nodes to produce an equivalent but nonladderized tree destroys this illusion (Fig. 12b). Conversely, although a reading across the tips in Fig. 12c would provide no indication of the fact, this tree indicates strong evidence for an evolutionary trend toward larger body size. In this case, information is available about the common ancestors, and it is clear that both descendants have been larger than their shared ancestor following every branching event. Only historical data or statistically rigorous inferences about history, and not a simple comparison of living species, can provide convincing support for claims of an evolutionary trend.

#### Misconception #4: Similarity versus Relatedness

The modern science of taxonomy is built upon the foundation laid by Carolus Linnaeus in the mid-eighteenth century. His system, which long predated the widespread scientific acceptance of common descent inspired by Darwin, categorized organisms on the basis of physical similarity. Notably, in the first edition of his *Systema Naturae* of 1735, whales were grouped with fishes—an oversight that he corrected in the tenth edition in 1758 by placing them with the other mammals. Today, the primary criterion for scientific classification is evolutionary relatedness, whereas differences in the degree of physical similarity across lineages are often a confounding variable. This can be so for two major reasons: First, as with whales and fishes, adaptation to similar environments can lead to a superficial convergence of physical appearance. Second, the rates of morphological change can vary considerably among lineages, with some remaining similar to a common ancestor and/or to more distantly related contemporary lineages and others becoming markedly different over the same time span (Baum et al. 2005).

By way of example, consider the phylogeny presented in Fig. 13. This tree shows one of the more prominent hypotheses regarding the relationships of major groups of nonmammalian tetrapods. Frogs are given as the outgroup in this tree, with turtles being the next most distantly related lineage to the others. Snakes are the sister group to lizards, and in fact, both modern lizards and snakes may be descended from a more ancestral lizard lineage. Most of these considerations are reasonably intuitive, but many

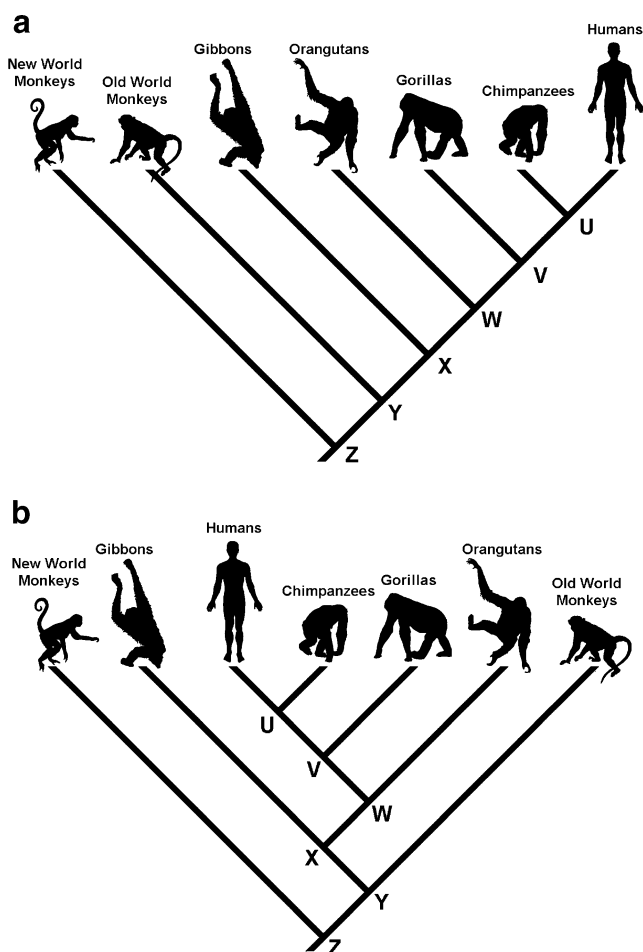


**Fig. 13** Evolutionary relatedness and physical similarity are not necessarily linked. The rates at which physical features change can differ among lineages (Fig. 8), and superficially similar morphologies can evolve independently in more than one lineage. As a result, close relatives may look different from one another or distant relatives may look misleadingly similar. This tree presents evolutionary relationships among “reptiles,” with frogs as the outgroup. Although they look very different, birds and crocodiles are actually more closely related to each other than either is to any other group of reptiles. This particular phylogenetic hypothesis shows birds, crocodiles, lizards, and snakes to all be equally related to turtles (a detail that remains a subject of debate) and birds and all “reptiles” to be equally related to frogs (which is well accepted). The similarities between birds and mammals (e.g., four-chambered hearts, homeothermy) evolved independently in the two lineages well after their split from a distant reptilian ancestor. As with “fishes,” the category of “reptiles” is phylogenetically inconsistent

people find it surprising that phylogenetically birds are located within the “reptiles” and represent the sister group to crocodilians. Although physical similarities would seem to suggest otherwise, crocodiles are more closely related to birds than they are to lizards. The reason for this is that the bird lineage has experienced significant modification, whereas crocodilians have remained largely unchanged for tens of millions of years. It is important to note that birds and crocodiles represent each other’s closest *living* relatives but that birds are not descended from crocodiles nor vice versa—as taxonomic groups, crocodilians and birds both arose long after their respective lineages diverged from a common ancestor well over 200 million years ago. Birds are, in fact, descended from a lineage of theropod dinosaurs, making *Tyrannosaurus rex* far more similar to the last nonavian ancestor of modern birds than anything resembling a crocodile (see Prothero 2007).

#### Misconception #5: Sibling versus Ancestor

Mistaken assumptions that the ancestor of two modern groups must have been very similar to, or perhaps even was, one of the modern groups extend well beyond the case of crocodiles and birds. Any claim that two species represent each other’s closest living relative should not be construed as implying that one of the modern groups itself



**Fig. 14** Cousins are not ancestors, and humans are not descended from chimpanzees. **a** shows an evolutionary tree of anthropoid primates as it is often depicted, namely as an unbalanced, right-ladderized tree with humans at the extreme end. Viewed in this way, several of the most common fallacies in interpreting trees can arise: for example, that humans are the endpoint of a “main line,” that there is a trend toward “human-ness” from left to right, that the human lineage includes a monkey ancestor, or that there has been no branching in the lineages leading to the other modern species of primates. All are absolutely false. This becomes clearer if a few internal nodes are rotated, as in **b**, which is an equally accurate depiction of primate relationships. Humans and chimpanzees are more closely related to each other than either is to gorillas, orangutans, or any other living primates. However, note that “chimpanzees,” although depicted as one terminal node here, includes both the common chimpanzee (*Pan troglodytes*) and the bonobo (*Pan paniscus*), and if this tree was drawn as recently as 30,000 years ago, it would also include Neanderthals (*Homo neanderthalensis*) as a sister species to humans. Humans are not descended from chimpanzees any more than chimpanzees are descended from humans; rather, the two share a common ancestor (*U*) that lived some 5–7 million years ago and that was neither a human nor a chimpanzee. “Monkeys” are divided into Old World and New World lineages. Old World monkeys share a more recent ancestor with apes (*Y*) than either does with New World monkeys (*Z*), which means that apes (including humans) and Old World monkeys are equally related to New World monkeys. Monkeys are not ancestral to humans: The two lineages are related as distant cousins, not as grandparents and grandchildren

is an ancestor of the other nor even that the common ancestor looked anything like either of the two groups. For example, the hypothesis that whales and hippopotamuses are sister groups (e.g., Boissarie et al. 2005) does not imply that the ancestor of whales was a hippo nor that it would even have been thought of as being similar to a hippo were it encountered when it was alive. Not surprisingly, the fossil record of whales, which is becoming increasingly extensive, shows that the early ancestors of whales (e.g., *Pakicetus*, *Ambulocetus*) bore no substantial resemblance to modern hippos at all (Thewissen and Bajpai 2001; Thewissen and Williams 2002).

Nowhere is this misconception more pronounced than in discussions of human evolution. One often hears it expressed in the rhetorical challenges offered by those who exhibit the poorest comprehension of evolutionary concepts: “If humans are descended from chimps,” so the question goes, “then why are there still chimps?” “If humans are descended from monkeys, then why has no one observed a monkey giving birth to a human baby?” The answer is simple because the premise is flawed: Humans are not descended from chimpanzees or monkeys, and no sane biologist suggests otherwise.

Figure 14a shows a ladderized phylogeny of the anthropoid primates. Humans and chimpanzees are sister taxa whose next (equally) close relatives are the gorillas, then the orangutans. Humans and chimpanzees share a common ancestor that lived around 5–7 million years ago. This ancestor was neither chimpanzee nor human, and as with whales, the increasingly detailed fossil record of the hominin lineage shows the extensive changes that have taken place since this divergence. Although the fossil record of chimpanzee ancestors is currently sparse, it can be presumed that a great deal of change characterized the evolutionary history of that branch as well.

The notion that other primates should have disappeared now that humans have evolved is based on a false understanding of species formation. Specifically, it assumes a process in which one species gradually transforms as a whole into another (called “anagenesis”). The reality of species diversification is that it most often proceeds by “cladogenesis,” the branching of new species from common ancestral populations. Chimps continue to exist because they are part of a separate branch that formed through cladogenesis when an ancestral population of a species, which was neither chimp nor human, split into independent lineages. Being confused about the coexistence of humans and chimpanzees is akin to being puzzled by the coexistence of Canada and Australia. Once again, rotating some internal nodes (Fig. 14b) can help to correct the misperception that other living primates are ancestors of humans or offshoots of a main line leading to humans or of

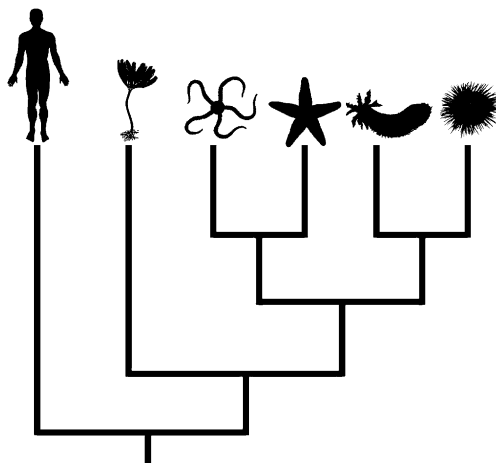


incorrectly assuming that the left- or bottom-most tip represents an ancestor to those at the terminal nodes of the other branches.

#### Misconception #6: Long Branch Implies no Change (or “Less Diverse Equals Basal Equals Ancestral”)

When viewing unbalanced trees such as those presented as Figs. 10a, 11a, 13, and 14a, there is a tendency among many people to misinterpret the long branch leading to the lone outgroup taxon in two ways. First, it is sometimes assumed that this species, although actually a contemporary of all others on the tree, is ancestral to the other lineages or at least is more similar to the root ancestor than any of the other species included in the tree (Crisp and Cook 2005). Second, this long branch is often taken to imply that no further branching has occurred along this lineage.

Figure 15 exposes the fallacy of both interpretations. In this case, humans are accurately included as the outgroup—the so-called basal lineage—to the echinoderms. It should go without saying that the branch leading from the common ancestor of echinoderms and vertebrates to modern mammals such as humans has not been devoid of additional divergence. In actuality, there have been hundreds of thousands, if not millions, of branching events along that lineage. The corollary of this observation, that humans do not resemble the ancestral echinoderm, should be even



**Fig. 15** A straight line does not mean that no change has occurred. This tree provides a simple illustration of the fact that the outgroup lineage cannot be assumed to be “basal,” “primitive,” or “ancestral” to the other species included on the tree. In this case, humans are accurately used as the outgroup to the echinoderms, which includes sea lillies, brittle stars, sea stars, sea cucumbers, and sea urchins. Of course, humans do not resemble the common ancestor of echinoderms, and there has been an enormous amount of branching among vertebrates since the very distant split of these two lineages from their common ancestor. It is most commonly argued that the sister group of the echinoderms is the hemichordates, both of which are more closely related to vertebrates than to any other group of animals, thus making the category “invertebrate” phylogenetically invalid

more obvious. Nonetheless, equally false interpretations of “basal” lineages are not uncommon, even in the scientific literature (Crisp and Cook 2005).

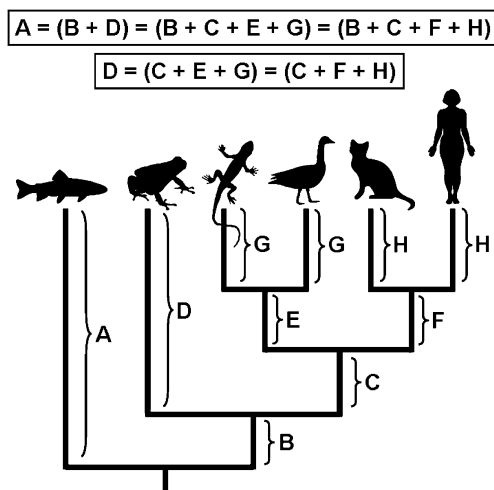
As with several of the other misconceptions discussed here, the problem of “basal equals primitive” is most likely to emerge when the tree under consideration is unbalanced and ladderized. It must be borne in mind that even if the unbalanced nature of a phylogeny reflects real differences in species diversity (which it often does not, as most trees include an incomplete sample of species), the relative diversity of major lineages can change over time, with one being the most diverse now and the other having been so in the past (Crisp and Cook 2005).

Finally, it must be pointed out that the relevant comparison is not between “primitive” (as in the sense of “poorly developed”) versus “advanced.” The only legitimate comparison is between “primitive” in the technical sense, meaning more like the last common ancestor (also called “ancestral”), versus “derived” (i.e., different from the ancestor). Any other interpretation runs the risk of invoking the fallacy of a progressive evolutionary scale. Moreover, as Crisp and Cook (2005) put it,

Once two lineages have separated, each evolves new characters independently of the other and, with time, each will show a mixture of plesiomorphic [inherited largely unchanged from the ancestor] and apomorphic [newly evolved and thus not possessed by the ancestor] character states. Therefore, extant species in both lineages resemble, to varying degrees, their common ancestor. Consequently, whereas character states can be relatively ancestral (plesiomorphic) or derived (apomorphic), these concepts are nonsensical when applied to whole organisms.

#### Misconception #7: Different Lineage Ages for Modern Species

Groups of species recognized as taxonomically distinct on the basis of particular characteristics (say, “flowering plants” or “beetles”) may have appeared at different times in the history of life and thus may be of different ages. However, the overall lineage leading to any modern species is of exactly the same age as that leading to any other modern species with whom an ancestor is shared (Fig. 16). This is a fundamental consequence of the principle of common descent, but there nevertheless can be a tendency to conflate taxon age with lineage age. For example, the group identified as teleost fishes is thought to be older—that is, to have appeared as a recognizable taxonomic group earlier—than mammals. Similarly, the first organisms that would be recognized as rainbow trout (*Oncorhynchus mykiss*) probably lived and died before the first individuals



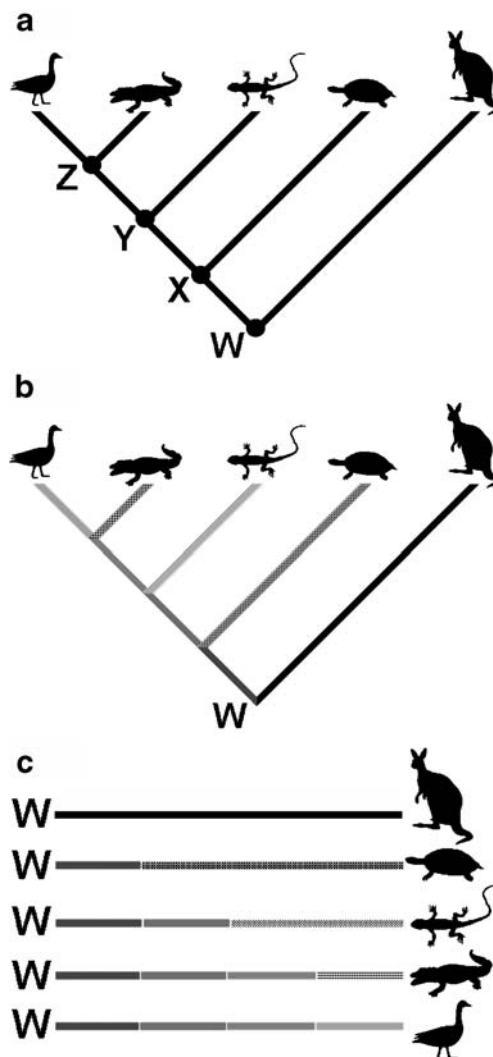
**Fig. 16** The lineages leading to contemporary species have all been evolving for exactly the same amount of time. Rates of morphological change may vary among lineages, but the amount of time that separates two living lineages from their common ancestor does not. This figure shows the relationships among a sample of vertebrate lineages, all of which have been evolving for exactly the same amount of time, even if some lineages have undergone more change or more branching than others or if some taxonomically identifiable subsets of those lineage (e.g., teleost fishes) arose earlier than others (e.g., mammals). It is therefore a fallacy to describe one modern species as “more evolved” than another. Note, however, that this is a cladogram rather than an ultrametric tree, such that one cannot assume that any or all of *G*, *H*, *E*, *F*, *C*, and *B* are equal, only that the total amount of time between root and tip is the same along each of the lineages

that would have been classified as *Homo sapiens* were born. However, rainbow trout and humans are contemporary species, meaning that the lineages of which they are currently terminal nodes have been evolving for exactly the same amount of time since their divergence from a distant common ancestor. As a result, any notion that one of these *lineages* is “more evolved” or that one has had more time to accumulate differences is flawed.

#### Misconception #8: Backwards Time Axes

Among the common misconceptions identified by Meir et al. (2007) was the tendency for many students to misread the time axis on evolutionary trees. Many students interpreted the location of the terminal nodes as indicating time, for example by reading from left to right or from the leftmost tip to the root. In Fig. 17a, for example, many students read time as proceeding from birds (oldest) to the root *W* (youngest) or from birds (oldest) to kangaroos (youngest). Neither is correct, as time extends from the root to the terminal nodes, all of which are contemporary. This misinterpretation may have been exacerbated by the fact that the tree used in the quiz placed mammals—which many students assume to be the most “advanced” and hence most recent group—alone on the less diverse branch at the far right of an unbalanced, ladderized tree (unfortunately, a

tendency to place humans or some other preferred taxon at the top or right of every tree appears to be an unshakable habit among many phylogeneticists, although there is no objective reason for doing so). As indicated in Fig. 5, even on cladograms, in which the lengths of the branches are not



**Fig. 17** The number of intervening nodes does not indicate overall relatedness between lineages. The tree in **a** is the same in topology as the one used in the study of Meir et al. (2007), which showed that many readers have a tendency misread the directionality of time on phylogenies and to count nodes when asked to determine evolutionary relatedness among species. Confusion may arise in this particular case because many people maintain the erroneous assumption that mammals are the most “advanced” and therefore must be the youngest group. More generally, because the tree is unbalanced, students may tend to consider birds and mammals (separated by four internal nodes on this tree, *Z*, *Y*, *X*, and *W*) as more distantly related than turtles and mammals (separated by two internal nodes, *X* and *W*). However, this is simply an artifact of the species chosen for inclusion on the tree. All species descended from ancestor *X* are equally related to kangaroos, with which they all share the same last common ancestor, *W*. To demonstrate this, **b** illustrates the same tree with different patterns for each branch, which are then spliced together in **c** to reveal the identical total distance from the common ancestor *W* to all of the terminal nodes

scaled to time, the historical order in which ancestors lived extends from the most distant past at the root, through a series of increasingly recent branching events, to the tips representing contemporary species.

#### Misconception #9: More Intervening Nodes Equals More Distantly Related

In the study by Meir et al. (2007), many students demonstrated a tendency to assess relatedness in a phylogeny like the one depicted in Fig. 17a by “counting nodes.” For example, because birds on this tree are separated from mammals by four internal nodes (Z, Y, X, W), whereas the separation of turtles and mammals consists of only two internal nodes (X, W), many students incorrectly concluded that birds must be more distantly related to mammals than are turtles. The important point in calculating relatedness is not the number of intervening nodes along a given branch but the number of shared ancestors.

In Fig. 17a, both turtles and birds share one ancestor with mammals (node W), making them equally closely related to mammals. By contrast, birds share three common ancestors with crocodilians (nodes Z, Y, and X) but only two with turtles (nodes X and W), which makes birds and crocodilians more closely related to one another than either is to turtles. To illustrate the basic notion that all modern species in a tree are equally distant from their common ancestor, one can plot the same phylogeny as in Fig. 17a with different patterns for each branch (Fig. 17b) and then splice those branches together to show that the total distance from the root (node W) to any of the terminal nodes is exactly equal (Fig. 17c). The only difference is the number of branching events that occurred within the lineages, whereas the relatedness of the lineages themselves is not affected by this. Misconceptions about relatedness based on node counting also could be countered by balancing the tree, for example by deleting all but one species of birds/reptiles, resulting in a symmetrical V-shaped tree, regardless of which species remains along with mammals, or by adding an equal number of mammals to the sample to even out the diversity along the major branches.

#### Misconception #10: Change Only at Nodes

There is a legitimate debate among professional evolutionary biologists regarding the patterns of species formation, such as whether it occurs comparatively rapidly (in a geological sense) or is more gradual. Proponents of the punctuated equilibrium model of speciation argue that species remain largely unchanged morphologically for the duration of their existence, with most physical diversification occurring concomitant with species formation events (Eldredge and Gould 1972; Gould 2002; Eldredge 2008). If

punctuated equilibrium were established conclusively to represent the exclusive mode of species formation in a clade and an accurate and complete phylogenetic tree were available for that clade that included all living and extinct species, then one could reasonably interpret the internal nodes as the points at which most morphological divergence took place among species. As Meir et al. (2007) noted, many students do draw such a conclusion, although of course this is not because they possess the requisite knowledge on which to base it.

The fact is that one should not assume that an internal node indicates the exact moment (again, geologically speaking) when particular physical changes came about, any more than one should interpret a long, node-free branch as indicating that no change has occurred. More accurately, an internal node represents the time at which a formerly cohesive population diverged into two genetically isolated descendant populations, with morphological change possible both at this time and long afterward (Baum et al. 2005).

Finally, one must bear in mind that terminal nodes can also be misinterpreted if the diversity that they sometimes represent is neglected. For example, the tree in Fig. 11 shows only a single fish, a frog, a lizard, a bird, a cat, and a human, but in actuality, these six terminal nodes together represent more than 50,000 species of living vertebrates and an untold number of ancestors. The important point is that any given node, whether internal or at the tips, represents a diverse assemblage of organisms with a complex evolutionary history.

#### Looking Ahead to Better Understanding the Past

Two points are abundantly clear when it comes to phylogenetic literacy: (1) It is crucial for an understanding of modern evolutionary concepts, and (2) it is insufficiently common. Misconceptions abound regarding evolutionary trees—sometimes because of, and sometimes creating, incorrect preconceptions about how, evolution operates. Many are holdovers of progressionist or even pre-evolutionary thinking about life’s diversity. Some, along with widespread misunderstandings of evolutionary mechanisms such as natural selection, undoubtedly contribute to the staggeringly low public acceptance of the principle of common descent in North America (Alters and Nelson 2002; Miller et al. 2006).

The way forward on this issue is unambiguous. Students, members of the public, and other nonspecialists must be better educated about the information that evolutionary trees do and do not convey. To this end, several teaching plans and software exercises for constructing and/or using phylogenetic hypotheses have become available (e.g., Bilardello and Valdes 1998; Gendron 2000; Singer et al.

2001; Goldsmith 2003; Meir et al. 2005). In addition, freely accessible online resources are making it possible for individuals to learn about and interact with evolutionary trees (see Appendix).

More generally, lessons at the high school and undergraduate level should de-emphasize the technical aspects of phylogeny reconstruction in favor of a focus on the concepts underlying tree thinking. In this regard, identifying, confronting, and clarifying misconceptions is perhaps the most important strategy. After all, a misconception corrected is a concept better understood. In few cases is this more relevant or more important than with Darwin's preferred metaphor of the Tree of Life.

**Acknowledgments** I thank Sarah Adamowicz, Alex Ardila Garcia, Martin Brummell, Niles Eldredge, Bruce Lieberman, Mark Pagel, Andy Purvis, Jillian Smith, Phillip Spinks, and Jonathan Witt for feedback on an early draft of the paper.

## Appendix. Online resources

### Understanding evolution

Understanding phylogenies:

[http://evolution.berkeley.edu/evolibrary/article/evo\\_05](http://evolution.berkeley.edu/evolibrary/article/evo_05)

Phylogenetic systematics, a.k.a. evolutionary trees:

[http://evolution.berkeley.edu/evolibrary/article/0\\_0\\_0/phylogenetics\\_01](http://evolution.berkeley.edu/evolibrary/article/0_0_0/phylogenetics_01)

### Tree-Thinking Group

<http://www.tree-thinking.org>

### The Tree-Thinking Challenge Supplemental Quiz

[www.sciencemag.org/cgi/content/full/310/5750/979/DC1](http://www.sciencemag.org/cgi/content/full/310/5750/979/DC1)

### Tree of Life Web Project

<http://tolweb.org/tree/phylogeny.html>

### Interactive Tree of Life (iTOL)

<http://itol.embl.de/index.shtml>

### TreeBASE

<http://www.treebase.org/treebase/index.html>

### Encyclopedia of Life

<http://www.eol.org/>

### Caminalcules teaching activity

<http://nsm1.nsm.iup.edu/rgendron/Caminalcules.shtml>

## References

- Alters BJ, Nelson CE. Teaching evolution in higher education. *Evolution* 2002;56:1891–901.
- Baldauf SL. The deep roots of eukaryotes. *Science* 2003;300:1703–6.
- Baum DA, DeWitt Smith S, Donovan SSS. The tree-thinking challenge. *Science* 2005;310:979–80.
- Benton MJ. *Vertebrate palaeontology*. 3rd ed. Malden, MA: Blackwell; 2005.
- Benton MJ, Ayala FJ. Dating the tree of life. *Science* 2003;300:1698–700.
- Bilardello N, Valdes L. Constructing phylogenies. *Am Biol Teach* 1998;60:369–73.
- Bininda-Emonds ORP, Cardillo M, Jones KE, MacPhee RDE, Beck RMD, Grenyer R, et al. The delayed rise of present-day mammals. *Nature* 2007;446:507–12.
- Boisserie J-R, Lihoreau F, Brunet M. The position of Hippopotamidae within Cetartiodactyla. *Proc Natl Acad Sci USA* 2005;102:1537–41.
- Cracraft J, Donoghue MJ, editors. *Assembling the tree of life*. Oxford, UK: Oxford University Press; 2004.
- Crisp MD, Cook LG. Do early branching lineages signify ancestral traits? *Trends Ecol Evol* 2005;20:122–8.
- Darwin C. *Notebook B*. 1837; p. 36, 74.
- Darwin C. *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life*. London, UK: Murray; 1859.
- Dawkins R. *Progress*. In: Keller EF, Lloyd EA, editors. *Keywords in evolutionary biology*. Cambridge, MA: Harvard University Press; 1992. p. 263–72.
- Doolittle WF. Uprooting the tree of life. *Sci Am* 2000;282:90–5. (Feb.).
- Doolittle WF, Bapteste E. Pattern pluralism and the tree of life hypothesis. *Proc Natl Acad Sci USA* 2007;104:2043–9.
- Eldredge N, Gould SJ. Punctuated equilibria: an alternative to phyletic gradualism. In: Schopf TJM, editor. *Models in paleobiology*. San Francisco, CA: Freeman Cooper; 1972. p. 82–115.
- Eldredge N. The early “evolution” of “punctuated equilibria”. *Evo Edu Outreach* 2008;2. DOI 10.1007/s12052-008-0032-0.
- Felsenstein J. *Inferring phylogenies*. Sunderland, MA: Sinauer; 2003.
- Gendron RP. The classification and evolution of Caminalcules. *Am Biol Teach* 2000;62:570–6.
- Goldsmith DW. The great clade race: Presenting cladistic thinking to biology majors and general science students. *Am Biol Teach* 2003;65:679–82.
- Gould SJ. *Bully for Brontosaurus*. New York: W.W. Norton; 1991. p. 168–81.
- Gould SJ. The evolution of life on the earth. *Sci Am* 1994;271:85–91. (Oct.).
- Gould SJ. *Full house*. New York: Harmony; 1996.
- Gould SJ. Branching through a wormhole. *Nat Hist* 1999;108:24–7, 84–9. (March).
- Gould SJ. *The structure of evolutionary theory*. Cambridge, MA: Harvard University Press; 2002.
- Gregory TR. Evolution as fact, theory, and path. *Evo Edu Outreach* 2008;1:46–52.
- Hall BG. *Phylogenetic trees made easy*. 3rd ed. Sunderland, MA: Sinauer; 2007.
- Harcourt-Brown KG, Pearson PN, Wilkinson M. The imbalance of paleontological trees. *Paleobiology* 2001;27:188–204.



- Hillis DM, Bull JJ, White ME, Badgett MR, Molineux IJ. Experimental phylogenetics: generation of a known phylogeny. *Science* 1992;255:589–92.
- Hillis DM, Moritz C, Mable BK, editors. *Molecular systematics*. 2nd ed. Sunderland, MA: Sinauer; 1996.
- Hodkinson TR, Parnell JAN, editors. *Reconstructing the tree of life: taxonomy and systematics of species rich taxa*. Boca Raton, FL: CRC; 2007.
- Huxley TH. On the application of the laws of evolution to the arrangement of the Vertebrata and more particularly of the Mammalia. *Proc Zool Soc Lond* 1880;43:659–62.
- Lamarck JB. *Philosophie zoologique*. Paris; 1809.
- Lecointre G, Le Guyader H. *The tree of life: a phylogenetic classification*. Cambridge, MA: Belknap; 2007.
- Mace GM, Gittleman JL, Purvis A. Preserving the tree of life. *Science* 2003;300:1707–9.
- Mace R, Holden CJ. A phylogenetic approach to cultural evolution. *Trends Ecol Evol* 2005;20:116–21.
- MacFadden BJ. *Fossil horses: systematics, paleobiology, and evolution of the family Equidae*. Cambridge, UK: Cambridge University Press; 1992.
- MacFadden BJ. Fossil horses—evidence for evolution. *Science* 2005;307:1728–9.
- Maddison DR, Schulz K-S, editors. *The tree of life web project*. <http://tolweb.org> (2007).
- Mayr E, Provine WB, editors. *The evolutionary synthesis*. Cambridge, MA: Harvard University Press; 1980.
- Meir E, Herron JC, Maruca S, Stal D, Kingsolver J. *EvoBeaker 1.0*. Ithaca, NY: SimBiotic Software. <http://www.simbio.com> (2005).
- Meir E, Perry J, Herron JC, Kingsolver J. College students' misconceptions about evolutionary trees. *Am Biol Teach* 2007; 69:71–6.
- Miller JD, Scott EC, Okamoto S. Public acceptance of evolution. *Science* 2006;313:765–6.
- Nee S. The great chain of being. *Nature* 2005;435:429.
- Nei M, Kumar S. *Molecular phylogenetics and evolution*. Oxford, UK: Oxford University Press; 2000.
- O'Hara RJ. Homage to Clio, or, toward an historical philosophy for evolutionary biology. *Syst Zool* 1988;37:142–55.
- O'Hara RJ. Telling the tree: narrative representation and the study of evolutionary history. *Biol Philos* 1992;7:135–60.
- O'Hara RJ. Evolutionary history and the species problem. *Am Zool* 1994;34:12–22.
- O'Hara RJ. Population thinking and tree thinking in systematics. *Zool Scr* 1997;26:323–9.
- Page RDM, Holmes EC. *Molecular evolution: A phylogenetic approach*. Malden, MA: Blackwell; 1998.
- Prothero DR. *Evolution: What the fossils say and why it matters*. New York: Columbia University Press; 2007.
- Rambaut A, Robertson DL, Pybus OG, Peeters M, Holmes EC. Phylogeny and the origin of HIV-1. *Nature* 2001;410:1047–8.
- Rokas A. Genomics and the tree of life. *Science* 2006;313:1897–8.
- Ruse M. *Monad to man*. Cambridge, MA: Harvard University Press; 1996.
- Salemi M, Vandamme A-M, editors. *The phylogenetic handbook*. Cambridge, UK: Cambridge University Press; 2003.
- Sanson GFO, Kawashita SY, Brunstein A, Briones MRS. Experimental phylogeny of neutrally evolving DNA sequences generated as a bifurcate series of nested polymerase chain reactions. *Mol Biol Evol* 2002;19:170–8.
- Singer F, Hagen JB, Sheehy RR. The comparative method, hypothesis testing and phylogenetic analysis. *Am Biol Teach* 2001;63: 518–23.
- Thewissen JGM, Bajpai S. Whale origins as a poster child for macroevolution. *BioScience* 2001;51:1037–49.
- Thewissen JGM, Williams EM. The early radiations of Cetacea (Mammalia): evolutionary pattern and developmental correlations. *Ann Rev Ecol Syst* 2002;33:73–90.
- Vogel G. Phylogenetic analysis: getting its day in court. *Science* 1997;275:1559–60.
- Wolf YI, Rogozin IB, Grishin NV, Koonin EV. Genome trees and the tree of life. *Trends Genet* 2002;18:472–9.