Tree thinking for all biology: the problem with reading phylogenies as ladders of progress

Kevin E. Omland,[1,2]* Lyn G. Cook,[3] and Michael D. Crisp[2]

Summary

Phylogenies are increasingly prominent across all of biology, especially as DNA sequencing makes more and more trees available. However, their utility is compromised by widespread misconceptions about what phylogenies can tell us, and improved ''tree thinking'' is crucial. The most-serious problem comes from reading trees as ladders from ''left to right''—many biologists assume that species-poor lineages that appear ''early branching'' or ''basal'' are ancestral—we call this the ''primitive lineage fallacy''. This mistake causes misleading inferences about changes in individual characteristics and leads to misrepresentation of the evolutionary process. The problem can be rectified by considering that modern phylogenies of present-day species and genes show relationships among *evolutionary cousins*. Emphasizing that these are extant entities in the 21st century will help correct inferences about ancestral characteristics, and will enable us to leave behind 19th century notions about the ladder of progress driving evolution. ***BioEssays*** **30:854–867, 2008.**

Introduction—the importance of "tree thinking"

During the last twenty years, phylogenies—evolutionary trees—have become increasingly important in nearly all subdisciplines in biology. The greater use of phylogenies can be traced to advances on several fronts: (1) ease and affordability of DNA sequencing, (2) advances in the bioinformatics programs used to analyse phylogenetic data, and (3) conceptual advances in how to think about and utilize phylogenies—tree thinking.[1–7] One of the main uses of phylogenies is to reconstruct how lineages and characteristics have evolved over time. Being able to infer the characteristics of extinct ancestral species or genes is a powerful advance made possible by modern phylogenies. For example, researchers studying protein evolution have been able to reconstruct ancestral amino acid sequences, and then test the activity of these reconstructed proteins.[8–10] More recently, researchers have reconstructed gene expression patterns[11,12] and ancestral vertebrate genomes.[13] On the organismal side, animal behaviorists have reconstructed features of inferred ancestral frog calls and played these calls back to test responses of living frogs.[14] Being able to make inferences about characteristics such as ancestral DNA sequences, gene order, behaviors and coloration, none of which generally fossilize, is an exciting research program that was generally not possible thirty years ago (see Refs 15,16).

However, many researchers bring misconceptions to this exercise that do not fit with current understanding of trees or the process of evolution. Our conception of evolution and our interpretation of phylogenetic trees are intimately linked—each affects the other. How we interpret phylogenetic trees directly impacts our understanding of evolution. The classic example of the intertwined misunderstanding of tree thinking and evolution is the so-called "ladder of progress". Also known as the *scala naturae*, this concept was central to early attempts to understanding the organization of life beginning with Aristotle, through Linnaeus and the ''chain of being''. The ladder of progress view is most clearly documented in Haeckel's phylogenetic trees[17] (Fig. 1) (see Refs 18–23). Haeckel's trees show extant groups such as ''amoeba'' and ''primitive worms'' low down and embedded within the main trunk of the tree. Further up, evolution passes through ''jawless animals'', ''pouched animals'', ''apes'' and finally, to the apparent pinnacle of evolution—''man''.

Now, with the benefit of nearly 150 years of evolutionary research, and the work of biologists such as Darwin,[24] Hennig,[1] Wilson,[25] Gould,[26] Felsenstein[27] and Avise[28] to name just a few, we know that evolution generally has not stopped in any lineages. Thus, it is misleading to think of an extant amoeba species as ancestral to humans, or an extant amphibian as ancestral to snakes. Evolution has continued in amoebas and amphibians. In particular, new mutations can

[1]Department of Biological Sciences, University of Maryland, Baltimore County, Baltimore MD 21250.
[2]School of Botany and Zoology, The Australian National University, Canberra, ACT 0200, Australia.
[3]School of Integrative Biology, The University of Queensland, Brisbane, QLD 4072, Australia
Funding agencies: Kevin Omland is funded by an US National Science Foundation, Systematics Program CAREER Grant (DEB–0347083). Lyn Cook and Mike Crisp were partially funded by Australian Research Council Discovery Project grants.
*Correspondence to: Kevin E. Omland, Department of Biological Sciences, University of Maryland, Baltimore County, Baltimore MD 21250. E-mail: omland@umbc.edu
DOI 10.1002/bies.20794
Published online in Wiley InterScience (www.interscience.wiley.com).

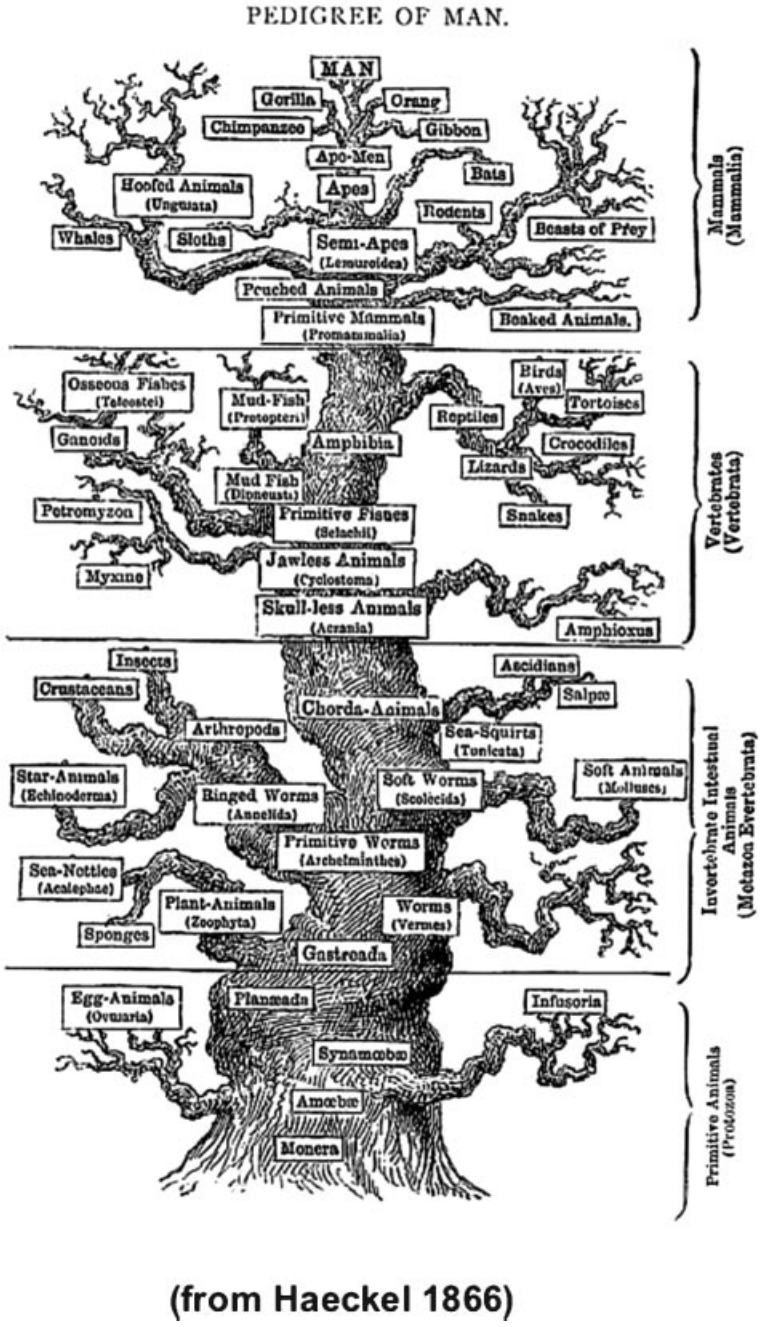


Figure 1. An early conception of a phylogenetic tree from the work of Haeckel.[17] This view of evolutionary history is based on a ladder of progress, *scala naturae*, view of life. Note in particular that the central feature of this tree is a trunk. Such a "main branch" is completely absent from current drawings or understanding of evolutionary trees. However, reading trees from "left to right", or from one extant species to the next to the next, perpetuates this outdated view of evolution.

lead to a better fit between any organism and its environment, especially in the face of changing environmental factors such as new pathogens. Furthermore, we know that evolution by genetic drift proceeds due to neutral processes for morphological and molecular traits.[29] Classic work by Wilson and colleagues[25,30] (also see Refs 31–33) demonstrated that, even though some gross morphological features may evolve slowly in some lineages (e.g. frogs, horseshoe crabs), evolution of other parts of the organism, especially molecular sequences, generally continue at similar rates. So even species that some might *consider* primitive, simple or ancestral, continue to evolve at some rate at least for some characteristics throughout evolutionary history. However, some researchers continue to incorrectly describe certain present-day species as "primitive" and to incorrectly imply that extant species may be ancestral to other extant species.

Improved tree thinking will not only help us better understand the evolution of the particular characters that we are studying, but will also improve our fundamental understanding of the process of evolution.

Understanding the problem—looking for ancestors in the treetops

Consider the tree in Fig. 2A with four hypothetical extant species, A, B, C and D. As with all figures in this paper, this is a rooted cladogram showing unscaled branches. Which of the species is the oldest? Which is youngest? Which is most ancestral? Most derived? Most primitive? Most advanced? Most simple? Most complex? The answer is that a phylogeny provides *no information* about any of these questions! While this answer may seem inconvenient to researchers looking to phylogenies to provide that information, these are the incorrect questions to be asking. Questions that could be asked include the following. Which trait evolved first? Which trait evolved most recently? What might the common ancestor have looked like? What was the ancestral protein sequence? Such questions involve *trait* evolution, or the *characteristics* of extinct ancestral organisms.

The pervasive misconceptions surrounding tree thinking can be demonstrated by presenting a correct tree drawn in a way that may seem incorrect to some readers (Fig. 2B). Looking at the tree that shows humans at the far left, ask which species seems the most-old, ancestral, primitive, or simple. We are used to seeing humans in the far right of trees (see Ref. 34), so showing *Homo sapiens* in the far left, with *E. coli* at the far right seems wrong to many readers (also see Refs 19,35). However, there is nothing wrong with this tree—it represents our best understanding of relationships among these four species (references Fig. 2B legend). First, this tree shows that *Escherichia coli* and *Salmonella enterica* are each other's closest relatives among these four taxa—that is, they share a more-recent common ancestor with each other than with *Rickettsia rickettsii* or *H. sapiens*. Second, it shows that *Rickettsia* shares a more-recent common ancestor with the other two proteobacteria than with *H. sapiens*.

Going back to Fig. 2A, some might think that one can read the trees "left to right", so that species A is considered the oldest, most ancestral and most primitive, followed in turn by species B, C and D. However, it is a fundamental mistake to assume that order on the page has some meaning, in effect to read the trees as if time moves forward from "left to right".[22,36]

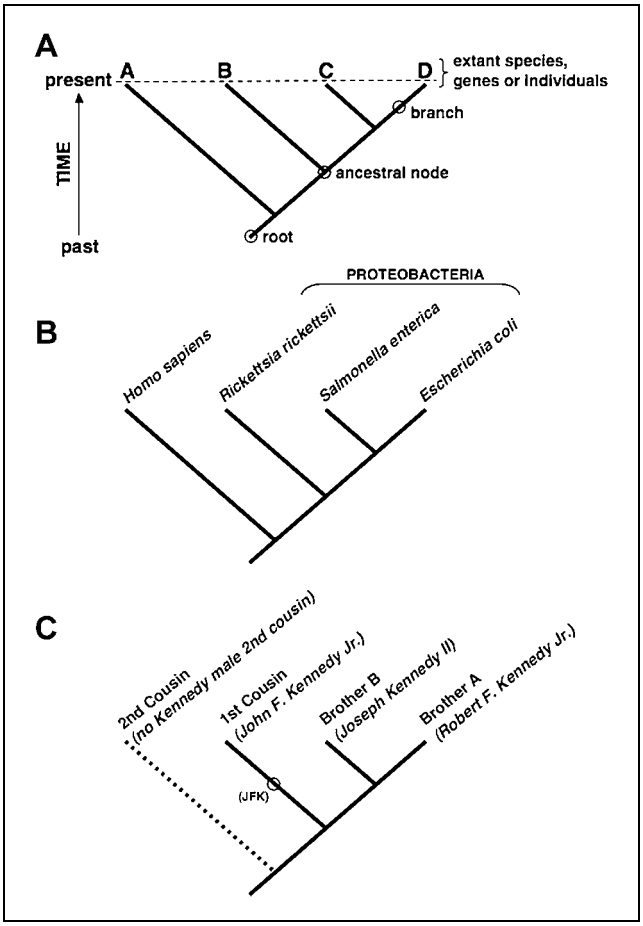


Figure 2. Phylogenetic tree interpretation. **A:** Phylogenetic tree illustrating relationships among four hypothetical species. Which of the four species is oldest? Which appears most ancestral? Which seems most primitive? As with all figures in this paper, this tree shows extant entities present today, and branch lengths are not drawn to scale—a cladogram. Most phylogenies published today, especially molecular phylogenies, have extant entities as their "terminals", and the dashed line emphasizes that all are present today. Note that while the Y axis in this orientation denotes relative time, the X axis has no meaning (see Ref. 4). **B:** A phylogenetic tree showing relationships among three Eubacteria and a representative of the Eukaryota. We are not used to seeing *H. sapiens* in the "far left" position on such a tree, and this position may make humans appear to some like the "older" or more "primitive" species. (For bacterial relationships see [34,73,74]; but see [75].) **C:** An evolutionary tree showing relationships among members of one generation of a human family, in this case the Kennedys. That John F. Kennedy Jr. appears to the left should not imply to us that he was older, more primitive, or more ancestral than his cousin Robert F. Kennedy Jr. (in fact he was 6 years younger). All phylogenies of extant species or genes show "evolutionary cousin" relationships and should not imply one species is more primitive, whereas another is more advanced. President John F. Kennedy does not really appear on this tree except along a branch. For that reason JFK Jr. may for some appear more closely related to their common ancestor, but all the cousins are equally related to that ancestor - their grandfather. Similarly, all extant evolutionary cousins in species trees are equally related to shared common ancestors. www.pbs.org/wgbh/amex/kennedys/

All of the taxa shown at the top (at the tips of the branches) are extant species present in the 21st century—i.e. the time axis runs from the bottom to the top of the page, going from the ancestral *nodes* in the past to the extant taxa or genes present today. We should not be looking for ancestors in the tops of the trees among the extant species.

Phylogenies show relationships among evolutionary cousins

Another way to view this problem is by considering a phylogeny of genes from the same generation of humans. Most phylogenies published today, especially molecular phylogenies, focus on relationships among extant species, genes and individuals. So consider a phylogeny of a human family just focusing on one recent generation. Fig. 2C shows the example of the American political family, the Kennedys. This tree focuses on the sons of Robert F. Kennedy (Robert F. Kennedy Jr. and Joseph Kennedy II), and also shows one first cousin (John F. Kennedy Jr.) and a hypothetical 2nd cousin. For human males, one should be able to draw such a male lineage tree with enough data from the Y chromosome, and for females one can consider the example of a phylogeny based on mitochondrial DNA. The first cousin, JFK Jr, is shown to the left, with RFK Jr to the far right. *Does that mean that JFK Jr is any older, more ancestral or more primitive than his cousin RFK Jr? We know that the answer is no*. Phylogenies of extant species should be interpreted in the same way—names at tips are *evolutionary cousins*, not ancestors. The ancestors are represented by internal nodes and internal branches on the tree—for example President John F. Kennedy (JFK—see labelled circle) is represented by the middle section of the branch leading to his son JFK Junior. (Although in the case of such human trees, ancestors can still be living.)

Reading trees "left to right" assumes a ladder of progress

A major cause of tree reading problems is thinking about evolution "left to right". Evolution is often presented in this way, for example, in introductory biology textbooks.[37,38] The evolutionary history of plants is presented as beginning with algae, proceeding along an orderly path to mosses, ferns, gymnosperms then angiosperms, the high point of plant evolution—i.e. from "lower plants" to "higher plants" (see Ref. 18). Animal evolution is presented as "beginning" with sponges, then "proceeding" to protostomes, to deuterostomes, to fish, early tetrapods, then humans. The problem occurs when extant species are used to illustrate this sequence, especially when they are presented as being "primitive" or "ancestral". Even in our own teaching, we do not emphasize enough that extant species are used to show the order in which key *characteristics* evolved (examples of ancestral versus derived character states). In fact, there is no general linear sequence of ancestral to derived, to more derived, except along the *internal* branches of the tree, following lineages through evolutionary time. Moreover, at each node in a bifurcating tree, there are two possible pathways that then can continue to branch through the tree, so cumulatively there are many different pathways from ancestors to different extant species. Reading the tree from "left to right" across the tip species gives one the false impression that there is just one pathway of evolutionary change.[7]

Unfortunately, in thinking of an ordered sequence of taxa (e.g. moss, fern, gingko, oak), some slip into saying or thinking that mosses are older than oaks and therefore that extant mosses are ancestral to oaks. However, the moss and oak lineages diverged approximately 400 million years ago.[39] Thus the moss lineage leading to *Sphagnum cuspidatum* is sister to the vascular plant lineage leading to *Quercus rubra*. Both lineages have been evolving independently for 400 million years into the plant species extant today (see Box 1 glossary, "crown group age" versus "stem group age"). The common ancestor to both these plant species is extinct. Perhaps because mosses are a member of a species-poor lineage, some think of their evolution as stopping once they diverged from what seems like the "main" group of extant plants.[7] This may be reinforced because some key morphological features of extant moss appear very similar to those in the fossil record.[40]

All extant species are a mix of ancestral and derived characteristics

The duck-billed platypus (*Ornithorhynchus anatinus*) provides a striking example[41] because it is often considered to be primitive. (Google Scholar finds over 900 hits for "platypus" and "primitive".) This perception is likely because of one set of characters relating to reproduction, specifically that the platypus lays eggs. However, this retained ancestral trait is only one aspect of its biology. The platypus has many other characteristics that are derived.[42] For example, the electroreceptor equipped bill is highly derived, as is the webbing on its feet.[43] In these latter characters, humans and most other placentals generally retain ancestral *states* that are expressed by many other tetrapods. The focus on characteristics tells us which traits are likely to be ancestral and which are derived—it is not the extant organism itself that is ancestral or "primitive".[44]

Clearly species vary in the number of ancestral versus derived character states. Indeed, it is possible to compare species to study rates of evolution using careful quantitative studies of extant and/or fossil taxa, but this character change information cannot come from molecular phylogenies of extant species alone, especially from cladograms with unscaled branch lengths. Several studies have addressed rates of molecular and/or morphological evolution in species-rich versus species-poor lineages[45–48] (see also Refs 49,50). There is also a series of studies that show that so-called "living fossil" taxa or lineages

Box 1. Glossary

Ancestral state reconstruction: using a phylogeny to infer the character states at ancestral nodes based on the character states of extant species or genes. (When working with species trees and organismal characters, the phylogeny is usually developed first based on an independent data set.) Also known by other terms including ''character-state optimization'', or ''phylogenetic character mapping''. (See definitions for character and character state.)

Ancestral haplotype: for closely related populations or species, it is common to find several individuals sharing a common haplotype or DNA sequence (in a central position in the network) that is likely to be ancestral to several more rare sequences.[79] Thus, when considering samples at the lowest levels of evolutionary divergence (i.e. within and among closely related species), extant entities can be ancestral to other extant entities (also see Refs 56,80,81).

Basal node: the oldest node on a given tree, at the ''base'' of the tree (see Fig. 3); this node can also be known as the ''ancestral node'' and depicts the most-recent common ancestor. Character states on both sides of this basal node should be considered when determining ancestral character states. (Other unambiguous uses of the term basal include ''basal divergence'', ''basal split'' etc. all of which refer to the two sides of the tree on both sides of the basal node.)

Basal lineage (or basal taxon): a problematic term referring to lineages with fewer terminals.[41,58] This use of the term basal is problematic because it can be confused with the correct use of the term (see basal node above).

Cladogram: a branching diagram (tree) that represents phylogenetic history in which branch lengths are unscaled (cf. phylogram).

Character: any heritable trait of an organism (e.g. specific site in a DNA or amino acid sequence, or eye color—cf. character state).

Character states: subdivisions of the variation within a character (e.g. presence of A, T, C or G in a DNA sequence, or red versus yellow eyes).

Clade: see monophyletic group.

Crown group: the monophyletic group that includes just the *extant* members of a lineage.[82] (Each such clade also has a *stem*, which comprises members of the lineage that evolved after it diverged from its sister group, but which went extinct before the coalescence time of the crown group—see below.)

Crown group age: the age of the most-recent common ancestor of the crown group. (Crown group age is always younger than the *stem group age*, which equals the time of divergence of the lineage from its sister group—see above Ref. 82).

Evolutionary tree: used here as synonymous with phylogeny. (Within cladistics, some authors have used a more-strict definition—for example Ref. 6 pp. 22–24.)

Gene tree: a phylogeny representing relationships among different copies of a gene. When all copies are sampled from a single species, all bifurcations represent gene duplication events (paralogous genes). When only single copies from an unduplicated gene are sampled from two or more species (orthologous genes), then all bifurcations represent speciation events. When multiple copies are included, both from within and among two or more species, then the gene trees include both paralogous and orthologous genes (e.g. Fig. 5A) (also see Refs 6,83).

Ladderized tree: a phylogeny in which every node is rotated to show the species-poor sister group on one side, and the species-rich sister group on the other side. Such views of trees may contribute to continued searching for ''ladders of progress'' in present day phylogenies.

Ladderize right: a tree drawing rotation setting available in programs such as PAUP*[84] and Treeview[85] (ladderize left is the opposite rotation). Trees shown in this paper generally use the ''ladderize right'' format, with the species-poor groups at left and the species-rich groups at right. This general format is common, from summaries of the tree of life[34] to introductory biology texts.[38]

Monophyletic group (clade): a group that includes the most-recent common ancestor and all of its descendants. (Generally recognised by a shared derived character state in all of its members—e.g. milk production in mammals. Note that when two groups are each other's sisters, and both are monophyletic, they are ''reciprocally monophyletic''.).

Paraphyletic group: a group that fails to include all the descendants of the most-recent common ancestor (e.g. great apes with humans left out). More generally, such groups are non-monophyletic, which also includes the related term ''polyphyletic'' (for diagrams see Ref. 6).

Phylogram: a branching diagram that represents evolutionary history in which branch lengths are drawn to scale, usually to amount of genetic change for molecular phylogenies (chronograms are scaled to absolute time).

Species tree: a phylogeny meant to represent relationships among species (or higher taxa); bifurcations in the tree represent speciation events (cf. gene tree).

Terminal: also known as an OTU or ''operational taxonomic unit'', this is a general term that can refer to phyla, species, individuals or genes. On molecular phylogenies, the terminals are generally the extant individuals or genes used to build the phylogeny. (The terminals are also used as the basis for inferring ancestral character states.)

Vertical tree: all trees in this paper are vertical trees, with the root node at the bottom of the page. However, trees can also be drawn in a horizontal orientation (with root node on one side, usually the left; another format is a ''circle tree'' with the root node in the center). Trees can be presented in many styles (e.g., slanted versus rectangular branches)—similar ''tree thinking'' applies to all these rooted trees. (Unrooted trees or ''networks'' without a time axis cannot be used to determine which characteristics are ancestral.)

Zig-zag rotation: a phylogenetic node rotation style in which the most-species-poor groups are alternated left, then right and so on (see Ref. 41; cf. ladderized tree)

For definitions of these and other phylogenetic terms see: http://www.sasb.org.au/glossary/html

with apparently slow rates of morphological evolution continue to accumulate many molecular changes.[25,30,32] Slow rates of speciation and/or slow rates of evolution of one type of character do not mean that evolutionary change in the lineage as a whole slows down dramatically or stops.[31]

All extant genes are a mix of ancestral and derived characteristics

Another powerful use of phylogenetic trees is to show evolutionary relationships among genes in multigene families, for example showing all the relationships among gap junction proteins in zebrafish[51] or MADS-box genes in *Arabidopsis*.[52] In such trees, many of the bifurcations are caused by gene duplications and such phylogenies are generally called "gene trees". As with species, genes contain a mixture of ancestral and derived characteristics, including single nucleotide substitutions, presence or absence of deletions, inversions, introns and other genomic features (both within and between transcribed regions). Again, as with species, just because some aspect of a gene (e.g. functional domain amino acid sequence) may be highly conserved, that does not mean that other aspects of the gene such as nucleotide sequence, intron length, expression patterns, gene order with respect to nearby parts of the genome, etc. either have slower evolutionary rates or stop evolving.

Solutions—thinking about characteristics

Whether using species trees or gene trees, researchers need to focus on characteristics, whether those be the presence or absence of limbs in vertebrates, or a serine or proline at a given site in a protein. Specific algorithms are available to help determine the likely character states of ancestral nodes on a phylogeny (for examples and detailed explanations see Refs 4,15,53). Here, we will use the method of parsimony because it is the simplest way to illustrate how the most-serious problem with tree thinking occurs.

Consider the simple situation of a tree including four chordate lineages with strong support for their relationships (Fig. 3). Using the example of an ancestral steroid receptor[54] reconstructing the common ancestor with serine at this site is most parsimonious (Fig. 3A). Reconstructing proline at this site would be less parsimonious because it would require two separate changes. Several papers discuss some of the assumptions inherent in such methods.[15,55–57]

Now consider a hypothetical case where only one lineage has Ser (Fig. 3B). Mis-reading the tree "left to right" following a ladder of progress, one might think that this tree again suggests that the common ancestor had serine with one change to proline. However, based on parsimony, the ancestral state is equivocal. It is clear that there must have been at least one change: either a change from Ser to Pro, or a change from Pro to Ser. But these two alternatives are equally parsimonious. (Box 2 compares parsimony, maximum likelihood and Bayesian methods of ancestral character mapping.)

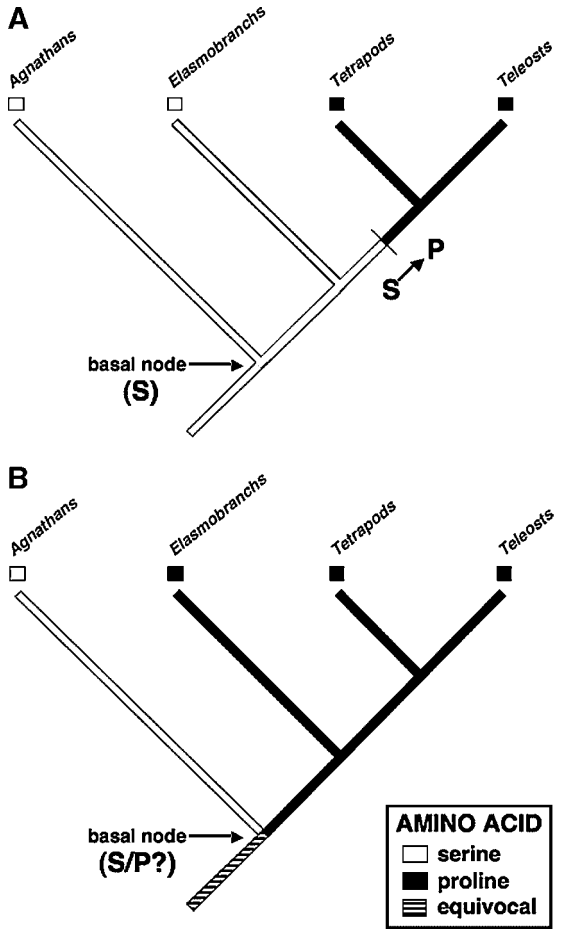


Figure 3. Ancestral states and the "primitive lineage fallacy". **A:** Ancestral state reconstruction of hormone receptor proteins for four chordate lineages based on Bridgham et al. (Fig. 4B residue 106[54]). Amino acid residues of extant genes are shown in boxes below the taxon names. Ser is the most-probable ancestral state based on parsimony, because that reconstruction only requires one change to Pro in the teleost—tetrapod ancestor. **B:** Demonstration of the "primitive lineage fallacy" in a hypothetical scenario when only one of the genes has Ser. Again, this reconstruction shows the most-probable ancestral states based on parsimony. In this case, it may seem to some that agnathans are in the "primitive" position thus must possess the ancestral state. Because some might think agnathans *appear* "basal" or "early branching" it may seem like the ancestor would have had a Ser. However, the reconstruction of sequence of the ancestral gene at the basal node is equivocal. Either Ser or Pro simply requires one change with these hypothetical states. No rigorous basic algorithm provides strong support for Ser being the ancestral state in such a scenario (adapted from Ref. 60).

We must consider traits from both sides of the tree (both sides of the *basal node*) to infer the most-likely ancestral

Box 2. Alternative methods for reconstructing ancestral character states

Here we review the three main methods used for discrete (categorical) data such as black versus white abdomen color, or C versus T for a given base pair of DNA. Methods are also available for continuous characters such as brain size or levels of gene expression (reviewed by Ref. 41).

1. PARSIMONY: This review has focused on parsimony because algorithms based on parsimony were the first proposed and are still the most-widely used methods. The principle of parsimony favors the reconstruction that requires the fewest changes. For example, in Fig. 3A, it is possible that the common ancestor of the group had Pro. However, that reconstruction would require two changes (e.g. Pro to Ser in both the agnathan and elasmobranch lineages) so is less parsimonious than the reconstruction shown. Parsimony reconstructions ignore branch length information—all that matters is the topology of the species tree or gene tree. Furthermore, parsimony can mislead by implying certainty even though slightly less parsimonious reconstructions might be highly likely. Only when alternatives are exactly 50/50 does parsimony show reconstructions as equivocal (basal node, Fig. 3A). For description of the algorithms used as well as some of the limitations of parsimony and other algorithms see Refs 4,15,41,86,87.

2. MAXIMUM LIKELIHOOD: Maximum likelihood reconstructions offer several potential advantages over parsimony.[15,88,89] First, information about branch lengths can be used, thereby allowing a higher probability of changes on branches that represent longer spans of time. Second, degree of uncertainty can be indicated (e.g. with pie diagrams). However, given that it is difficult to determine appropriate models of evolution for phenotypic characters,[90] many researchers may still favor simple parsimony. Consider a hypothetical example based on Figure 3B using maximum likelihood with the branch lengths as shown. This reconstruction indicates 0.74 probability of a Pro ancestor, and 0.26 probability of a Ser ancestor (reconstructions done in Mesquite[87] also see Ref. 60). Note that this reconstruction *strongly disagrees* with the "primitive lineage fallacy", which holds that the ancestral state likely would be Ser, the state of the agnathans. The states of all the taxa including tetrapods are relevant to inferring the ancestral amino acid.

3. BAYESIAN STOCHASTIC MAPPING: Bayesian methods attempt to incorporate even more components of the real uncertainty inherent in character mapping methods. First, Bayesian methods can reconstruct evolution over a set of probable tree topologies, as opposed to using just one 'best' estimate of the topology.[91] Second, Bayesian methods allow the other parameters involved to vary stochastically (e.g. branch lengths, rates of change from state 0 to 1 and back).[92] Like maximum likelihood, Bayesian methods are model-based, so there is a trade-off between incorporating complexity versus the limits of our knowledge of the detailed biology behind the characters being studied. Bridgham et al.[54] apply both ML and Bayesian reconstruction methods to ancestral amino acid sequences.

characteristics. It is not sufficient to consider only the traits from the "left" side of the tree just because that species or gene *appears* ancestral for some reason—this is the "primitive lineage fallacy".[41,58–61] Rather than looking at the "left", species-poor side of the tree to find the ancestral state, it is important to look for patterns of character nesting (e.g. Fig. 3A, see below). When species with character state A are nested among species that have state B, then that suggests state B is ancestral (based on parsimony). For larger complicated trees with multiple trait changes, it is advisable to use explicit computer algorithms (reviewed in Refs 15,41). Likely ancestral states are often misinterpreted when a species-poor lineage with state X (e.g. serine) is sister to a species-rich lineage with state Y (e.g. proline). In this case, even if one adds hundreds of gene copies to the agnathan lineage, or hundreds of gene copies to the sister lineage, there is still a 50/50 equivocal result for the reconstruction based on simple parsimony. Box 3 provides examples of organismal and molecular characters that demonstrate how patterns of nesting can be used to look for likely ancestral character states.

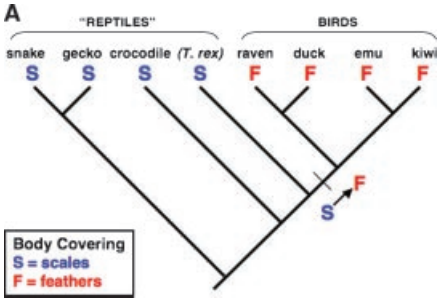
Focus of tree has a major influence on what *seems* "primitive"

Which species are shown at the "left" can be strongly influenced by many factors. In particular, this often depends on which species or genes are the focus of the study. One of the reasons that Fig. 2B appears incorrect to some readers is that the focus of the tree is on *E. coli*, with humans included as the one representative of eukaryotes. We are used to seeing trees with *H. sapiens* as the focus, therefore generally appearing at the far right of a tree (or at the top if the tree is drawn vertically[35]). Many readers may be thinking: "But these trees are incomplete". That is true. Every phylogenetic tree that has ever been published or will ever be published is incomplete. However, the point is that both sides of the node of interest must be sampled. For example, in the case of Fig. 2B, it is crucial to have at least one eukaryote sampled and at least one proteobacterium. Researchers publishing, interpreting and using trees should be aware of how tree completeness can alter the *apparent* "meaning" of the trees, especially with regard to ancestral characteristics. Even in trees that are more complete (e.g., good sampling of all species in a genus), position left to right may not tell us anything about which trait is more ancestral.

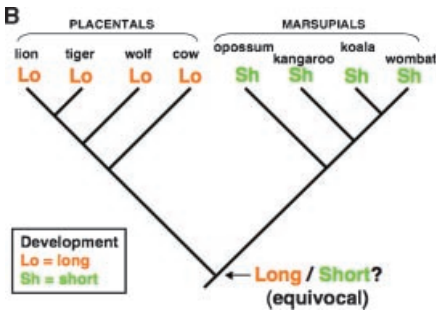
In an example involving extant mammals, opossums (e.g. Didelphidae) are sometimes incorrectly considered a "primitive mammal" (Table 1, for example see Refs 62,63), likely because they are marsupials (ergo not like "advanced" humans). In a tree that focused on placentals, especially primates, the tree could include one opossum as a representative of the marsupials, and thus appear at the left in a ladderized right tree (Fig. 4A). In contrast, if the sampling is

Box 3. Ancestral characteristics: looking for nesting of character states

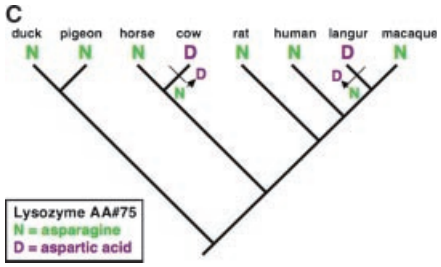
When thinking about likely ancestral states, it is crucial to look for patterns of nesting rather than looking for what to some could incorrectly interpret as "early branching lineages". For example, organisms with feathers (birds) are clearly nested within lineages that lack feathers (generally having bare scaly skin). Because "reptiles" are *paraphyletic* with respect to birds (see glossary), this helps us know that scaly skin is the ancestral state, and that feathers are recently derived. (The shared ancestral state of scaly skin was likely a major reason that "reptiles" were classified together. Phylogeny based on Ref. 34.) (Fig. A)



In contrast, here is an example in which there is no nesting of character states of extant species—the uterine development time of marsupials versus placentals. Extant placentals have long intrauterine development, whereas extant marsupials have short intrauterine development (for details on marsupial reproduction see Ref. 93). Assuming these are two sister clades that are reciprocally monophyletic,[94] these extant species provide no strong information on the likely ancestral state for their most-recent common ancestor. In fact, both clades have a derived state that is generally not shared with any other tetrapod groups. Because there is no nesting of character states, we are generally unable to use the tree alone to infer the likely traits of the most-recent common therian ancestor. Since placentals are more-species rich, and many trees focus on placentals (thus show only a few marsupials—e.g. Fig. 5B), marsupials may *appear* to some to be "basal", "older branching" and thus more "primitive". However, based on the data in this figure, no basic algorithm of rigorous ancestral state reconstruction would provide strong support for short intrauterine development (presumably accompanied by use of a pouch). (Data from other species including fossil taxa may provide additional information on the likely therian ancestral state—see Ref. 95. Phylogeny based on Ref. 96.) (Fig. B)



Here is an example from molecular evolution, mapping amino acids from the lysozyme gene onto the species tree for vertebrates. Parsimony suggests that the common ancestor likely had an N (asparagine), and that two lineages have independently converged on D (aspartic acid). This is a classic example of parallel evolution at the molecular level (Refs 97,98; amino acid position #75). The tree shows the species with D nested within the species with N. This reconstruction does not depend on the fact that N happens to be in the branch that is at the left, which to some may *seem* "early branching" or "ancestral". Using ancestral state reconstruction to reveal such examples of convergent evolution is a powerful application of modern tree thinking for both molecular and organismal biologists. (Fig. C)



Finally, for an example involving molecular characters on a *gene tree*, consider the globins from Fig. 5A. At a hypothetical site, if beta-hemoglobins have a histidine, and lysine is present at that site for all the other globins, then the genes with His are nested within the gene copies with Lys. Thus, lysine is the most-likely ancestral state. In contrast, if only lamprey has a Lys (marked with asterisk), and all the globins from mammals have His, then there is no nesting and the ancestral state is equivocal (based on parsimony). Although the lone lamprey globin to some may *appear* "basal", "primitive" or "older", in this second case the Lys could be due to a mutation that occurred in the lamprey lineage (e.g. as little as one million years ago). Looking for patterns of nesting can often give a quick sense of ancestral character states. However, using computer programs and explicit algorithms decreases the chances of simple errors. Furthermore, other scientists can then replicate or modify the specific methods used. (Fig. D)

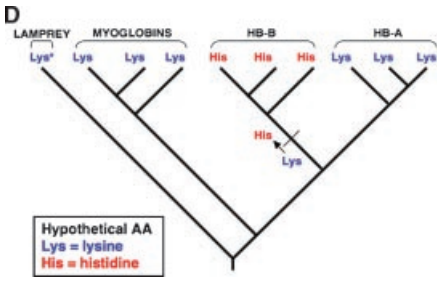


Table 1. Use of the term "primitive mammal" referring to extant species in publications from the last 10 years (1997–2006, Google Scholar, November 2007*)

Species	# Occurrences
opossum (*Didelphis* spp., *Monodelphis* spp.)	10
platypus (*Ornithorhynchus anatinus)*	3
echidna (*Tachyglossus aculeatus*)	3
shrew (*Suncus murinus*)	3
tenrec (*Echinops telfairi*)	3
armadillo (*Dasypus novemcinctus*)	2
other species+	3
Total cases	27

*Examples using the word "fossil" or geological periods were excluded. Only cases in which the exact context could be determined from the Google summary were counted.
+Other species comprise one citation each of: hedgehog (*Erinaceus europaeus*), anteater (species not specified), and capybara (*Hydrochaeris hydrochaeris*).

focused on marsupial opossums, the tree could include humans as one representative of the placentals, with humans appearing in the left position, in what seems to be the "basal" lineage (Fig. 4B) (also see Ref. 64, Fig. 1). In either case, thinking of either the opossum or humans as an ancestral or "primitive" species is misleading (also see Fig. 4C). (See Box 4, "Do outgroups have to be primitive?")

Tree focus and tree completeness are also important issues when investigating gene trees and gene family evolution. As with organismal biologists working on species trees, molecular biologists working on gene trees will focus their efforts on the gene families in which they are most interested. First consider a phylogeny that includes many representatives of the globin gene family (Fig. 5A). A laboratory focused on alpha-hemoglobins might include one copy of a myoglobin gene, leading some to see myoglobin as "basal" and more ancestral (Fig. 5B). However, a laboratory focusing on myoglobin evolution would likely include many myoglobins, including only one or a few alpha hemoglobins, thus leading some to see HB-A as more ancestral (Fig. 5C). Most gene trees and species trees will include more than five genes or more than five taxa, but the issues of tree completeness, tree focus and which taxa or genes are presented at the "left" will still remain. The important point is that both sides of a tree need to be considered when trying to infer ancestral appearance, behaviors, amino acids or gene rearrangements.

The misinterpretation of trees is most pronounced when a tree is "unbalanced"[65] with a series of species-poor lineages seeming to "branch off". This interpretation is partly a result of how trees are represented. When trees are always ladderized, this may promote poor tree reading. With such unbalanced trees, people tend to view the straight line from the root to, say,

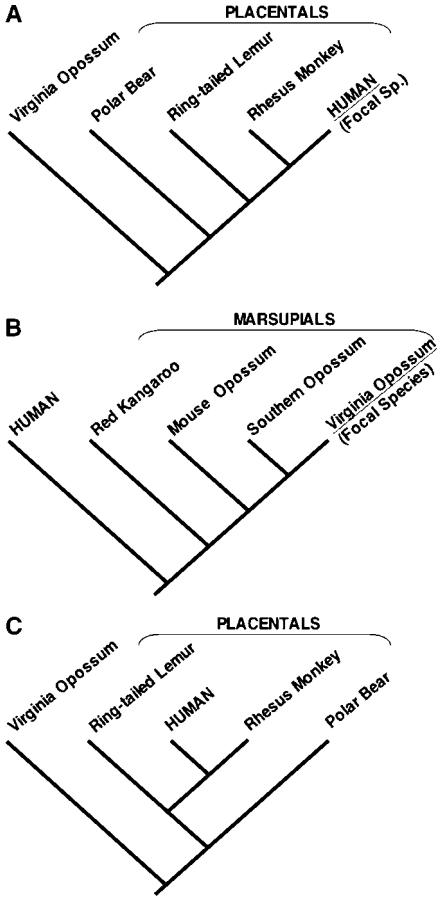


Figure 4. Species tree focus—an example with placental and marsupial mammals. **A:** Phylogeny of five mammal species focused on primates. With this rotation, humans appear at the "right", and the Virginia opossum appears at the "left". This is often interpreted as meaning opossums are "ancestral" or "primitive" (Table 1), which is how some often mistakenly think of marsupials. (Tree based on Ref. 76.) **B:** Another phylogeny of mammals including two of the same species, but here the focus of the tree is on opossums. With this rotation humans now are at the "left", which may make them appear to some as older, more ancestral, more primitive, or more simple. Position on the tree left to right, top to bottom, or even middle to edge does not mean that one extant species is any more ancestral or derived than another extant species. (Tree based on Ref. 77.) (Also see Ref. 23.) **C:** Zig-zag rotation of the same five taxa as in Fig. 4A. With this rotation, it is less tempting to think that there is a "main branch" leading directly from an ancestral possum, through lemurs and rhesus monkeys to human (see Ref. 41). Zig-zag rotation, or random rotation of branches may make it less likely that researchers read trees "left to right" as a ladder of progress. This rotation helps emphasize that the X axis has no meaning, and that all nodes can be rotated.[7] (Species included: human, *Homo sapiens*; rhesus monkey, *Macaca mulatta*; ring-tailed lemur, *Lemur catta*; polar bear, *Ursus maritimus*; Virginia opossum, *Didelphis virginiana*; southern opossum, *Didelphis marsupialis*; Mexican mouse opossum, *Marmosa mexicana*; red kangaroo, *Macropus rufus*.)

Box 4. Do outgroups have to be primitive?

Considering outgroups provides a good case study because they often are depicted on the "left" or "bottom" and thus *appear* "early branching". As a result, readers may be bothered by several of the examples, especially Fig. 1b because they think that outgroups have to be primitive. Outgroups are species or genes used to help root the phylogeny and thus provide a relative time axis for the tree.[4,99] Although it may seem correct to try to find a "primitive" or "ancestral" species or gene to be the outgroup, there is no way to find such a species or gene—all extant species are a mix of ancestral and derived traits,[41] and all present day genes contain a mix of older ancestral and more-recently derived characteristics. Therefore, perceived primitiveness should not be a basis for choosing an outgroup.

Instead, the key is to choose close outgroups that are in the sister group to the taxon or gene family being studied.[99,100] As the mammal example in Fig. 4 demonstrates, any one, two or three marsupials could serve as outgroups to root a tree of placentals (Fig. 4A), or any one, two or three placentals could serve to root a tree of marsupials (Fig. 4B). For rooting gene trees, outgroup genes should be chosen from one or more duplicate genes that are in the sister group to the focal ingroup genes.[101] Whether rooting species tree or gene trees, primitiveness is a problematic term and concept in current evolutionary biology.

humans as the "main path" of evolution (e.g. Fig. 4A). An alternative is "zig-zag" rotation, which alternates species-poor lineages left, right, left, right so that the most-species-rich lineages appear in the center of the tree (Fig. 4c, see Ref. 41). Another alternative is to use some sort of random rotation. Regardless, researchers should consider how issues such as branch rotation,[7] tree balance and tree completeness alter tree appearance thus perceptions and beliefs about evolutionary history.[5]

Problematic terms for leftists and rightists

There are a number of problematic terms related to the imprecise evolutionary thinking that we are describing, most prominently *primitive*. For example, in birds the only extant species that are labelled as "primitive bird" are those from species-poor lineages. For example, ratites are frequently referred to as primitive (e.g. ostrich, *Struthio camelus*;[66,67] emu, *Dromaius novaehollandiae*;[68,69] kiwi, *Apteryx australis*—"the most-primitive bird"[70]). Speaking of ratites as primitive is surprising given that one of their most-obvious characteristics—lack of flight—is known to be secondarily derived.[71] Little about any of these birds strongly suggests a close affinity or resemblance to *Archaeopteryx* or other ancient fossil birds (for example Ref. 72).

While speaking of any extant species as primitive or ancestral is problematic, it is also misleading to speak of any species as "older" just based on a tree, especially a cladogram. Similarly, speaking of some species as belonging to a "deeper" lineage is misleading. Finally, it is incorrect to speak about such species as "branching off first" or "earlier" than another species—this would imply a main branch or a known end point of evolution. However, unlike Haeckel's 19th century phylogenies (Fig. 1), there is no "main branch" or "main stem" in 21st century phylogenies. This language about lineages "branching off" is misleading about the process of evolution. *There is no one main path; there is no goal to evolution*. Stating which is the lineage to "branch off" requires a fixed reference that defines where evolution is heading (e.g. that *Drosophila melanogaster* is the pinnacle of animal evolution, or that the rose is the epitome of plantness—see Ref. 5). Although one of the lineages may *appear* to some as "early branching" or more "basal",[41,58] in phylogenies of extant species or genes, for any node there are at least two descendent lineages, each of which have continued to evolve to the present. With any of these problematic terms, a one-sided view of branching, focusing just on the "left" side of trees, is misleading.

There is a corresponding set of problematic terms meant to apply to species on the "other side" of the tree, the "right" or more-species-rich side. To speak of any such species, branches or lineages as derived, young, recent or the "last to branch off" is also misleading. Just as it is incorrect to speak of any extant species as ancestral or primitive, it is also problematic to speak of "derived species" or species that "branched off last" simply based on position on a cladogram. Nevertheless, "more-derived species" and "molecular phylogeny" appear together in 46 papers tracked by Google Scholar (November 2007).

So instead, "*sister group*" should be used whenever possible when describing trees.[41,58] Marsupials can be described as the sister group to placentals, whether the focus is on placentals and only one marsupial is sequenced or vice versa (Fig. 4A,B). Gibbons are the sister group to the great apes, regardless of how many Hylobatidae or how many Hominidae are sampled. The myoglobin gene family is sister to the alpha-beta hemoglobin clade regardless of which or how many gene copies are included (Fig. 5). Furthermore, rather than writing about every extant species or gene on a tree, it is often best to refer to well-labelled internal nodes (e.g. node 1, 2, 3), to help avoid problematic evolutionary thinking. Both careful use of sister group language and references to figures should decrease chances that researchers think that species-poor or gene-poor lineages generally express the ancestral state.

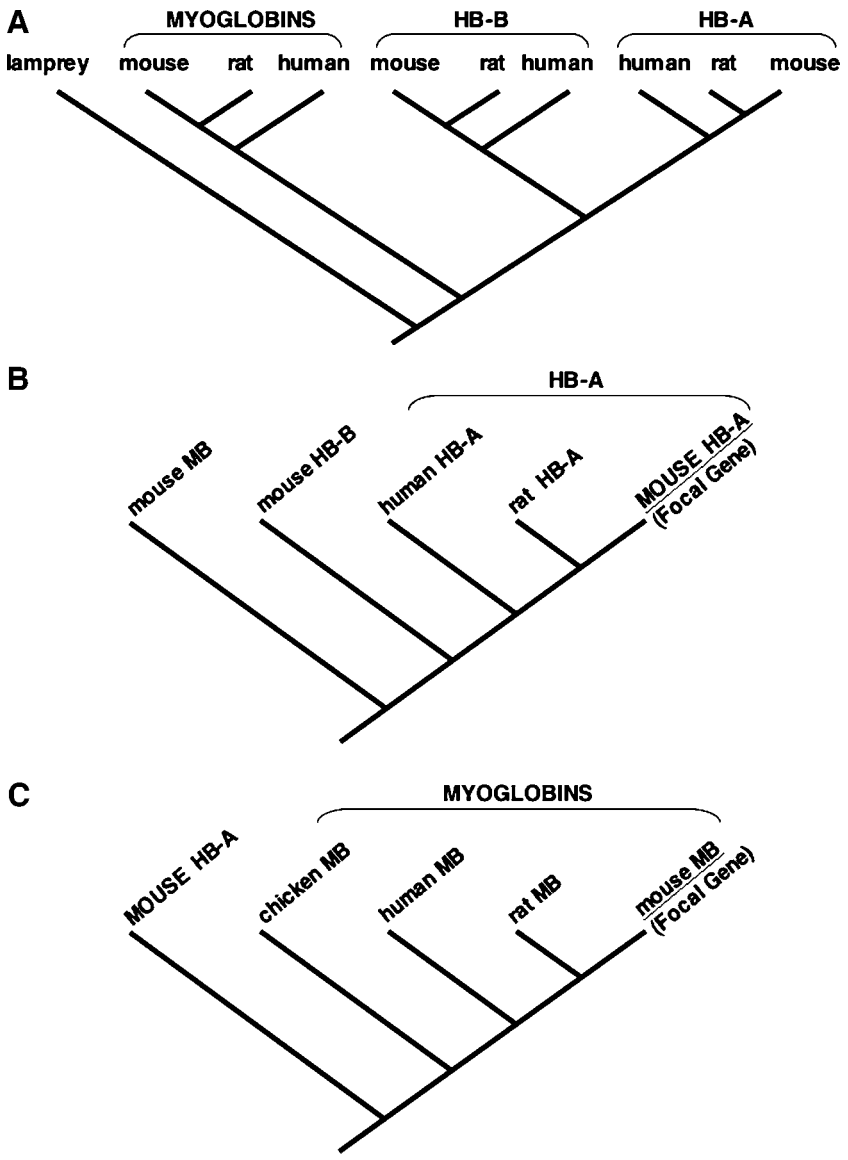


Figure 5. Gene tree focus—the globin gene family. **A:** Gene tree of the globin gene family with balanced sampling for three main vertebrate globin groups (tree based on Ref. 78). Note that, with this rotation, an incorrect reading of the tree could cause some to interpret myogobin as "intermediate" between lamprey globin and beta hemoglobins. However, gene sampling and branch rotation[7] strongly influence which genes *appear* "intermediate", or more "primitive" or "ancestral". Order on the page and which genes or species are next to each other provide little or no information about which *characteristics* are ancestral. **B:** Another globin gene phylogeny, this time with sampling focused on the alpha hemoglobins, thus an incorrect reading of the tree might consider beta hemoglobin and myoglobin "older" or "more ancestral". But again, the myoglobins and the hemoglobins are sister groups. **C:** A third globin gene phylogeny, but this time the focus is on myoglobins. Thus the alpha hemoglobin included for comparison may *appear* to some to be "older" and "more primitive". However, position on a gene tree does not indicate age, number of ancestral features, or "primitiveness".

Conclusions—cousin thinking needed

So indeed just as all humans alive today are at least distant cousins of each other, all extant species are also our evolutionary cousins. Furthermore, the gene copies that we carry are evolutionary cousins of the gene copies in Charles Darwin's lineage, the chimp lineage, the Galapagos finch lineage, and the zebrafish lineage. Gene copies in humans are evolutionary cousins to the gene copies found today in *E. coli*. Looking for ancestors in the tops of trees showing extant species should be replaced by determining which *characteristics* are ancestral. None of these extant species is ancestral to other extant species—none of these individuals, genes or

species are parents or grandparents to other present generation entities. We are all evolutionary cousins! Fully grasping this crucial point will help researchers and teachers in all fields of biology better understand and use phylogenies to understand how characters evolve.

The problems with tree thinking that we focused on are pervasive; they appear in the full range of journals from the highest impact general science publications, across all taxonomic groups, and in subfields as wide ranging as genetics, development, and cell biology as well as evolution and systematics. Such a widespread problem will need to be tackled from a variety of levels. However, one way to address this problem is through editors and editorial boards of journals. Just as a journal might have statistical consultants, it would be prudent for journals across all of biology to have phylogenetic consultants or editors charged with ensuring that authors present, interpret and discuss phylogenetic trees correctly. The problem of the ''primitive lineage fallacy'', confusion about apparent ''early branching lineages'', and reading trees ''left to right'' represent fundamental, deep-seated and widespread misinterpretation that needs to be addressed across the full breadth of modern biology. Phylogenies of extant species and gene families do not show us which extant entities are ancestral, they show us relationships among 21st century evolutionary cousins.

Acknowledgments

We thank Steve Cameron, Richard Carter, Robert Edwards, Phil Farabaugh, Nick Friedman, Chris Hofmann, Beatrice Kondo, Celeste Linde, Tamra Mendelson, Steve Miller, David Morris, Lindsay Popple, Jordan Price, Sonja Scheffer, Andrew Thornhill and David Yeates, as well as other colleagues at ANU, UQ and UMBC who gave us comments on the manuscript. Matt Phillips provided feedback on the marsupial–placental examples. KEO thanks the ANU School of Botany and Zoology, and Heads of School, William Foley and Andrew Cockburn, for an ANU Visiting Fellowship.

References

1. Hennig W. 1966. Phylogenetic Systematics. Urbana: University of Illinois Press.
2. Cavalli-Svorza LL, Edwards AWF. 1967. Phylogenetic analysis: models and estimation procedures. Evolution 21:550–570.
3. Brooks DR, McLennan DA. 1991. Phylogeny, Ecology, and Behavior: a Research Program in Comparative Biology. Chicago: University of Chicago Press.
4. Maddison DR, Maddison WP. 1992. MacClade Manual. Sunderland: Sinauer.
5. O'hara RJ. 1997. Population thinking and tree thinking in systematics. Zoologica Scripta 26:323–329.
6. Page RDM, Holmes EC. 1998. Molecular Evolution: A Phylogenetic Approach. London: Blackwell Science.
7. Baum DA, Smith SD, Donovan SSS. 2005. The tree-thinking challenge. Science 310:979–980.
8. Jermann TM, Opitz JG, Stackhouse J, Benner SA. 1995. Reconstructing the evolutionary history of the artiodactyl ribonuclease superfamily. Nature 374:57–59.
9. Sassi SO, Benner SA. 2007. The resurrection of ribonucleases from mammals: from ecology to medicine. In: Liberles DA, editor. Ancestral Sequence Reconstruction. Oxford: Oxford University Press. p 208–224.
10. Chang BSW, Matz MV, Field SF, Müller J, van Hazel I. 2007. Dealing with model uncertainty in reconstructing ancestral proteins in the laboratory: examples from ancestral visual pigments and GFP-like proteins. In: Liberles DA, editor. Ancestral Sequence Reconstruction. Oxford: Oxford University Press. p 164–180.
11. Serb J, Oakley TH. 2005. Hierarchical phylogenetics as a quantitative analytical framework for evolutionary developmental biology. BioEssays 27:1158–1166.
12. Oakley TH, Ostman B, Wilson ACV. 2006. Repression and loss of gene expression outpaces activation and gain in recently duplicated fly genes. Proc Natl Acad Sci USA 103:11637–11641.
13. Muffato M, Crollius HR. 2008. Paleogenomics in vertebrates, or the recovery of lost genomes from the mist of time. BioEssays 30:122–134.
14. Ryan MJ, Rand AS. 1995. Female responses to ancestral advertisement calls in the túngara frog. Science:390–392.
15. Cunningham CW, Omland KE, Oakley TH. 1998. Reconstructing ancestral character states: a critical reappraisal. Trends Ecol Evol 13: 361–366.
16. Liberles DA, editor. 2007. Ancestral Sequence Reconstruction. Oxford: Oxford University Press.
17. Haeckel E. 1866. General Morphology of Organisms. Berlin: Reimer.
18. Lovejoy AO. 1936. The Great Chain of Being: The History of an Idea. Cambridge: Harvard University Press.
19. Gould SJ. 1997. Redrafting the tree of life. Proc Am Phill Soc 141:30–54.
20. Bowler BJ. 2003. Evolution: the History of an Idea. Berkeley: University of California Press.
21. Dayrat B. 2003. The roots of phylogeny: how did Haeckel build his trees? Systematic Biology 52:515–527.
22. Sandvik H. 2008. Tree thinking cannot taken for granted: challenges for teaching phylogenetics. Theory in Biosci 127:45–51.
23. Gregory T. 2008. Understanding evolutionary trees. Evolution: Education and Outreach 1:121–137.
24. Darwin C. 1859. The Origin of Species by Means of Natural Selection. London: John Murray.
25. Wilson AC, Carlson SS, White TJ. 1977. Biochemical evolution. Annual Reviews of Biochemistry 46:573–639.
26. Gould SJ. 2002. The Structure of Evolutionary Theory. Cambridge: Harvard University Press.
27. Felsenstein J. 2003. Inferring Phylogenies. Sunderland, MA: Sinauer.
28. Avise JC. 2004. Molecular Markers, Natural History and Evolution, Second Edition. Sunderland: Sinauer.
29. Kimura M. 1983. The Neutral Theory of Molecular Evolution. Cambridge: Cambridge University Press.
30. Cherry LM, Case SM, Wilson AC. 1978. Frog perspective on morphological difference between humans and chimpanzees. Science 200:209–211.
31. Schopf TJM. 1984. Rates of evolution and the notion of ''living fossils''. Ann Rev Earth Planetary Sci 12:245–292.
32. Avise JC, Nelson WS, Sugita H. 1994. A speciational history of ''living fossils'': molecular evolutionary history in horseshoe crabs. Evolution 48:1986–2001.
33. Papadopoulos D, Schneider D, Meier-Eiss J, Arber W, Lenski RE, Blot M. 1999. Genomic evolution during a 10,000-generation experiment with bacteria. Proc Natl Acad Sci USA 96:3807–3812.
34. Sugden AM, Jasny BR, Culotta E, Pennisi E. 2003. ''A Tree of Life'', charting the evolutionary history of life. Science 300:1691–1695.
35. Nee S. 2005. The great chain of being. Nature 435:429–429.
36. Doolittle WF. 2005. The origin and early evolution of life. In: Cracraft J, Bybee RW, editors. Evolutionary Science and Society: Educating a New Generation. Washington, D.C: Biological Sciences Curriculum Study, American Institute of. Biological Sciences.
37. Knox RB, Ladiges P, Evans B, Saint R. 2001. Biology, 2nd Edition. New South Wales, Australia: McGraw-Hill.
38. Campbell NA, Reece JB. 2005. Biology, 7th Edition. Menlo Park: Benjamin Cummings.
39. Palmer JD, Soltis DE, Chase MW. 2004. The plant tree of life: an overview and some points of view. Am J Bot 91:1437–1445.

40. Kenrick P, Crane PR. 1997. The Origin and Early Diversification of Land Plants: A Cladistic Study. Washington, D.C: Smithsonian Institution Press.
41. Crisp MD, Cook LG. 2005. Do early branching lineages signify ancestral traits? Trends in Ecology and Evolution 20:122–128.
42. Warren W, Hillier L, Graves J, Birney E, Ponting C, et al. 2008. Genome analysis of the platypus reveals unique signatures of evolution. Nature 453:175–183.
43. Grant T. 1995. The Platypus, A Unique Animal, 2nd Edition. Sydney: New South Wales University Press.
44. Milinkovitch MC, Tzika A. 2007. Escaping the mouse trap: the selection of new Evo-Devo model species. J Exp Zool Part B: Molec Develop Evol 308B:337–346.
45. Barraclough TG, Harvey PH, Nee S. 1996. Rate of rbcL gene sequence evolution and species diversification in flowering plants (angiosperms). Proceedings of the Royal Society of London Biological Sciences Series B 263:589–591.
46. Omland KE. 1997. Correlated rates of molecular and morphological evolution. Evolution 51:1381–1393.
47. Bromham L, Woolfit M, Lee MSY, Rambaut A. 2002. Testing the relationship between morphological and molecular rates of change along phylogenies. Evolution 56:1921–1930.
48. Pagel M. 2006. Large punctuational contribution of speciation to evolutionary divergence at the molecular level. Science 314:119–121.
49. Bromham L. 2003. Molecular clocks and explosive radiations. J Molec Evol 57:13–20.
50. Venditti C, Meade A, Pagel M. 2006. Detecting the node-density artifact in phylogeny reconstruction. Systematic Biol 55:637–643.
51. Eastman SD, Chen THP, Falk MM, Mendelson TC, Iovine MK. 2006. Phylogenetic analysis of three complete gap junction gene families reveals lineage-specific duplications and highly supported gene classes. Genomics 87:265–274.
52. Alvarez-Buylla ER, Liljegren SJ, Pelaz S, Gold SE, Burgeff C, et al. 2000. MADS-box gene evolution beyond flowers: expression in pollen, endosperm, guard cells, roots and trichomes. Plant Journal 24:457–466.
53. Chang BSW, Donoghue MJ. 2000. Recreating ancestral proteins. Trends in Ecology and Evolution 15:109–114.
54. Bridgham JT, Carroll SM, Thornton JW. 2006. Evolution of hormone-receptor complexity by molecular exploitation. Science 312:97–101.
55. Maddison DR. 1994. Phylogenetic methods for inferring the evolutionary history and processes of change in discretely valued characters. Ann Rev Entomol 39:267–292.
56. Omland KE. 1997. Examining two standard assumptions of ancestral reconstructions: repeated loss of dichromatism in dabbling ducks (Anatini). Evolution 51:1636–1646.
57. Cook LG, Crisp MD. 2005. Directional asymmetry of long-distance dispersal and colonization could mislead reconstructions of biogeography. J Biogeog 32:741–754.
58. Krell FT, Cranston P. 2004. Which side of the tree is more basal? (Editorial). System Entomol 29:279–281.
59. Jenner RA. 2006. Unburdening the evo-devo: ancestral attractions, model organisms, and basal baloney. Developmental Genes and Evolution 216:385–394.
60. Omland KE, Hofmann CM. 2006. Adding color to the past: ancestral state reconstruction of bird coloration. In: Hill GE, McGraw KJ, editors. Bird Coloration Volume 2: Function and Evolution. Cambridge, MA: Harvard University Press. p 417–454.
61. Jenner RA, Wills MA. 2007. The choice of model organisms in evo-devo. Nature Reviews Genetics 8:311–319.
62. Ahnelt PK, Hokoc JN, Rohlich P. 1995. Photoreceptors in a primitive mammal, the South-American opossum, *Didelphis marsupialis aurita*—characterization with anti-opsin immunolabeling. Visual Neurosci 12: 793–804.
63. Frost SB, Milliken GW, Plautz EJ, Masterton RB, Nudo RJ. 2000. Somatosensory and motor representations in cerebral cortex of a primitive mammal (*Monodelphis domestica*): A window into the early evolution of sensorimotor cortex. J Comp Neurol 421:29–51.
64. Collins AG, Cartwright P, McFadden CS, Schierwater B. 2005. Phylogenetic context and basal metazoan model systems. Integrative and Comparative Biology 45:585–594.
65. Mooers AH, Heard SB. 2002. Using tree shape. System Biol 51:833–834.
66. Ozegbe P, Aire T, Soley J. 2006. The morphology of the efferent ducts of the testis of the ostrich, a primitive bird. Anat Embryol 211:559–565.
67. Wicher KB, Fries E. 2006. Haptoglobin, a hemoglobin-binding plasma protein, is present in bony fish and mammals but not in frog and chicken. Proc Natl Acad Sci USA 103:4168–4173.
68. Fischer FP. 1998. Hair cell morphology and innervation in the basilar papilla of the emu (*Dromaius novaehollandiae*). Hearing Res 121:112–124.
69. Koppl C, Manley GA, Konishi M. 2000. Auditory processing in birds. Curr Opn Neurobiol 10:474–481.
70. Duncker H-R. 2004. Vertebrate lungs: structure, topography and mechanics: A comparative perspective of the progressive integration of respiratory system, locomotor apparatus and ontogenetic development. Resp Physiol Neurobiol 144:111–124.
71. Proctor NS, Lynch PJ. 1993. Manual of Ornithology: Avian Structure and Function. New Haven: Yale University Press.
72. Sereno PC, Chenggang R. 1992. Early evolution of avian flight and perching: new evidence from the lower Cretaceous of China. Science 255:845–848.
73. Ciccarelli FD, Doerks T, von Mering C, Creevey CJ, Snel B, Bork P. 2006. Toward automatic reconstruction of a highly resolved tree of life. Science 311:1283–1287.
74. Lake JA, Herbold CW, Rivera MC, Servin JA, Skophammer RG. 2007. Rooting the tree of life using nonubiquitous genes. Molec Biol Evol 24:130–136.
75. Doolittle WF, Bapteste E. 2007. Pattern pluralism and the Tree of Life hypothesis. Proceedings of the Natl Acad Sci 104:2043–2049.
76. Yoder AD, Cartmill M, Ruvolo M, Smith K, Vilgalys R. 1996. Ancient single origin for Malagasy primates. Proceedings of the National Academy of Sciences USA 93:5122–5126.
77. Jansa SA, Voss RS. 2000. Phylogenetic studies on didelphid marsupials I. Introduction and preliminary results from nuclear IRBP gene sequences. J Mam Evol 7:43–77.
78. Holmquist R, Jukes TH, Moise H, Goodman M, Moore GW. 1976. The evolution of the globin family genes: Concordance of stochastic and augmented maximum parsimony genetic distances for alpha hemoglobin, beta hemoglobin and myoglobin phylogenies. J Molec Biol 105:39–74.
79. Castelloe J, Templeton AR. 1994. Root probabilities for intraspecific gene trees under neutral coalescent theory. Molec Phylogen Evol 3: 102–113.
80. Harrison RG. 1991. Molecular changes at speciation. Ann Rev Ecol Syst 22:281–308.
81. Crisp MD, Chandler GT. 1996. Paraphyletic species. Telopea 6:813–844.
82. Doyle JA, Donoghue MJ. 1993. Phylogenetics and angiosperm diversification. Paleobiology 19:141–167.
83. Funk DJ, Omland KE. 2003. Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. Annual Rev Ecol, Evol Syst 34:397–423.
84. Swofford DL. 2002. PAUP*: Phylogenetic Analysis Using Parsimony (*and other methods). Version 4.0. Sunderland: Sinauer Associates.
85. Page RDM. 1996. TreeView: An application to display phylogenetic trees on personal computers. Compar Appl Biosci 12:357–358.
86. Omland KE. 1999. The assumptions and challenges of ancestral state reconstructions. Systematic Biology 48:604–611.
87. Maddison WP, Maddison DR. 2007. Mesquite: a modular system for evolutionary analysis. Version 2.01 http://mesquiteproject.org
88. Schluter D, Price T, Mooers AO, Ludwig D. 1997. Likelihood of ancestor states in adaptive radiation. Evolution 51:1699–1711.
89. Pagel M. 1999. The maximum likelihood approach to reconstructing ancestral character states of discrete characters on phylogenies. Syst Biol 48:612–622.
90. Lewis PO. 2001. A likelihood approach to estimating phylogeny from discrete morphological character data. Syst Biol 50:913–925.
91. Huelsenbeck JP, Nielsen R, Bollback JP. 2003. Stochastic mapping of morphological characters. Systematic Biology 52:131–158.
92. Ronquist F. 2004. Bayesian inference of character evolution. Trends in Ecology and Evolution 19:475–481.

93. Freyer C, Zeller U, Renfree MB. 2003. The marsupial placenta: A phylogenetic analysis. J Experi Zool Part A: Comp Exper Biol 299A:59–77.
94. Phillips MJ, Penny D. 2003. The root of the mammalian tree inferred from whole mitochondrial genomes. Molec Phylogen Evol 28:171–185.
95. Reilly SM, White TD. 2003. Hypaxial motor patterns and the function of epipubic bones in primitive mammals. Science 299:400–402.
96. Bininda-Emonds ORP, Cardillo M, Jones KE, MacPhee RDE, Beck RMD, et al. 2007. The delayed rise of present-day mammals. Nature 446:507–512.
97. Stewart C-B, Schilling JW, Wilson AC. 1987. Adaptive evolution in the stomach enzymes of foregut fermenters. Nature 330:401–404.
98. Kornegay JR, Schilling JW, Wilson AC. 1994. Molecular adaptation of a leaf-eating bird: stomach lysozyme of the hoatzin. Molecular Biology and Evolution 11:921–928.
99. Smith AB. 1994. Rooting molecular trees: problems and strategies. Biological Journal of the Linnean Society 51:279–292.
100. Maddison WP, Donoghue MJ, Maddison DR. 1984. Outgroup analysis and parsimony. Syst Zool 33:83–103.
101. Mathews S, Donoghue MJ. 2000. Basal angiosperm phylogeny inferred from duplicate phytochromes A and C. Internl J Plant Sci 161:S41–S55.