Phylogenies and Tree-Thinking

Article *in* The American Biology Teacher · January 2009

DOI: 10.1662/0002-7685(2008)70[222:PT]2.0.CO;2

David A Baum
University of Wisconsin–Madison

PHYLOGENIES & TREE-THINKING

David A. Baum Susan Offner

Phylogenetic trees, once peculiarities of systematics, now permeate almost all branches of biology and are appearing in increasing numbers in biology textbooks. While few state standards explicitly require knowledge of phylogenetics, most require some knowledge of evolutionary biology and many scientists and educators would hold that it is impossible to really understand evolution without an ability to accurately interpret phylogenetic trees (O'Hara, 1988, 1997). Evolution, at its core, is a claim that living species are related by descent from common ancestry, and as such it is a theory of evolutionary trees. Additionally, trees help integrate evolutionary concepts throughout the curriculum (e.g., Offner, 2001) and provide students with an organizational framework for structuring knowledge of biological diversity. Therefore, biological literacy requires some exposure to tree-thinking – the ability to conceptualize evolution in terms of phylogenetic trees. As noted by O'Hara (1997): "... just as beginning students in geography need to be taught how to read maps, so beginning students in biology should be taught how to read trees and to understand what trees communicate."

Such research as exists suggests that students hold significant misconceptions about trees and that these views may be deeply held and persistent (Baum et al., 2005). Therefore, the challenge faced by teachers, most of whom have had little exposure to phylogenetics, is significant. In this article we will provide a brief overview of some important principles of tree-thinking and a list of specific skills in which high school and college students should become proficient. We will also briefly discuss strategies for bringing trees into the broader biology curriculum.

What a Tree Represents

A phylogenetic tree is a depiction of the inferred evolutionary relationships among a set of species (or other taxa). When introducing trees to students it can be helpful to make clear the connection between reproduction within populations over short time frames and the evolution along the branches of a tree over a longer period of time. A useful strategy is to "zoom out" from a single population at a single point in time to a phylogeny representing much longer periods of time. One of us (DB) uses Figures 1-3 for this purpose, both in a lecture format and in an assigned reading.

character	a trait or feature that varies among a set of taxa (e.g., hair color)
character-state	a variant of a character that occurs in a particular taxon (e.g., black hair)
chronogram	a style of tree diagram in which each branch's length is proportional to its inferred duration
clade	a monophyletic group: A common ancestor and all its descendants
cladistics	more or less synonymous to "phylogenetics," but generally connoting a particular philosophical approach built around the principle of parsimony
cladogram	either a general term for a tree diagram, or a particular style of tree diagram in which neither the amount of change nor time is depicted
monophyly	when a set of organisms are all descended from a common ancestor and that ancestor did not give rise to any organisms that are not in the set
paraphyly	a variant of non-monophyly
phylogeny	an evolutionary tree
phylogenetics	the study of phylogenetic relationships and the use of phylogenetic trees to elucidate evolutionary phenomena
phylogram	a style of tree diagram in which each branch's length is drawn proportional to the inferred amount of change on that branch
plesiomorphy	an ancestral character state (e.g., limb presence is plesiomorphic in reptiles)
polyphyly	a variant of non-monophyly
polytomy	a node on a phylogenetic tree that depicts an ancestral lineage dividing into three or more descendant lineages (opposite = dichotomy)
synapomorphy	a derived character state that is shared by a group of organisms (e.g., hair is a synapomorphy of mammals; Loss of the hind limb is a synapomorphy of snakes)
systematics	the branch of biology concerned with using evolutionary and phylogenetic principles to organize classification systems and to understand the origin and maintenance of biological diversity
taxon	a formally-named group of organisms
taxonomy	the scientific discipline concerned with discovering, naming, and classifying taxa

Table 1. Glossary of Terms Used in Phylogenetics

David A. Baum is Professor, Department of Botany, University of Wisconsin, Madison, WI 53706; e-mail: dbaum@wisc.edu. Susan Offner is a science teacher at Lexington High School, Lexington, MA 02421; e-mail: soffner@ix.netcom.com.

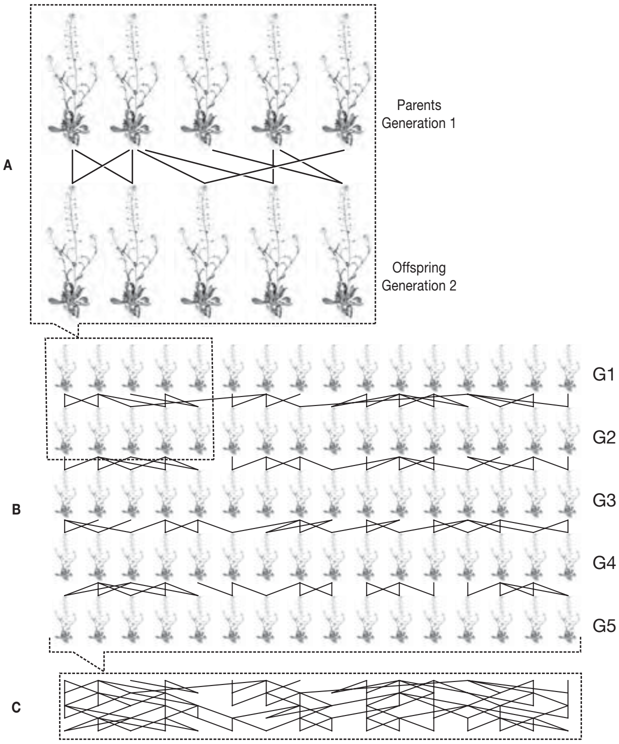

Figure 1. A hypothetical genealogy of a local population of plants. **A.** Two generations of five organisms. **B.** A local population viewed over five generations that includes the organisms illustrated in A. **C.** The same genealogy shown in B with only the lines of descent shown.

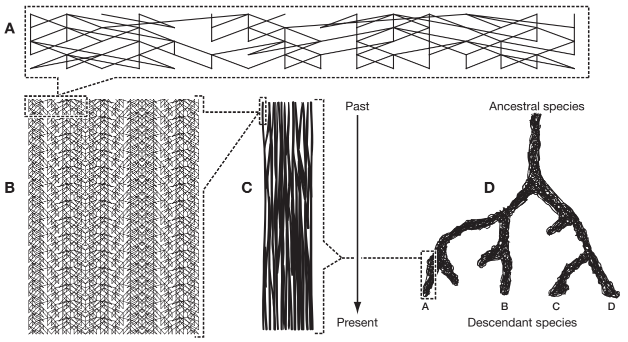

Figure 2. An illustration of how the genealogy of a sexual population is included in a phylogenetic tree. **A.** The local population illustrated in Figure 1C. **B.** An expanded view showing the population over many generations, including the organisms and generations illustrated in Figure 2A. **C.** A species lineage consisting of population lineages that are interconnected by occasional interbreeding. **D.** A branching phylogeny in which one ancestral population gives rise to four living species.

Students are asked to imagine one generation of plants of a particular species, for example, shepherd's purse, *Capsella bursa-pastoris*, growing side by side in a meadow and producing offspring by exchanging pollen. Five individual plants in a parental generation (G1) and an offspring generation (G2) could have a pedigree like that shown in Figure 1A. You can expand the frame to encompass all the plants in this population and several generations (Figure 1B). To encourage students to examine this figure closely you can give students a version without the time axis included . They can usually figure out the direction of time from the fact that each individual has two parents, but gives rise to a variable number of offspring.

The next step is to imagine taking the preceding figure and getting rid of the organisms, keeping only the descent relationships, since it is these that "glue" together the members of a sexual population. The resulting image might look like Figure 1C. One can then expand the field of view to include many more individuals and generations. For example, Figure 2B is like Figure 2A except it includes about 250 individuals and 80 generations. As you can see, if one were to try to represent a typical population of several thousand individuals that persists for hundreds or thousands of generations, all one would see would be a fuzzy line. Individual populations may be fairly isolated for some period of time. However, on an evolutionary timescale, seeds and pollen occasionally move between the discrete populations that comprise a typical species. This gene flow between populations, has the effect of "braiding" the population lineages into a single species lineage (Figure 2C).

The next question is to consider what happens when populations become genetically isolated from one another for a long period of time, for example because of dispersal of a few individuals to a new, isolated region (e.g., an island), or the splitting of a formerly contiguous range by geological or climatic events (e.g., when mountains, rivers, or barriers of inhospitable environments arise). Would these populations be expected to remain identical to one another indefinitely? Students generally see that genetically-isolated populations will tend to diverge and further, that given enough time, it would become impossible for individuals from the separate lineages to mate successfully and create viable offspring. Through this reasoning, students can discover for themselves the principle of allopatric speciation and that speciation does not require "special" evolutionary phenomena, just "normal" evolution in isolated populations. They can also extrapolate to imagine a multi-species phylogenetic tree (as in Figure 2D).

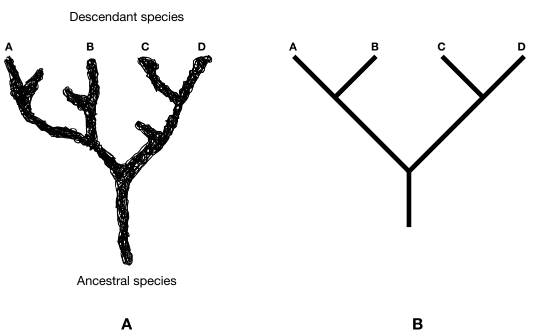

Figure 3. A tree diagram is meant to represent an actual history of evolutionary lineages that have branched over time. The most critical facts about the real history shown in panel A are summarized in panel B; namely that species A and B derived from an ancestral lineage that did not give rise to C and D, and vice versa.

When drawing trees, it is common to invert the arrow of time, placing the past at the bottom and the present at the top (Figure 3). This convention probably arose because in fossil beds, older (ancestral) fossils tend to lie in lower strata than fossils of lesser age. Also, drawn in this orientation, the figure looks more like a conventional tree rooted in the ground.

Students are likely to encounter a variety of different shapes of trees in their reading (Figure 4). Some trees are drawn with diagonal lines, others with rectangular lines, and still others with curved lines. All these formats are valid, but authors typically select tree formats that they think will present the data in the most understandable or accessible way. Our experience suggests that students find diagonal trees to be especially confusing. It is an open question whether this argues for specifically using diagonal trees, knowing that students will likely encounter them at some time, or begin by using other tree styles that resonate more with a typical student's preconceptions. Figure 5 shows a simple, diagonal rooted tree with some terms indicated.

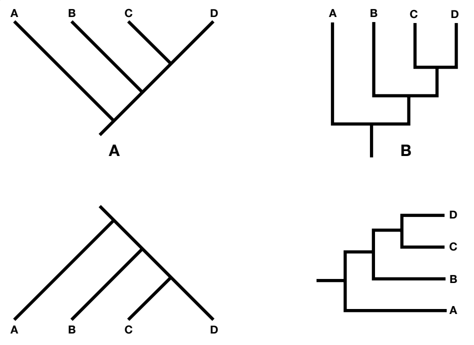

Figure 4. Four alternative trees all of which depict the same evolutionary history. In each case, taxa B, C, and D share an ancestral lineage that is not ancestral to A, and taxa C and D share an ancestral lineage that is not ancestral to A or B.

Evolutionary Relatedness & Phylogenies

When biologists talk of "relatedness," they are usually referring to recency of common ancestry: Two living species are closely related if their most recent common ancestor lived close to the present, and more distantly related if their most recent common ancestor lived in the more distant past. A helpful introduction to this material is to stress the parallels between relationships among species and among individuals within families. The last common ancestors of you and your first cousins are your grandparents, whereas the last common ancestors of you and your second cousins are your great-grandparents. Your grandparents are situated only two generations before you, whereas your great-grandparents are situated three generations back. This provides a basis for the assertion that you are more closely related to your first than your second cousins. Students can be quizzed on the degree of relatedness of certain individuals in family trees (either their own pedigree or published trees, for example, of royal families), thereby training them to pick out points of common ancestry. Phylogenetic trees contain information about the relative recency of common ancestry of species and, thus, provide a succinct way of representing their degree of relationship.

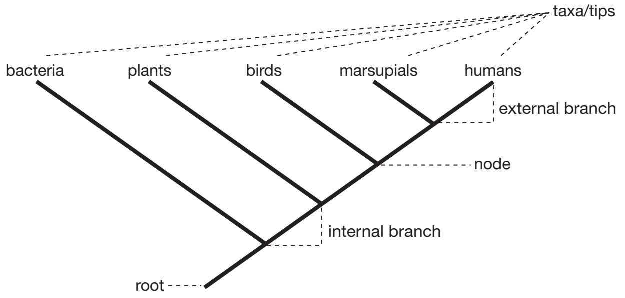

Figure 5. Some important terms used to describe phylogenetic trees.

Students need to learn to focus on the relative branching order of a tree, because it is this that contains information about relatedness. In so doing they must avoid being distracted by the shape of the tree or the proximity of tips to each other. A quick look at the tree in Figure 6A might suggest that taxa A and B are closely related because the tip labels are right next to each other. In fact A and B are as distantly related as any pair of taxa on the tree. Indeed B is more closely related to E than to A. To see this, find the last common ancestor of A and B (Node 1). Then find the last common ancestor of B and E (Node 2). Because the latter is further from the root than the former, we can see that B had a common ancestor with E more recently than it had a common ancestor with A and is therefore more closely related to E than it is to A.

If students have a hard time reading the tree this way, it might help to put arrows on all branches that point away from the root (Figure 6B). Again the last common ancestor of B and E is labeled "Node 2" and that of B and A is labeled "Node 1." Note that an arrow points from Node 1 to Node 2. This shows that Node 2 is a descendant of Node 1. It also means that there is a part of the evolutionary history that was shared by B and E, that was not shared by A and B. This should help clarify why this tree implies that B is more closely related to E than to A. For simplicity, this figure includes only a few species, but the same principles can be applied to larger trees that are more representative of the real magnitude of biological diversity.

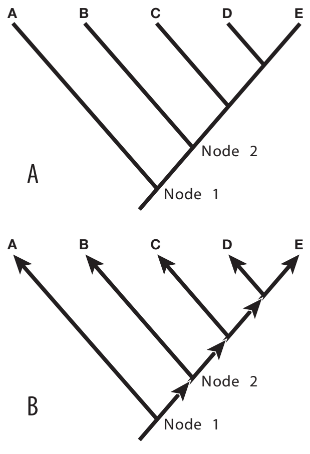

Figure 6. How to read a tree in terms of evolutionary relationships. B is more closely related to E than to A. This can be seen because the last common ancestor of B and A (Node 1) is an ancestor of Node 2, the last common ancestor of B and E. The lower tree has arrows added to help clarify the direction of descent from the root.

A minor point of confusion that may arise is a tendency to read some tips as actually being ancestors. For example, a tree that includes humans and chimpanzees might be misread as showing that humans descended *from* chimpanzees. Except in very special circumstances (involving very recent divergences), biologists never view one living species as the ancestor of another. If this is unclear, an analogy can be drawn to human pedigrees. Just as you are related to, but not descended from your cousin, so are humans related to, but not descended from chimpanzees. Just as your grandparent is neither you nor your cousin, so the common ancestor of a human and a chimpanzee was neither a human nor a chimpanzee.

One of the reasons that students make mistakes in reading relatedness from trees is that they tend to be distracted by looking at the order of taxa along the tips of a tree. However, the ordering of taxa is not

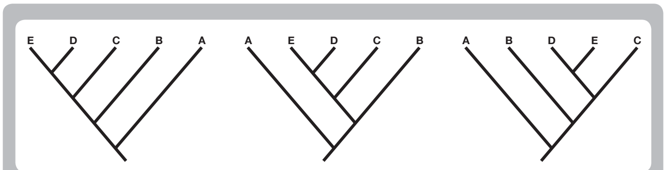

Figure 7. Three alternative trees all of which depict the same evolutionary history. The trees can be inter-converted simply by rotating parts of the tree, without ever having to cut a branch off and reattach it somewhere else.

meaningful: Two trees showing the same fundamental relationships can have the taxa in a different order. Recalling the way that a phylogeny grows by lineage splitting, it is arbitrary which descendant lineage one draws to the right or left. Thus, one can spin parts of the tree around any internode without changing the implied relationships. A tree should be thought of, not as a static path drawn on a map, but as a mobile, with flexible branches and joints: If you can change one tree into another tree by simply twisting or bending branches, without ever having to cut and re-attach branches, then the two trees depict the same relationships—they are really just different views of the same tree. For example, the three trees shown in Figure 7 are identical except that the order of tips has been changed by rotating internal branches.

The flexibility of trees can be illustrated in the classroom in several ways. You can simply have a mobile hanging in the room and point out how you can swing the branches without changing the mobile's structure, whereas cutting a branch and putting it somewhere else would fundamentally change it. Alternatively, you can have students construct simple mobiles from wire, string, paper, and paper clips or construction kits. They can then directly observe what happens to the order of tips when the branches are rotated. Similarly, the Great Clade Race activity (Goldsmith, 2003) gives students practice in seeing when two trees (race courses in the activity) are the same. We have found that computer programs such as MacClade, or its open source descendant Mesquite (Maddison & Maddison, 2006), can be used to give students practice in graphically rearranging trees. Lastly, Julius and Schoenfuss (2006) offer a sophisticated multi-day activity in which students construct their own phylogenetic trees from morphological and genomic data. Such an exercise is likely to lead to improvements in tree-thinking.

Clades & Taxonomy

A clade is a group of organisms that includes a common ancestor and all the descendants of that ancestor. This group of organisms has the property of monophyly (from the Greek for "single clan") and, thus, may also be called a monophyletic group. Systems of classification strive to reflect evolutionary history (see Nickels & Nelson, 2005), which is today achieved by formally naming only groups that are monophyletic. A clade/monophyletic group is easy to identify visually on a tree: It is simply a piece of a larger tree that can be cut away from the root with a single cut (Figure 8A). If one needs to cut the tree in two places to extract a set of taxa (Figure 8B), then that group is non-monophyletic. Biologists sometimes distinguish two different brands of non-monophyly (*polyphyly* and *paraphyly*), but we have not found it useful to draw this distinction in introductory teaching (and, indeed, there is disagreement among specialists as to what they mean). When using tactile models such as mobiles, the

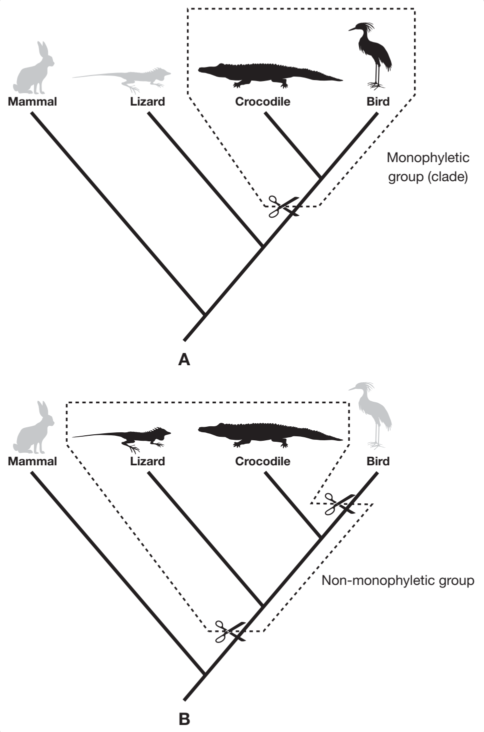

Figure 8. A clade or monophyletic group includes all and only the descendants of a particular ancestor. One way to see if a group of organisms is a clade is to imagine trying to cut those organisms off the tree (away from the root). A. If this can be done with a single cut, the group is monophyletic (a clade). B. If two or more cuts are needed, the group is non-monophyletic.

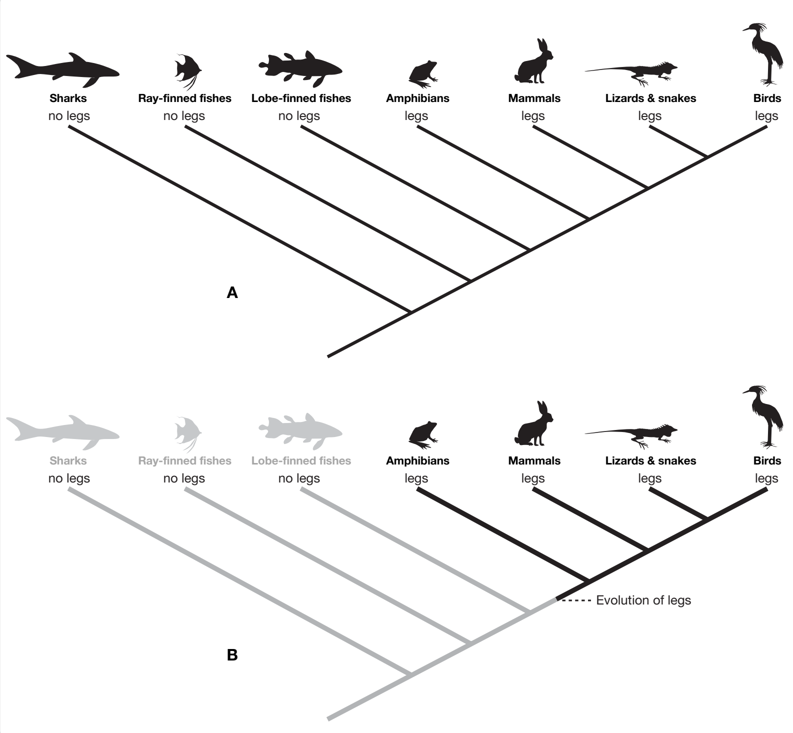

Figure 9. An example showing how trees help us understand evolutionary history. Given the tree shown in A, and the distribution of legs among the included taxa, it is simplest (most parsimonious) to assume that there was a single origin of legs on the branch shown in Figure B. In the absence of additional information, this historical pattern is considered the most plausible explanation for the distribution of legs in the vertebrates.

difference between monophyletic and non-monophyletic groups is easily demonstrated.

Every living species has a unique "clade address," the list of nested clades to which it can be assigned. For example, humans are included in the following nested clades (using the informal names): eukaryotes, animals, deuterostomes, vertebrates, gnathostomes, tetrapods, amniotes, mammals, eutherians, primates, monkeys, apes, and great apes (see Dawkins, 2004 for more information on the ancestry of humans). It can be helpful to point this out to students and see if they can write down the clade address of various familiar organisms. In doing this, it is important to make sure that students understand that the clade address does not depict the descent of one group *from* another, but the nesting of one group *within* another. It is inaccurate to say that humans descended from apes: Because humans are members of the ape clade, it is more accurate to say that we *are* apes (and primates, and animals ...).

Using Phylogenetic Trees To Structure Knowledge of Diversity

One of the main reasons that phylogenetic trees are useful is that they provide a simple way to infer when particular characteristics of a living organism evolved relative to each other. For example, given the tree in Figure 9A, when did legs evolve? Many scenarios are possible, for example, legs could have evolved four times in the four species to the right, or they could have been present in the common ancestor of all seven taxa but have been lost in the three leftmost tips. However, the simplest or *most-parsimonious* explanation, is that there was a single origin of legs, as shown in Figure 9B. The parsimony criterion states that the most plausible mapping of a character onto a tree is that which invokes the fewest changes. While we cannot be sure that the most parsimonious scenario is correct, in the absence of contrary evidence, it is our best hypothesis for the true evolutionary history of a trait.

Presenting a tree and marking the origin of a trait on an internal branch is a very efficient way to show that every descendant of that node should have this trait (unless it was secondarily lost, which could be indicated by a later mark). Any biologically-important trait (for example glycolysis, photosynthesis, meiosis, the nucleus, homeothermy, possession of hair, etc.,) can be mapped either to a single location on the tree of life or, in cases of traits that evolved multiple times, to a few places. In an educational setting this means that major characteristics of living organisms can all be presented within a unified organizational scheme for what can otherwise appear to be an intimidating assemblage of unrelated facts (Offner, 2001). This is becoming easier as knowledge of the tree of life is becoming more accessible in books (for example, Dawkins [2004] and Cracraft & Donoghue [2004]) and Internet sites, most notably the Tree of Life (tolweb.org) and the University of California Museum of Paleontology (ucmp.berkeley.edu).

During every unit, you can take out the appropriate phylogenetic tree and ask students when the trait they are studying arose. This helps make evolution a central organizing theme of the course rather than an isolated unit. We will illustrate the approach by focusing on one example. However, a similar approach can be used to cover all major living groups: animals, mammals, plants, fungi, etc.

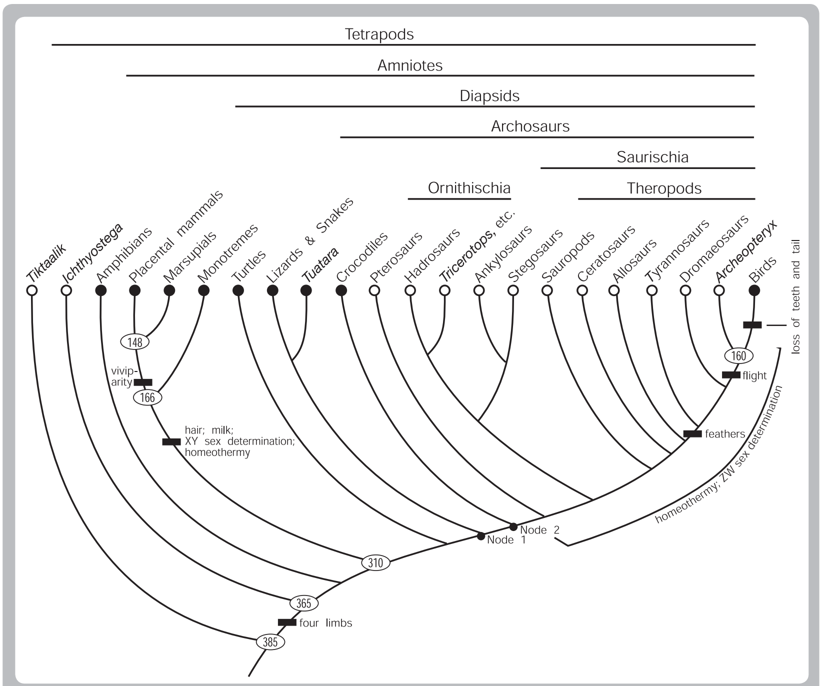

Figure 10. A tree that may be used to explore the evolution of tetrapods. The topology of this tree is based on assorted sources including Lee et al. (2004), Daeschler et al. (2006), and the Tree of Life Web site (www.tolweb.org). Extinct taxa are marked with lineages that end in an open circle. Approximate ages of certain nodes are shown in millions of years before present (Ma) and the composition of some named clades are shown with lines above the tree. Selected traits are mapped to the branches on which they are thought to have arisen. Two traits of birds (homeothermy and ZW sex determination) are not easily scored for fossils, hence there is uncertainty as to which among the bracketed branches is the correct place to mark their evolution.

The Tetrapod Tree Is a Useful Teaching Tool

The tetrapod tree in Figure 10 can be used to illustrate many of the principles discussed in this article and also to show students how trees become richer as our knowledge of biology increases. This tree shows our most current understanding of the evolutionary relationships of tetrapods. The dates on the trees, which come from a combination of the fossil record and statistical analysis of the rate of molecular evolution, give students some glimpse of evolutionary time (though they will have to work hard to grasp the true magnitude of the numbers).

This tree can be used to illustrate principles such as relatedness. In this regard it is helpful to focus on a case that highlights the distinction between the concepts of relatedness and similarity. Referring to Figure 10, students can be asked whether a crocodile is more *similar* to a lizard or a bird, and, at the same time, which it is more closely *related* to. Students will easily see that a crocodile has more external similarities to a lizard than to a bird, but may find it counter-intuitive that crocodiles are more closely related to birds than to lizards. This is best explored by numbering the internal nodes in this part of the tree (Figure 10) and identifying the last common ancestor a crocodile shares with a lizard (Node 1) and a bird (Node 2). The question is then, which of these two nodes (representing common ancestors) lived closer to the present? It should become clear that because Node 2 is more distant from the root and, hence, a descendant of Node 1, it must have lived more recently. As discussed earlier, the addition of arrows pointing away from the root might help some students. Because the crocodile had a common ancestor with a bird more recently than with a lizard, it is more closely related to the bird, even though it looks more like the lizard. Relatedness, by definition, is based on how recently two organisms had a common ancestor, not by how much they look alike.

The inference that crocodiles are more closely related to birds than to lizards is explained by the fact that the similarities of crocodiles and lizards, such as the sprawling gait, elongated tail, and scales, are features that trace back to the common ancestors of lizards, crocodiles, *and* birds (and mammals too, actually). The lizard and crocodile lineages have both retained this ancestral body form whereas birds underwent dramatic evolutionary divergence, acquiring many novel features: feathers, flight, a wishbone, a keeled sternum, reduced forelimb digits, loss of a bony tail, and loss of teeth. Although crocodiles and birds do share some evolutionary-derived anatomical traits, these are less obvious to a casual observer than the shared ancestral similarities of crocodiles and lizards. Thus, despite the many visible differences between birds and crocodiles, a crocodile is more closely related to a bird than to a lizard. Relatedness is about descent not similarity.

One of us (SO) gives the full tetrapod tree to the students after teaching them the characteristics of the major groups of vertebrates. These characteristics are taught in a very traditional way. After all, knowledge of the characteristics of the groups of vertebrates has not changed significantly with the improved understanding of the evolutionary origins of those characteristics. A point that arises is that, despite the fact that birds and mammals are both endotherms with four chambered hearts and insulating skin coverings (feathers and fur, both of which are made of keratin and are modified scales), birds and mammals are *not* especially closely related. Indeed, as you can see from the tree, birds and mammals are each more closely related to ectothermic organisms than they are to each other. This is more obvious for birds because they are closely related to living ectothermic species (crocodiles, lizards, snakes, etc.). However, extinct fossil species are known that are more closely related to mammals than to birds and are inferred to have been ectothermic (e.g., pelycosaurs). Therefore, the similarities between birds and mammals arose independently, making examples of convergent evolution.

The story of how specific groups of organisms originated and acquired their distinctive traits makes for interesting story telling, while also allowing for the introduction of diverse biological principles. For example, the evolution of birds from terrestrial dinosaurs always engages students at both the college and high school levels. Specifically, it has become clear that modern birds are part of a clade that included a number of fast-running, bipedal, meat-eating dinosaurs such as *Tyrannosaurus* and *Velociraptor*. These dinosaurs had feathers and are thought to have been homeothermic. The high metabolic needs of these animals probably explains the fact that their respiratory surfaces (like those of birds) extend into the cavities of hollow bones. This in turn resulted in light bones that would likely have been advantageous for running speed. Combined with the elongated forelimbs of coelurosaurs (these are thought to have allowed them to grab and hold prey), students may come to visualize that a transition to eating flying insects might have favored the evolution of leaping with the aid of feathered forelimbs and, ultimately, to powered flight (Garner et al., 1999). By working through this example, students learn not only more about birds, but also come to appreciate how scientists construct and test adaptive hypotheses.

A wonderful way to keep students interested in research is to ask when different adaptations of birds evolved. For example, students will be amazed and inspired to learn that *Tyrannosaurus rex* had feathers (Xu et. al., 2004), putting the origin of feathers well back before the origin of flight. Other questions might come up that are harder to answer, such as when homeothermy and ZW sex determination evolved in the ancestry of birds. While you will be unable to answer these questions, asking students to ponder how we might ultimately answer them encourages them to think like an evolutionary biologist.

Trees Change as Our Knowledge of Biology Increases

Often when one is teaching, new discoveries or analyses are reported in the popular media and these should be exploited as an opportunity to expose students to the ways that scientists probe the past. For example, in April 2006, the journal *Nature* (Daeschler et.al., 2006, and Shubin et. al., 2006) reported a beautifully preserved fossil named *Tiktaalik*, an animal with front appendages like tetrapod limbs, and rear fins like fish. *Tiktaalik* lived a little over 380 million years ago, and its structure fills in a gap between that of the fully aquatic fish and the first known tetrapod, *Ichthyostega*. Coverage of this find extended into the popular media and included articles written at a level that students could follow. What was particularly exciting was being able to add this group to the tetrapod tree. Students could then see how the tree provides a predictive framework for the existence of other, as yet undiscovered, fossil forms.

At the same time, it can be powerful to explore cases in which there is uncertainty as to the correct phylogenetic tree and/or cases in which views on the correct tree have changed over time. For example, the position of turtles on the tetrapod tree has been a source of controversy. Some early molecular analyses placed the turtles close to crocodiles, which disagreed with studies of anatomy that suggested that turtles are an isolated tetrapod lineage. At this point in time, the preponderance of evidence points to the resolution shown in Figure 10, but this is far from definitive and could change as new genomic data are collected and analyzed.

The Challenge of Teaching Tree-Thinking

Tree-thinking is important because it embodies a clear understanding of the principle of descent from common ancestry, because it helps students organize knowledge of biological diversity, and because phylogenetic trees are playing a more central role in scientific discourse. Indeed, a case can be made that understanding tree-thinking is at least as important to evolutionary biology instruction as understanding natural selection (O'Hara, 1997). Nonetheless, because phylogenetics is a fairly young field, there has been little educational research on tree-thinking, and textbooks generally overlook trees or cover them only superficially and/or inaccurately. Fortunately, there is growing awareness of the importance of this subject, and some creative approaches to teaching tree-thinking have begun to be developed (e.g., Goldsmith, 2003; Gilbert, 2003; and others posted at www.tree-thinking.org). Hopefully, in the coming few years there will be development and assessment of additional instructional materials for teaching tree-thinking and this will result in long-term gains in students' understanding of evolution and biological diversity.

Acknowledgments

Comments on drafts of this article were kindly provided by Margaret Koopman and six anonymous reviewers. Artwork was expertly prepared by Kandis Elliott. Funding was provided by the National Science Foundation (DEB-0416096).

References

Baum, D. A., DeWitt Smith, S. & Donovan, S. S. (2005). The tree thinking challenge. Sci*ence, 310*, 979-980.

Cracraft, J. & Donoghue, M. J. (Editors). (2004). *Assembling the Tree of Life*. Oxford, UK: Oxford University Press.

Daeschler, E. B., Shubin, N. H. & Jenkins, F. A., Jr. (2006). A Devonian tetrapod-like fish and the evolution of the tetrapod body plan. *Nature*, *440,* 757-763.

Dawkins, R. (2004). *The Ancestor's Tale: A Pilgrimage to the Dawn of Life*. New York, NY: Houghton Mifflin.

Garner, J. P., Taylor, G. K. & Thomas, A. L. R. (1999). On the origin of birds: The sequence of character acquisition in the evolution of avian flight. *Proceedings of the Royal Society of London Series B – Biological Sciences, 266*, 1259-1266.

Gilbert, S. F. (2003). Opening Darwin's black box: Teaching evolution through developmental genetics. *Nature Reviews Genetics, 4*, 735-741.

Goldsmith, D. W. (2003). The great clade race. *The American Biology Teacher*, *65*(9), 679-682.

Julius, M.L. & Schoenfuss, H.L. (2006). Phylogenetic reconstruction as a broadly applicable teaching tool in the biology classroom. *Journal of College Science Teaching, 35,* 41-45.

Lee, M. S. Y., Reeder, T. W., Slowinski, J. B. & Lawson, R. (2004). Resolving reptiles relationships: Molecular and morphological markers. In J. Cracraft and M. J. Donoghue (Eds.), *Assembling the Tree of Life* (pp. 451-467). Oxford, UK: Oxford University Press.

Maddison, W. P. & Maddison, D.R. (2006). Mesquite: a modular system for evolutionary analysis, version 1.12. Available online at: http://mesquiteproject.org.

Nickels, M. K. & Nelson, C. E. (2005). Beware of nuts and bolts: Putting evolution into the teaching of biological classification. *The American Biology Teacher, 67*(5), 283-289.

Offner, S. (2001). A universal phylogenetic tree. *The American Biology Teacher*, *63*(3), 164-170.

O'Hara, R. J. (1988). Homage to Clio, or, toward an historical philosophy for evolutionary biology. *Systematic Zoology*, *37*, 142-155.

O'Hara, R. J. (1997). Population thinking and tree thinking in systematics. *Zoologica Scripta, 26*, 323-329.

Shubin, N. H., Daeschler, E. B. & Jenkins, F. A., Jr. (2006) The pectoral fin of *Tiktaalik roseae* and the origin of the tetrapod limb. *Nature, 440,* 764-771.

Xu, X., Morell, M.A., Kuang, X., Wang, X., Zhao, Q. & Jia, C. (2004). Basal tyrannosauroids from China and evidence for protofeathers in tyrannosauroids. *Nature, 431,* 680-684.

Appendix: Phylogenetic Inference

The actual methods of tree construction are beyond the scope of this article and, indeed, would take another entire article. Fortunately, one can develop a solid understanding of what a phylogenetic tree represents without knowing much about how scientists actually infer the structure of those trees. Nonetheless, a few words of explanation may be helpful. There are sophisticated computer programs that estimate phylogenetic trees of organisms based on morphological data, molecular data (most commonly, DNA sequences), or a combination of different types of data. They use a variety of statistical criteria for selecting the optimal tree, with four main approaches.

1. Distance methods favor trees on which the difference between each pair of tips is most similar to the evolutionary distance between those two tips (measured along the tree's branches).
2. Parsimony methods favor trees that minimize the number of evolutionary changes needed to explain the traits of the tips.
3. Likelihood methods favor trees that, if they were true, would be most likely to have generated the observed data (given a particular statistical model of how traits evolve).
4. Bayesian methods favor trees that are most likely to be true given the data, a model of evolution, and our prior beliefs about the evolutionary process.

In most cases computer programs sort among the set of possible trees (of which there may be very many) to find one or a subset that are favored by the particular optimality criterion. Methods are also available for assessing how much more strongly the data favor one set of trees over competing trees. When a clade is strongly supported by multiple, independent datasets, it comes to be accepted as a well-supported hypothesis of evolutionary relationships. Generally, results are stronger with more data. This is why the rapid growth in DNA sequence data over the last few decades has greatly improved our understanding of the tree of life. Nonetheless, some phylogenetic problems (such as those that involve multiple branching events that occurred in a short window of time in the distant past) are very difficult to solve even with abundant data. For example, the relationships among the protostome phyla or the classes of arthropods continue to be points of contention. Additionally, resolving the phylogenetic placement of fossil forms, which are often quite fragmentary, can be problematic.