Evolutionary Biology
Entry ID:
Citation Style: Scientific
Version date:

INTERPRETING PHYLOGENETIC TREES

Stacey D. Smith
University of Colorado-Boulder

INTRODUCTION

At first glance, a phylogenetic tree appears to be a relatively simple diagram, beginning at some single starting point and giving rise to a set of tips through a series of branching events. Trees have long served as models for relating different species, even before the advent of Darwin's theory of evolution. However, correctly interpreting phylogenetic trees can be challenging, especially given the pervasiveness of alternative depictions of the history of life (e.g., the scala naturae or "ladder of life"). The majority of biologists today interpret phylogenetic trees as representations of the branching evolutionary history of the organisms at the tips. Because of the relevance of evolutionary history to understanding many biological patterns (e.g., in ecology, biogeography, and epidemiology), developing a clear understanding of how to read and interpret phylogenetic trees has become paramount for all students of biology. This bibliography will cover the core literature related to understanding the information communicated by phylogenetic trees, and to some extent, how this information is applied to address broader questions. Related topics, such as phylogenetic inference and the use of phylogenies to create classifications, are covered by other bibliographies.

BOOKS

Although there are quite a few texts that cover phylogenetic inference, relatively few deal explicitly with interpreting phylogenetic trees or applying them to address broader evolutionary questions. In this section, I list some of the classic books that cover areas of phylogenetics beyond tree inference itself. Of these, the Baum and Smith (2013) text provides the most in-depth introduction to reading phylogenetic trees while the others discuss some of the ways that phylogenies are important for addressing questions about evolutionary patterns and processes.

Eldredge, N., and J. Cracraft. 1980. *Phylogenetic patterns and the evolutionary process*. New York : Columbia University Press.

The authors cover the basics of cladistics, but also devote special attention to how phylogenetic analysis can be applied to addressing macroevolutionary questions, such the role of adaptation to new niches in diversification. Advances in statistical methods in recent years have resulted in renewed interest in applying phylogenies to testing such questions.

Brooks, D. R., and D. A. McLennan. 1991. *Phylogeny, ecology, and behavior: A research program in comparative biology*. Chicago: University of Chicago Press.

This book is mainly focused on interpreting patterns of speciation and adaptation in a parsimony framework and includes a large number of fascinating biological examples. It is important to note that many of the trees are based on morphological characters, which are not commonly used today for phylogenetic studies of adaptation because of the potential for circularity.

Harvey, P. H., and M. D. Pagel. 1991. *The comparative method in evolutionary biology*. Oxford: Oxford University Press.

A classic book in comparative biology that lays out the need for incorporating phylogenetic history in any analysis that spans multiple taxa. It describes how phylogenetic approaches can be used to test adaptive hypotheses using both discrete and continuous data.

Wiley, E. O., and B. S. Liberman. 2011. *Phylogenetics: The theory and practice of phylogenetic systematics* (2nd edition). Hoboken, New Jersey: Wiley & Sons.

An updated version of Wiley's 1981 book, this text covers many concepts relevant to interpreting trees. For example, they discuss different kinds of trees and different ways that character evolution is represented graphically on trees.

Baum, D. A., and S. D. Smith. 2013. *Tree-thinking: An introduction to phylogenetic biology*. Greenwood Village, Colorado: Roberts & Co.

This textbook provides a broad introduction to tree-thinking, with several chapters devoted to interpreting evolutionary relatedness and patterns of trait evolution using phylogenies. It is written to be accessible to undergraduates and biologists in fields outside of phylogenetics.

JOURNALS

Phylogenetic trees are relevant to many areas of biological sciences, and thus appear in a wide range of scientific journals. Only a handful of journals (e.g., *Trends in Ecology and Evolution* and *Cladistics*), however, regularly contain articles discussing how to interpret phylogenetic trees. Indeed, many of the relevant articles are found in journals aimed at teachers, e.g., *American Biology Teacher* and *Evolution: Education and Outreach*.

**American Biology Teacher* [http://www.nabt.org]*

This is the journal of the National Association for Biology Teachers. Its content is targeted at teachers of K-16, but the journal is also useful to students of evolutionary biology or those teaching introductory undergraduate courses. The articles include assessments of teaching strategies, reviews, and activities designed for the classroom.

**Bioessays* [http://www.wiley.com/WileyCDA/WileyTitle/productCd-BIES.html]*

This journal contains reviews and discussions in all areas of biology. Its articles are written for a broad audience and often cover topics in phylogenetics, including how to read trees and interpret patterns of trait evolution.

**Cladistics* [http://cladistics.org]*

Cladistics, the journal of the Willi Hennig Society, focuses on systematics and regularly has papers discussing tree interpretation, terminology, and applications to taxonomy.

**Evolution* [http://onlinelibrary.wiley.com/journal/10.1111/(ISSN)1558-5646]*

This journal is sponsored by the Society for the Study of Evolution, and is the premier journal for research in evolutionary biology. It covers a wide range of topics related to phylogenetics and comparative biology, including divergence time estimation, trait evolution, tests of adaptive hypotheses, biogeography, and diversification analyses.

**Evolution: Education and Outreach*

[http://www.springer.com/life+sciences/evolutionary+%26+developmental+biology/journal/12052]*

This journal targets a similar audience as American Biology Teacher, but is specifically focused on the teaching of evolution. It contains many articles related to teaching tree-thinking and clarifying misconceptions about phylogenetics and evolution.

**Systematic Biology* (formerly *Systematic Zoology*) [http://sysbio.oxfordjournals.org/]*

This journal of the Society of Systematic Biologists is the principal source for new statistical methods in phylogenetics and comparative methods. Most papers are targeted at specialists in the field, but the journal contains many classic papers in phylogenetic biology that are accessible for a general audience.

**Trends in Ecology and Evolution* [http://www.cell.com/trends/ecology-evolution/]*

TREE provides short opinion articles and review s about current topics in ecology and evolutionary biology. The articles are easy-to-read and often cover subjects related to phylogenetics and evolution.

TREES AS A MODEL FOR EVOLUTION

The preponderance of trees in modern biology is largely thanks to Charles Darwin, who established phylogenetic trees and the common ancestry they imply as a core component of the theory of evolution. Nonetheless, the use of a branching tree-like diagram to portray the history of life predates Charles Darwin, and indeed the metaphor can be found in the writings of his grandfather, Erasmus Darwin. Recent studies have revealed abundant evidence for the tree-like structure of evolutionary history (see **TREES AS EVIDENCE FOR EVOLUTION***), effectively rejecting other models of macroevolution, such as separate ancestry (e.g. spontaneous generation) or a scala naturae (ladder-of-life). The articles below describe the fascinating and sometimes controversial history of trees as a model for evolution. Stevens (1983), Archibald (2009), Ragan (2009), and Tassy (2011) discuss the early tree diagrams in the literature while Bowler (2003), Gould (1997) and Nee (2005) write about the tree of life versus the ladder of life as metaphors for evolutionary history. Mindell (2013) discusses the importance of tree as a metaphor and as a model in evolutionary biology.

Stevens, P. F. 1983. Augustin Augier's "Arbre Botanique" (1801), a remarkable early botanical representation of the natural system. *Taxon* 32:203-211.

Stevens describes the work of the French botanist Augier who used a 'botanical tree' to show the relationships among plant families and devise a natural system of classification, over 50 years before the *Origin of Species*.

Gould, S. J. 1997. Redrafting the tree of life. *Proceedings of the American Philosophical Society* 141:30-54.

An insightful essay discussing the "ladder of progress" and the "tree of life" as depictions of evolutionary history. Although Gould notes that a phylogeny is not well suited for portraying morphological disparity, often appearing as a "cone of increasing diversity."

Bowler, P. J. 2003. *Evolution: the history of an idea*. University of California Press: Berkeley, CA.

This historical book details the rise of Darwinism and the subsequent efforts to reconstruct the tree of life. The spread of tree-thinking lead to inherent conflicts with progressive or "ladder-like" views of evolutionary change.

Nee, S. 2005. The great chain of being. *Nature* 435: 429.

Nee discusses the psychological drive to describe evolution as a ladder of progress, with humans at the apex. Despite the recognition that the ladder runs counter to modern evolutionary theory, phylogenies are often drawn in a ladder-like fashion that allows the perception of an apex to persist.

Archibald, J. D. 2009. Edward Hitchcock's Pre-Darwinian (1840) "Tree of Life". *Journal of the History of Biology* 42:561-592.

Archibald reviews the use of tree diagrams to portray relationships among species, beginning with Augier's 1801 tree of plant families to those presented by Darwin and Haeckel in the mid-1800's. The rest of the paper discusses the paleontological trees published by the geologist E. Hitchcock, who nonetheless rejected evolutionary theory.

Ragan, M. A. 2009. Trees and networks before and after Darwin. *Biology Direct* 4:43.

Ragan describes the history of trees and networks as representations of relationships among species, from the 1700's to present day.

Tassy, P. 2011. Trees before and after Darwin. *Journal of Zoological Systematics and Evolutionary Research* 49:89-101.

This paper considers how the first trees of species were used for the purpose of classification, while those published after Darwin began to incorporate a time dimension and explicitly portray evolution.

Mindell, D. 2013. The tree of life: metaphor, model, and heuristic device. *Systematic Biology* 62: 479-489.

The author discusses the implications of reticulate processes (introgression, horizontal transfer) for the tree of life as a model for evolutionary history. He argues that the tree retains important heuristic utility in evolutionary biology, even though it overly simplifies the true nature of history (i.e. a mix of divergent and reticulate processes).

TREE TERMINOLOGY

Communicating the information portrayed in phylogenetic trees requires a vocabulary for discussing the parts of trees. As the field of phylogenetics, and statistical phylogenetics in particular, has grown, so has

the terminology for describing trees. Many of these terms, such as root, branch, and leaf, are borrowed from botany. Since phylogenies are essentially graphs (specifically directed acyclic graphs), terms derived from mathematics and network literature (nodes, edges, splits) are also common. The choice of terms when there are multiple equivalent descriptors (e.g. leaf, tip, and taxa) depends on the context. For example, the term "leaf" is often used in statistical literature where there is less potential confusion with leaves in the botany sense, while tips and taxa are more commonly used in systematics literature. This section provides both broad overviews of tree terminology as well as papers focused on particular concepts.

General overviews

Much of the terminology relating to phylogenetics was codified by Hennig's 1966 Phylogenetic Systematics. Mayr (1978), Mayden and Wiley (1994), and Wheeler (2012) give helpful summaries of while Keller and Lloyd (1992) give more conceptual and philosophical viewpoints. Semple and Steel (2003) is a useful complement to these texts as a mathematical perspective.

Hennig, W. 1966. *Phylogenetic systematics*. Translated by D. D. Davis and R. Zangerl. Urbana, IL: University of Illinois Press.

> This is a foundational book in the field and is credited with spurring the tremendous growth in systematics in subsequent decades. Hennig defined many new terms and concepts in this book, including those relating to characters (e.g. apomorphies and plesiomorphies) as well as groups of taxa (e.g. paraphyletic groups).

Mayr, E. 1978. Origin and history of some terms in systematic and evolutionary biology. *Systematic Zoology* 27: 83-88.

> Mayr reviews some commonly used terms and describes when they were first proposed. These include cladogram, dendrogram, phenogram and phylogram,

Mayden, R. L., and E. O. Wiley. 1992. "The fundamentals of phylogenetic systematics." In *Systematics, historical ecology, and North American freshwater fishes*, Edited by R. L. Mayden, 114-185. Stanford, CA: Stanford University Press.

> This overview of phylogenetic systematics has a large vocabulary section, which covers trees, taxa, and characters.

Keller, E. F., and E. A. Lloyd, eds. 1992. *Keywords in evolutionary biology*. Cambridge, MA: Harvard University Press.

> An excellent collection of essays about common terms used in evolutionary biology. Several of the essays are relevant to tree thinking including those on the terms "character", "monophyly," and "species".

Semple, C., and M. Steel. 2003. *Phylogenetics*. New York, NY: Oxford University Press.

> A phylogenetics text targeted at computer scientists, mathematicians and statisticians. This book provides a starting point for becoming acquainted with the mathematical tree terminology, such as cherry, edge, vertex, and forest.

Wheeler, W. C. 2012. *Systematics: A course of lectures*. Oxford, UK: Wiley-Blackwell.

> This book is structured as a series of lectures at the introductory graduate level and has a chapter covering fundamental terms and concepts (characters, taxa, trees, etc.)

Characters and character states

The concept of a character is central to phylogenetics: characters are the heritable features that are compared across species to infer relationships. The first phylogenetic analyses relied on morphotypic characters, such as morphology, behavior, or chemistry, which required delimitation of character states before scoring and analysis. The rise of molecular phylogenetics has reduced the focus on delimiting states (as the states can only be one of four nucleotides, A, C, G, T). Nonetheless, choosing to treat characters as discrete or continuous and deciding how to divide a character into character states is still extremely important for applying phylogenetic comparative methods. For example, a study of the evolution of feather color across the phylogeny of birds would require first determining whether color variation was continuous or discrete, and if the latter, how many different color states to recognize. Colless (1985), Pimentel and Riggins (1987), Thiele (1993), and Harris and Mishler (2009) discuss the meaning of characters and character states and describe the challenges of delimiting states. Nixon and Wheeler (1990) suggest that it is important to distinguish between traits and characters in phylogenetics.

Colless, D. 1985. On "character" and related terms. *Systematic Zoology* 34: 229-233.

Colless discusses the various ways in which biologists use the term character and describe character states. He supports maintaining a clear distinction between characters and character states and suggests a uniform format for naming them.

Pimentel, R. A., and R. Riggins. 1987. The nature of cladistic data. *Cladistics* 3: 201-209.

The authors argue that only discrete characters with two or more mutually exclusive states are suitable for use in cladistics analysis. Others (e.g. Thiele, 1993) conclude that the division between discrete and continuous variables is not so clear and that data should not be considered *a priori* to be of less use for inferring relationships.

Nixon, K. C., and Q. D. Wheeler. 1990. An amplification of the phylogenetic species concept. *Cladistics* 6: 211-223.

As part of developing their species concept, Nixon and Wheeler distinguish between characters and traits, with characters being features that are found in all individuals of a putative species.

Thiele, K. 1993. The holy grail of the perfect character: the cladistics treatment of morphometric data. *Cladistics* 9: 275-304.

A useful paper that describes different types of data and how to represent this variation as character states for phylogenetic analysis.

Harris, E. S. J., and B. D. Mishler. 2009. The delimitation of phylogenetic characters. *Biological Theory* 4: 230-234.

A very accessible paper that summarizes much of the above literature and emphasizes the fact that, in phylogenetics, characters always exist relative to some purpose (e.g. distinguishing different species of spiders).

Groups

One of the most basic elements of tree interpretation is identifying natural groups of taxa, i.e. monophyletic groups or clades. A monophyletic group is a branch of the tree, containing an ancestor and all of its descendants. The list below provides some of the classic references for this and related terms. Baum (2008) gives an accessible introduction to monophyly, and Ashlock (1971), Nelson (1971), Farris (1974) discuss different ways to define non-monophyletic groups (paraphyletic and polyphyletic groups). The Huxley (1957) argues for the evolutionary significance of grades (paraphyletic groups); this and similar papers may be one reason for the drive to recognize non-monophyletic groups in classification. De Queiroz (1988) rejects the recognition of any groups of other than natural (monophyletic) groups.

Huxley, J. 1957. The three types of evolutionary process. *Nature* 180: 454-455.
In this paper, Huxley defines three major evolutionary processes (anagenesis, cladogenesis, and stasigenesis) and their products (clades and grades). The emphasis on "improvement" and the recognition of grades, however, reflects the persistence of ladder-of-life thinking in evolutionary biology (see **MISCONCEPTIONS***).

Ashlock, P. D. 1971. Monophyly and associated terms. *Systematic Zoology* 20: 63-69.
This paper aimss to clarify the meanings of monophyly, paraphyly, and polyphyly. Seeking something of a middle ground that would allow the recognition of paraphyletic groups in classification, Ashlock proposed that paraphyly be considered a type of monophyly and proposed a new term, holophyly, to correspond to Hennig's strict monophyly.

Nelson, G. J. 1971. Paraphyly and polyphyly: redefinitions. *Systematic Zoology* 20: 471-472.
Nelson offered an alternative solution to the ambiguity of Hennig's definitions, namely to distinguish paraphyly and polyphyly by the number of descendants of the common ancestor that were not included in the group.

Farris, J. S. 1974. Formal definitions of paraphyly and polyphyly. *Systematic Zoology* 23: 548-554.
Farris points out that Hennig defined monophyly by common ancestry while paraphyly and polyphyly were defined by characters, created a problem in comparing the three types of groups. He proposes a parsimony-based method for determining if previously named group (e.g., "Mammals") constitutes a monophyletic, paraphyletic or polyphyletic group.

De Queiroz, K. 1988. Systematics and the Darwinian revolution. *Philosophy of Science* 55: 238-259.
The author discusses the meaning of the terms including relationship, synapomorphy, taxa, and monophyly, and he argues that only phylogenetic relationships (as opposed to phenotypic similarity or grades) should be used to create classifications.

Baum, D. 2008. Reading a phylogenetic tree: The meaning of monophyletic groups. *Nature Education* 1:1.
This paper is written for a broad audience and explains how to identify monophyletic groups as well as other basic principles of tree interpretation.

Trees

Phylogenetic trees can appear in several different forms that are intended to convey different information. For example, cladograms communicate only the tree topology, i.e. relationships among the tips, but not the time interval between branching events. Phylograms, by contrast, convey both topological information and branch lengths (the inferred amount of evolutionary change, e.g., number of substitutions per site in a DNA sequence). Chronograms also portray branch length information but unlike phylograms, the unit of change is time. Each of these dendrograms can be drawn in a several ways, with the most common being rectangular, slanted, and circular. Relatively few papers have directly compared the different types of trees and their effectiveness in communicating phylogenetic information, although the ever-increasing size of phylogenetic trees (often now with thousands of tips) is driving improvements in tree visualization. Hendy an.d Penny (1984) describe what trees are as a mathematical construct and in and evolutionary perspective. Catley and Novick (2008), Pavlopoulos et al (2010), and MacDonald and Wiley (2012) survey different types of tree diagrams in the literature, textbooks and museums. Avise (2009) focuses on trees that include a time component, and Page (2012) expand to trees that link to geography.

Hendy, M. D., and D. Penny. 1984. Cladograms should be called trees. *Systematic Zoology* 33:245-247.
Hendy and Penny argue that cladograms are technically trees, even if some authors prefer not to interpret them as evolutionary trees. They briefly review the definition of trees from the standpoint of mathematics and graph theory.

Catley, K. M., and L. R. Novick. 2008. Seeing the wood for the trees: An analysis of evolutionary diagrams in biology textbooks. *BioScience* 58: 976-987.
The authors document the frequency of different tree representations in textbooks. They describe how professional journals tend to use rectangular cladograms (which they call "tree" format) while textbooks often show slanted cladograms ("ladder" format), which may contribute to the ladder-of-life misconceptions about evolution.

Avise, J. C. 2009. Timetrees: beyond cladograms, phenograms, and phylograms. In *The Timetree of Life*, Edited by S. B. Hedges and S. Kumar, 19–25. Oxford: Oxford Oxford University Press.
This book chapter discusses the increasing interest in chronograms ('timetrees') in evolutionary biology, and possible causes for the lag in inferring chronograms relative to cladograms and phylograms (e.g., concerns about the molecular clock).

Pavlopoulos, G. A., T. G. Soldatos, A. Barbosa-Silva, and R. Schneider. 2010. A reference guide for tree analysis and visualization. *BioData Mining* 3:1.
This article reviews the different tree formats and includes a summary of software available for tree visualization.

Page, R. D. M. 2012. Space, time, form: viewing the Tree of Life. *Trends in Ecology and Evolution* 27: 113-120.
Page describes the challenges of creating viewing and navigating trees with thousands of taxa and other areas of interest in tree visualization, such as portraying uncertainty in phylogenies and plotting tanglegrams of host-parasite trees.

MacDonald, T., and E. O. Wiley. 2012. Communicating phylogeny: Evolutionary tree diagrams in museums. *Evolution: Education and Outreach* 5:14-28.
The authors survey the different types of tree diagrams used in museum, and describe the different diagrammatic elements of trees (orientation, geometry, etc.). They include a useful table of tree types (slanted cladograms, rectangular cladograms, curvograms, circular trees, radial trees, etc.).

READING RELATIONSHIPS FROM TREES

The most fundamental piece of information communicated by a phylogenetic tree is the order of branching events that gave rise to the sampled taxa. The process of identifying these events and correctly translating them into evolutionary relationships is often referred to as "reading" a tree. More broadly, interpreting phylogenetic trees and using them to understand evolutionary history is termed "tree-thinking" (O'Hara, 1997). Baum et al. (2005) and Baum and Offner (2008) provide accessible introductions to how to read trees, and Gregory (2008) and Omland (2013) also discuss what *not* to infer from trees.

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

O'Hara traces the history of different schools of thought in systematics and contrasts 'development thinking', which paints evolution as a linear narrative, with 'tree thinking', which emphasizes the branching tree-like history of life.

Baum, D. A., S. D. Smith, and S. S. Donovan. 2005. The tree-thinking challenge. *Science* 310: 979–980.
This article describes how to read phylogenetic trees and discusses the meaning of relatedness in an evolutionary context. The authors provide a quiz in the supplementary material to test basic tree-thinking skills.

Baum, D. A., and S. Offner. 2008. Phylogenies and tree-thinking. *American Biology Teacher* 70: 222-229.
Targeted for teachers of introductory biology, this article reviews tree interpretation, including how individuals and populations relate to lineages on a phylogeny and how to read relationships from trees.

Gregory, T. R. 2008. Understanding evolutionary trees. *Evolution: Education, and Outreach* 1: 121-137.
The author argues for the importance of phylogenetic literacy, and discusses how to read trees as well as how *not* to read trees (see also **MISCONCEPTIONS***).

Omland, K. 2013. Interpretation of phylogenetic trees. Pp. 51-59 in J. Losos, D. Baum, and D. Futuyma, eds., *The Princeton guide to evolution*. Princeton University Press: Princeton, NJ.
Omland reviews tree terminology, reading trees, and misconceptions of trees as ladders of progress.

MISCONCEPTIONS

Misinterpretations of phylogenetic trees abound in both popular culture and in the primary literature, even evolutionary biology journals. For example, trees are often used to assert that some extant taxa are "basal" or "primitive" relative to others. There is small but growing body of literature aimed at dispelling misconceptions about phylogenies and evolutionary history. Hodos and Campbell (1969) describe how misconceptions about evolution lead to poor experimental design in comparative psychology, while Meir et al. (2007) and Omland et al. (2008) discuss problems relating to tree thinking more broadly. Krell and Cranston (2004), Crisp and Cook (2005) and Rigato and Minelli (2013) tackle the issue of how misconceptions about evolution trickle into the language used to describe trees and relationships, and Sandvik (2009) writes about how anthropocentric world views can shape tree depiction. Finally, Casane and Laurenti (2013) address the misconception that some extant taxa are "living fossils", i.e. have ceased evolving.

Hodos, W., and C. B. Campbell. 1969. Scala naturae: Why there is no theory in comparative psychology. *Psychological Review* 76: 337-350.
Hodos and Campbell describe the persistence of ladder-of-life thinking in studies of behavioral evolution, and they argue that subjective ranking species as higher or more advancement in comparative studies is misleading and counterproductive.

Krell, F. T., and P. S. Cranston. 2004. Which side of the tree is more basal? *Systematic Entomology* 29: 279-281.
The authors clarify that, although there are basal (earlier) nodes in a tree, there are no basal clades or taxa because each of the sister groups arising from any node is by definition of equal age. This incorrect phrase can be easily avoided by using sister group terminology instead.

Crisp, M. D., and L. G. Cook. 2005. Do early branching lineages signify ancestral traits? *Trends in Ecology and Evolution* 20:122-128.

An excellent article that aims to debunk the association of species-poor sister groups (often termed "basal" or "early-diverging" lineages) with ancestral character states. They point out that inference of ancestral states should be carried out objectively using one of the many available methods for trait reconstruction.

Meir, E, J. Perry, J. C. Herron, and J. Kingsolver. 2007. College students' misconceptions about evolutionary trees. *American Biology Teacher* 69:71–76.

Meir et al. review four common misconceptions about trees relating to the flow of time on the diagram, tip order, node counting, and drawing format.

Omland, K., L. G. Cook, and M. D. Crisp. 2008. Tree thinking for all biology: the problem with reading phylogenies as ladders of progress. *Bioessays* 30:854-867.

The authors address the misconception that phylogenies represent ladders of progress. They explain this problem is related to what they call the "primitive lineage fallacy", that species-poor lineages are "basal" or represent ancestors of other species. They discuss the implications of such misinterpretation of trees for research in evolutionary biology.

Sandvik, H. 2009. Anthropocentrism in cladograms. *Biology & Philosophy* 24: 425-440.

Sandvik describes how the ordering of tips across a tree can lead to an anthropocentric view of evolution, and devises a metric for measuring the anthropocentrism in tree diagrams.

Casane, D., and P. Laurenti. 2013. Why coelacanths are not 'living fossils'. *Bioessays* 35: 332-338.

The authors point out that the concept of a living fossil is at odds with tree-thinking and use data on the evolution of coelacanths as case in point.

Rigato, E., and A. Minelli. 2013. The great chain of being is still here. *Evolution: Education and Outreach* 6:18.

Rigato and Minelli examine recent literature for evidence of progressive misconceptions about evolution. They find that ladder-of-life thinking and terminology (e.g., "higher" and "lower" taxa) remains common, particularly (and troublingly) in evolutionary biology journals.

TRAIT EVOLUTION

One of the primary applications of phylogenetic trees is the reconstruction of history of character evolution. Together with a model of how characters change, trees are used to infer the ancestral states at nodes, the number of transitions between character states, and the location of these changes along the tree. There are quite a few methods for ancestral state reconstruction and character mapping (e.g., parsimony, maximum likelihood, Bayesian stochastic mapping), and these are covered by other articles in this series. The goal of this section is to give citations for some of the articles that provide a basic understanding of concepts related to character evolution and explain how to interpret trees showing mapped characters. Maddison et al. (1984) outlines the importance of outgroup sampling in understanding character history. Maddison and Maddison (1992) and Baum (2008) give a basic introduction to how trait evolution is portrayed and interpreted, and Schulter et al. (1997) and Ronquist (2004) review statistical methods for studying trait evolution.

Maddison, W. P., M. J. Donoghue, and D. R. Maddison. 1984. Outgroup analysis and parsimony. *Systematic Zoology* 33: 83-103.

In a parsimony framework, outgroup analysis is key in establishing character polarity, i.e. the direction of evolutionary change. This paper discusses outgroup analysis largely with respect to

phylogenetic inference ('tree-building'), but the same principles apply to studying character evolution.

Maddison, W. P. and Maddison, D. R. 1992. *MacClade: Analysis of phylogeny and character evolution*. Sunderland, MA: Sinauer Associates.

The user's manual for this program (which is now freely available online) provides detailed and helpful explanations about how to trace character history and how to interpret the resulting reconstructions. The authors also give empirical examples of most methods.

Schluter, D., T. Price, A. O. Mooers, and D. 1997. Likelihood of ancestor states in adaptive radiation. *Evolution* 51: 1699-1711.

This paper applies maximum likelihood methods based on a Markov model of discrete trait evolution to the estimation of ancestral states. The authors introduce the pie diagram notation at the nodes, which displays relative support for different states. The paper also extends to ancestral state reconstruction for continuous traits.

Ronquist, F. 2004. Bayesian inference of character evolution. *Trends in Ecology and Evolution* 19: 475-481.

Ronquist provides a helpful comparison of different methods for ancestral state reconstruction, and discusses approaches for accommodating uncertainty in the phylogeny (with a focus on Bayesian methods).

Baum, D. A. 2008. Trait evolution on a phylogenetic tree: relatedness, similarity, and the myth of evolutionary advancement. *Nature Education* 1:1.

This article describes how trait evolution is shown on trees, and how trait change at the macroevolutionary level relates to population genetic processes (e.g., drift, natural selection).

CONNECTING GENES, INDIVIDUALS, POPULATIONS AND SPECIES

One of the reason that phylogenetic trees are so widely used in biological research is the basic structure can be used to model genealogical relationships among many different biological entities, from genes to species. Moreover, interpreting gene histories in the context of species histories has become a major research focus in systematic biology. For example, instance of gene duplication and gene loss can be inferred by reconciling gene trees with species trees. Also, large samples of gene trees are often analyzed jointly to infer the most likely species history. The literature on gene trees and species trees is large and growing; the list below provides starting points for understanding how gene trees and species trees relate. Pamilo and Nei (1988) describe the probability of gene trees differing from the species topology in a coalescent framework, and Doyle (1992) and Maddison (1997) discuss processes beyond the sorting of alleles that result in gene tree / species tree conflict. Baum and Shaw (1995) consider these processes in the context of speciation while Avise (2000) covers gene trees in a phylogeographic context. Page and Charleston (1997) focus specifically on the history of duplication and loss in gene families.

Pamilo, P., and M.Nei. 1988. Relationships between gene trees and species trees. *Molecular Biology and Evolution* 5: 568-583.

This mathematically-oriented paper shows how short branches in the history of species divergence lead to conflicting gene trees. They point out that the probability of finding the true species history is increased by sampling more loci across the genome, but not by sampling more alleles from the locus.

Doyle, J. J. 1992. Gene trees and species trees: molecular systematics as one-character taxonomy. *Systematic Botany* 17: 144-163.

Doyle reminds readers that gene trees may not have the same history as the species from which they were sampled. He reviews the biological phenomena that cause these conflicting histories and emphasizes the dangers of relying on a single gene to infer species history.

Baum, D. A., and K. L. Shaw. 1995. "Genealogical perspectives on the species problem." In *Experimental and molecular approaches to plant biosystematics,* Monographs in systematics, Volume 53, Edited by P. C. Hoch and A. G., Stevenson, 289-303. St, Louis: Missouri Botanical Garden Press.

This paper presents a genealogical species concept, wherein members of the same species share gene copies that are more closely related to each other than to gene copies of other species. In the process of laying out this new concept, the authors review the connections between genes, individuals, and species.

Maddison, W. P. 1997. Gene trees in species trees. *Systematic Biology* 46: 523-536.

Maddison notes that many processes can lead to differences between gene trees and the species trees that contain them, and that species trees might be best envisioned as a cloud of gene trees.

Page, R.D.M. and Charleston, M.A. 1997. From gene to organismal phylogeny: reconciled trees and the gene tree/species tree problem. *Molecular Phylogenetics and Evolution* 7: 231–240.

This paper focuses on using species trees to infer the history of gene duplication and loss through 'reconcilation' and provides biological examples from several gene families.

Avise, J. C. 2000. *Phylogeography: the history and formation of species*. Cambridge, MA: Harvard University Press.

This is a highly accessible text that explores the connections between microevolutionary and macroevolutionary processes. It has helpful diagrams showing gene trees within population trees, and describes how barriers to gene flow lead to genetic structure and, eventually, reciprocal monophyly of sister species.

TREES AS EVIDENCE FOR EVOLUTION

The branching history of life and the common ancestry of all extant organisms are core components of evolutionary theory. This history predicts that trait variation across organisms will show tree-like patterns, supporting hierarchically nested clades of taxa. That is, we predict that traits will group species like leaves on a twig, twigs on a branch, and branches on a trunk, with each smaller clade nested within a larger clade. The presence of tree-like structure in various biological data (morphology, DNA sequences, geography) can thus be interpreted as support for evolution. Several methods are available for quantifying phylogenetic (tree-like) signal in data, and empirical studies find that signal is widespread, for across traits and across organisms. This literature highlights the pervasiveness of phylogenetic structure, despite reticulate processes, such as hybridization and lateral gene transfer. Archie (1989), Penny et al. (1982) and Theobald (2010) propose and apply different tests for common ancestry based on data from extant organisms. Sober and Steel (2002), Penny et al. (2003), and Steel and Penny (2010) discuss the strengths and weaknesses of these approaches and future directions in this area of research.

Archie J. W. 1989 A randomization test for phylogenetic information in systematic data. *Systematic Zoology* 38:239-252.

Archie proposed a new method (known as the permutation tail probability, or PTP, test) that tests for phylogenetic structure by comparing randomized (permuted) data to actual data. His test

demonstrated significant structure in morphological datasets from 28 different groups, from flies to salamanders.

Penny, D., L. R. Foulds, and M. S. Hendy. 1982. Testing the theory of evolution by comparing phylogenetic trees constructed from five different protein sequences. *Nature* 297:197–200.
Evolutionary theory predicts that genes from the same organisms should share the same or similar history. Penny et al. demonstrate that the agreement across different genes from the same organisms is significantly higher than would be predicted by chance, providing strong support for the existence of an evolutionary tree.

Sober, E., and M. Steel. 2002. Testing the hypothesis of common ancestry. *Journal of Theoretical Biology* 218: 395-408.
The authors suggest that because common ancestry is a widely accepted tenet of evolution, it has not been rigorously tested. They suggest new approaches for competing the tree model against competing models.

Penny, D., M. S. Hendy, and A. M. Poole. 2003. Testing fundamental evolutionary hypotheses. *Journal of Theoretical Biology* 223: 377-385.
This paper is a response to Sober and Steele (2002) and stresses that evolution remains a fundamentally testable hypothesis and that alternatives can be modeled and compared statistically with a tree.

Theobald, D. L. 2010. A formal test of the theory of universal common ancestry. *Nature* 465: 219-222.
Theobald uses a model selection approach to show that that a single common ancestor of all life is significantly more likely than separate origins of the domains of life (Bacteria, Eukarya, and Archaea).

Steel, M., and D. Penny. 2010. Origins of life: Common ancestry put to the test. *Nature* 465:168–169.
A perspective piece about Theobald's (2010) paper that reviews the literature in this area.

TOOLS FOR TEACHING

Many teaching tools are available for expanding and enhancing tree-thinking as part of biology curricula. Some classic activities are based around the Caminalcules (Gedron, 2000) or the Great Clade Race (Goldsmith, 2003) and new ones include the pipe-cleaner trees (Halverson 2010) and programs like EvoBeaker (Perry et al. 2008). Incorporating these tree-thinking activities into basic evolutionary biology education is necessary for a full understanding of evolutionary process across macro- and micro-scales (Catley 2006).

Gendron, R. P. 2000. The classification and evolution of Caminalcules. *American Biology Teacher* 62: 570-682.
The Caminalcules are imaginary animals created by J. H. Camin in the 1980s and have been widely used as teaching tools. Gendron describes his use of the animals in a series of exercises teaching phylogenetic inference (with and without fossils) and patterns of trait evolution.

Goldsmith, D. W. 2003. The great clade race. *American Biology Teacher* 65: 679-682.
Goldsmith presents an activity that uses runners in a race as an analogy for trait evolution and phylogenetic inference. This is a useful tool for teaching the basic concepts about trait evolution, such as how descendants inherit traits from their ancestors and how lineages accumulate character changes over time.

Catley, K. M. 2006. Darwin's missing link – A novel paradigm for evolution education. *Science Education* 90: 767-783.

Catley points out the overemphasis on microevolutionary processes in the teaching of evolution, and argues that both micro- and macroevolution should be given equal weight.

Perry, J., E. Meir, J. C. Herron, S. Maruca, and D. Stal. 2008. Evaluating two approaches to helping college students understand evolutionary trees through diagramming tasks. *CBE Life Sciences Education* 7: 193–201.

Perry et al. compare the effectiveness of the Great Clade Race and the computer software EvoBeaker for teaching tree-thinking skills.

Halverson, K. L. 2010. Using pipe cleaners to bring the tree of life to life. *American Biology Teacher* 74: 223-224

Halverson describes an activity that uses pipe cleaners to teach tree-thinking skills. It is especially useful for learning to compare topologies and rotating nodes.