Syst. Biol. 62(3):479–489, 2013

DOI:10.1093/sysbio/sys115

The Tree of Life: Metaphor, Model, and Heuristic Device

DAVID P. MINDELL*

Department of Biochemistry & Biophysics, University of California, San Francisco, CA 94158, USA
**Correspondence to be sent to: E-mail: dpmindell@gmail.com.*

Received 12 August 2012; reviews returned 29 October 2012; accepted 18 February 2013
Associate Editor: Elizabeth Jockusch

> The meaning of any conception in the mind is the practical effect it will have in action. (Peirce 1878)
>
> The price of metaphor is eternal vigilance. (attributed to Rosenblueth and Wiener, Lewontin 2001)

The tree of life (ToL) is used variously as metaphor, model, and research tool to explore life's evolution and genealogical relationships. It has broad appeal for multiple scientific communities and the public as a descriptive term covering life's diversity. However, as biologists identify more entities and events to describe—including species, organisms, cells, genomes, gene families, extra-chromosomal genetic elements, endosymbioses, hybridizations, recombination types, and lateral gene transfer (LGT) events—the ToL strains under the weight of multiple uses and expectations. The multiple uses span levels of biological organization, from genes to clades, and they extend to all life forms despite their disparate mechanisms of reproduction and inheritance. Furthermore, some biologists consider the purpose of the ToL to be communicating patterns of relationship, whereas others consider evolutionary process to be a primary focus (e.g., Woese 1998). Use of both metaphors (figures of speech) and models (diagrams) comparing evolution to branching trees go back at least to Darwin (1859). Current multiple uses of the ToL, despite their differences, share reliance on phylogenetic methods. The ToL works well as a strictly bifurcating tree in the absence of reticulate evolution, which results from hybridization, lineage merger, and LGT. However, for organisms where inheritance of traits is often reticulate, as with LGT in Bacteria and Archaea, the evolutionary histories of organismal lineages and their molecular traits can differ dramatically. As a result, identifiable histories for organismal lineages can fade or disappear entirely for lack of characters sharing that same history, complicating recognition, and tracking of species or other taxa. Similarly, the ToL applied to life's earliest evolution and origins is more specifically about molecular sequence lineages and their precursors than organismal lineages. For these reasons, the ToL has been criticized as disproven, obsolete, or restricted, applying to some life forms but not to others (e.g., Doolittle 1999; Bapteste et al. 2009; Puigbo et al. 2009). In this article, I argue for a pluralistic view of the ToL, not as a falsifiable hypothesis about the nature of the branching patterns, but as a broad metaphor for life's common descent, and as an ongoing effort to model genealogical relationships, such as reticulations, for molecular and organismal lineages. The ToL has also become an important heuristic device. This stems from its use in studying evolutionary process. I seek to show the heuristic value of ToL studies by discussing the mutually informing nature of phylogenetic analysis and evolutionary biology, including the impact of ToL studies on core concepts of species, genes, and homology. I suggest that both instrumentalist and realist philosophical approaches to systematics are applied in estimating the ToL. The expanding number of distinctive entities whose common descent with modification can be traced via phylogenetics (genes, genomes, cells, organisms, populations, and clades) requires a pluralistic approach if the ToL is to be useful in developing and integrating new knowledge about evolution. Although the overwhelming majority of the ToL's details are lost to time, much remains to be learned.

HISTORICAL OVERVIEW

The meaning and uses of "the tree of life" by biologists have changed over time, in response to changes in scientific understanding. In 18th and 19th century classifications for organisms, both trees and networks were invoked in summarizing observed similarities in form and function of organisms, without imputing evolution or other mechanisms of change (e.g., Rühling 1774; Batsch 1802; Eichwald 1829; see reviews in Archibald 2009, Ragan 2009, and Tassy 2011) (Fig. 1). Italian naturalist Vitaliano Donati developed an arrangement for aquatic organisms noting (1750, pp. xx–xi), "When I observe the productions of Nature, I do not see one single and simple progression, or chain of beings, but rather I find a great number of uniform, perpetual and constant progressions … [these]

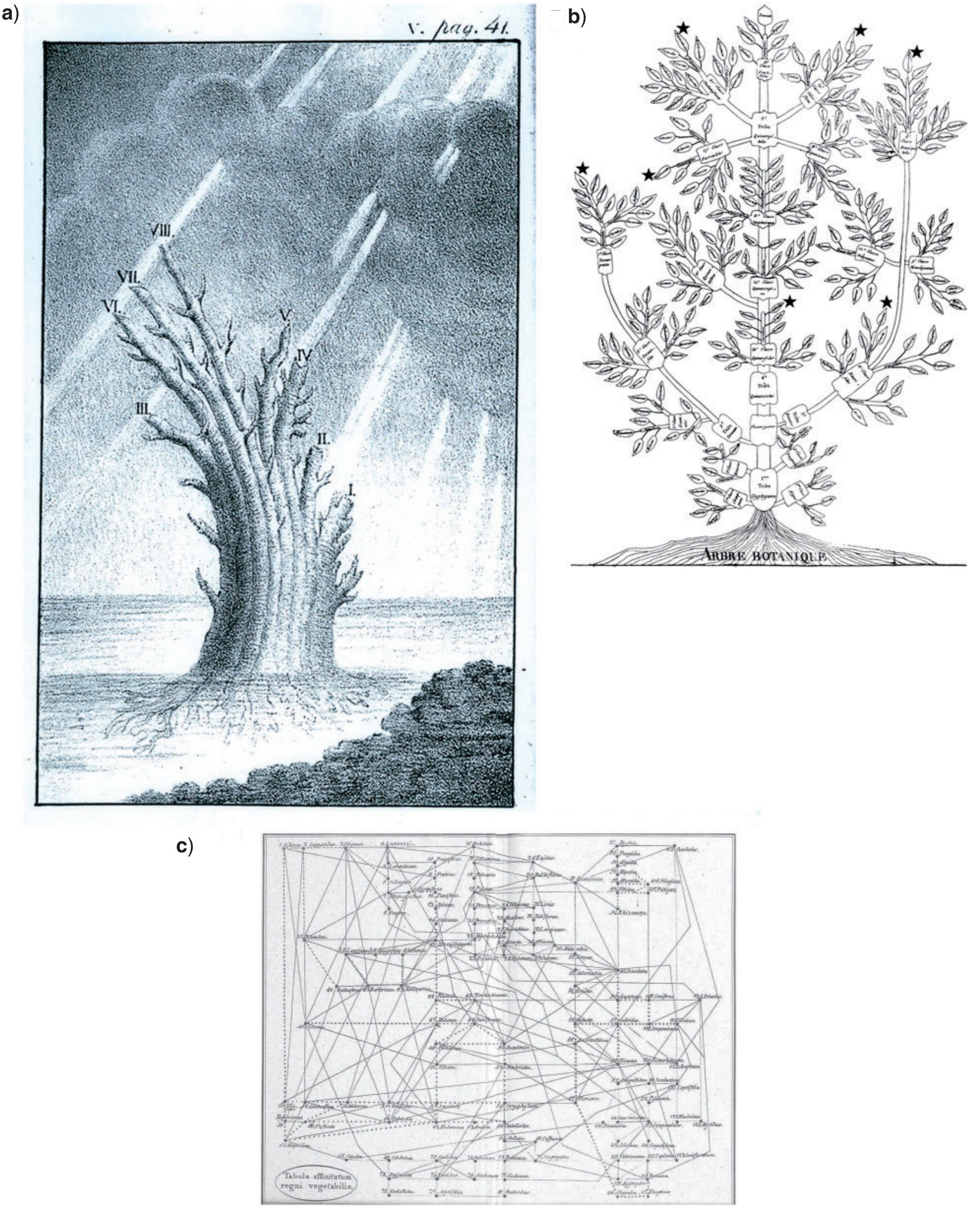

FIGURE 1. Diagrams used in pre-Darwinian classification of organisms. Tree of animal life from Eichwald (1829); Mikulinskii (1972) considered this as representing the tree described by Pallas (1766; see Pallas quote in text below) a). Augier's (1801) *Arbre Botanique* showing plant affinities b). Network of plant affinities from Batsch (1802) c).

should have to be compared more to a Net [rete] than to a chain, that Net being, so to speak, woven with various threads which show, between them, changing communications, connections, and unions." German naturalist Peter Simon Pallas described a classification of organisms using the image of a tree, writing (1766, pp. 23–24; translated from the Latin in Ragan 2009, p. 5), "As Donati has already judiciously observed, the works of Nature are not connected in series in a Scale, but cohere in a Net. On the other hand, the whole system of organic bodies may be well represented by the likeness of a tree that immediately from the root divides both the simplest plants and animals, [but they remain] variously contiguous as they advance up the trunk, Animals and Vegetables; those leading, from Mollusca advancing to Pisces, with great lateral branches of Insects sent out among themselves, from here to Amphibia; and at the extreme top of the tree the Quadrupeds are supported, Aves truly thrust out as an equally great lateral branch below the Quadrupeds." French botanist Augustin Augier also used a tree metaphor for classifying plants, stating (1801, p. 2; see Stevens 1983; Fig. 1) "A figure like a genealogical tree appears to be the most proper to grasp the order and gradation of the series or branches which form classes or families. This figure, which I call a botanical tree, shows the agreements which the different series of plants maintain amongst each other, although detaching themselves from the trunk; just as a genealogical tree shows the order in which different branches of the same family came from the stem to which they owe their origin."

Later, following Lamarck, Darwin, and Haeckel, tree diagrams were used to communicate life's history of common descent with the strong inference of evolution and, eventually, natural selection as a mechanism (see Penny 2011). Commenting on the pre-existing use of trees in classifying organisms, Darwin (1859, p. 99) invoked common descent as an explanation for similarities among organisms, "The affinities of all the beings of the same class have sometimes been represented by a great tree. I believe this simile largely speaks the truth … The limbs, divided into great branches, and these into lesser and lesser branches, were themselves once, when the tree was young, budding twigs, and this connection of the former and present buds by ramifying branches may well represent the classification of all extinct and living species in groups subordinate to groups."

With the advent of molecular sequencing and orthologous gene sets for organismal groups in the 1960s and 1970s, efforts to reconstruct phylogeny for all known life forms in a universal tree with bifurcating or multifurcating branching diagrams came to be associated with the tree of life (ToL) (e.g., Woese 1987). However, growing knowledge about pervasive lateral gene transfer (LGT) in prokaryotes based on incongruent genes trees did not fit a simple, bifurcating-only view of the ToL. This simple view of the ToL gained prominence as it provided a useful foil in arguing for the evolutionary importance of LGT. The simple view was also fostered, if unintentionally, by phylogenetic methods and programs that only provided bifurcating or multifurcating trees, and the much slower development of phylogenetic methods identifying reticulation.

However, existence of reticulation in the form of hybridization among species had long been recognized by biologists, including Darwin, and it takes some imagination to interpret his ToL simile as precluding reticulation among prokaryotes—especially when considering he knew little of prokaryotes and nothing at all of heritable molecular sequences. Darwin's association of a branching ToL giving rise to "groups subordinate to groups" for organismal lineages, is also apparent for molecular sequence lineages involved in LGT. In such cases, the "groups within groups" denote gene genealogies, arising from LGT events with subsequent sequence divergence, duplication, further LGT and organismal divergences. This extrapolation stems from the ToL now spanning levels of biological organization as noted earlier.

Adding reticulations for hybridization, endosymbioses, or LGT, making a bifurcating tree a network or web, and by extension making the ToL a network or web of life which includes both organismal and molecular sequence lineages, is broadly accepted by evolutionists as necessary and useful for explaining many data sets and observations of common descent (Ragan et al. 2009).

Some recent refinements to the ToL accommodating reticulation and focusing on relationships are as follows (see Franklin-Hall 2010). A tree of cells (or cell divisions) aims to trace the evolution of cell lineages back to the last common cellular ancestors of all life using phylogenetic analyses of molecular sequence from all major taxa. Tree branch bifurcations denote 1 cell becoming 2 progeny cells (Zhaxybayeva et al. 2004; Bapteste et al. 2005; Doolittle and Bapteste 2007). A statistical ToL model tests for significant bifurcating tree-like trends in the "forest" of individual gene trees (O'Malley and Koonin 2011; see Puigbo et al. 2009). A rooted network of life model seeks to combine the relationships for many discrete evolutionary units of organisms (e.g., open reading frames, operons, and plasmids) in a single reticulated network, with a rooted tree based on ribosomal genes providing a scaffold, and incongruent gene family phylogenies providing reticulations (Williams et al. 2011a). A recent comprehensive ToL diagram demonstrating a method for summarizing extensive gene transfer (LGT), duplication, and loss, as well as sister relationships among higher level taxa are provided by David and Alm (2011) (Fig. 2).

HEURISTIC VALUE OF THE TOL FOR EVOLUTIONARY BIOLOGY

Phylogeny Informs Evolution

In this section, I discuss the heuristic value of the ToL based on its uses in developing new knowledge about evolution, including aspects of process which can lead, in turn, to improved phylogenetic methods and improved understanding of the assumptions inherent in

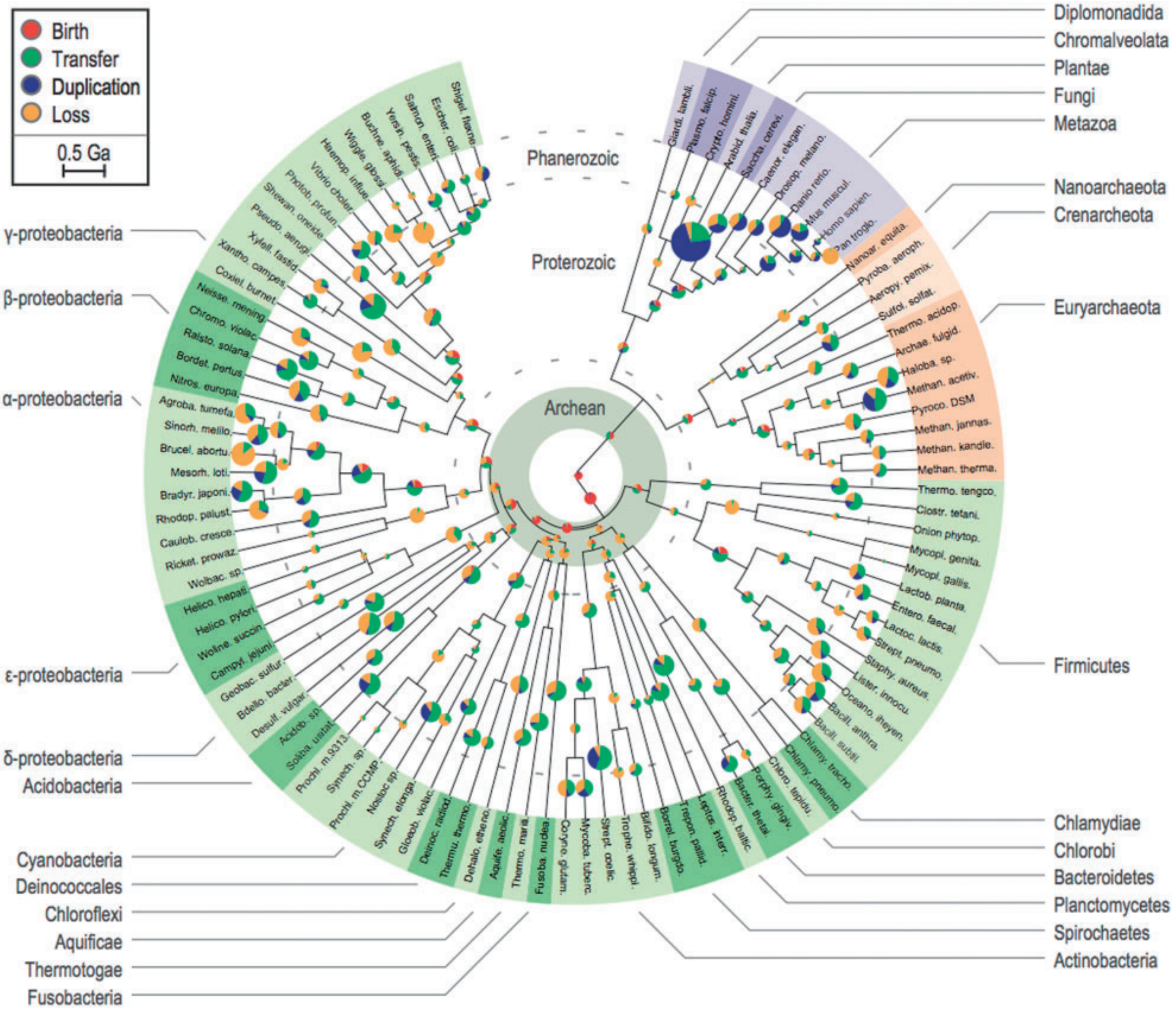

FIGURE 2. ToL diagram showing evolutionary events for diverse lineages from David and Alm (2011, Supplementary Figure S15). The number of macroevolutionary events is mapped to each lineage on an ultrametric ToL and visualized using the iToL website (Letunic and Bork 2007). Pie chart area denotes the number of events, and color indicates event type: gene birth (red), duplication (blue), HGT (green), and loss (yellow).

concepts, such as species and genes. I use the term "ToL studies" to denote phylogenetic analyses and explicitly link them to the ToL. Phylogenetic trees, with or without reticulations, model the genealogical relationships that comprise the ToL. These models become heuristic devices when they are used to learn about evolutionary processes.

Phylogenetic analyses have played primary roles in elucidating the distribution and abundance of LGT in both prokaryotic and eukaryotic lineages, improving our understanding of evolutionary patterns and processes (reviewed in McInerney et al. 2008 and Andersson 2009; see Fig. 2). This includes new details about the chimeric nature of genes, genomes, organisms, and species. Dagan et al. (2008) estimated that at least 81% of the genes in the 22 archaebacterial and 159 bacterial genomes that they analyzed had been laterally transferred among lineages at some time in the past. Thus, 81% of gene trees would differ, at least in part, from corresponding consensus whole genome and organismal species trees. Overall rates of recombination generating LGT in prokaryotes have been estimated to be as high as or higher than their mutation rates (Fraser et al. 2007; Boucher and Bapteste 2009), indicating the evolutionary importance of LGT as a source of genetic variation and adaptive evolution (Marri et al. 2007).

LGT in eukaryotes is also being discovered with increasing frequency, with both prokaryotes and eukaryotes as LGT sources (see Keeling and Palmer 2008; Syvanen 2012). Examples of animal nuclear genomes that have received prokaryotic genes via LGT include bdelloid rotifers (Gladyshev et al. 2008), ciliates (Ricard et al. 2006), arthropods, and vertebrates (Kondrashov et al. 2006). Examples of LGT among eukaryotic lineages include instances among fungi (Slot and Hibbett 2007), between fungi and oomycetes (Richards et al. 2006), and between ancestral opisthokonts and ancestral diplomonads (Simpson et al. 2008). In viruses, LGT and recombination are common in some groups (e.g., bacteriophage, influenza viruses, and retroviruses) and can happen whenever there is coinfection of a host cell and genetic exchange between virus lineages (e.g., Simon-Loriere and Holmes 2011).

Interspecific hybridization and endosymbioses are also evolutionarily important forms of reticulate evolution. Though phylogenetic analyses have not been primary in their discovery, they have been helpful in learning about the degree of relatedness of hybridizing parents, the relative timing of hybridization events, and the creative roles of hybridization in facilitating species diversification within plants and animals (see reviews in Arnold 2006; Mallet 2008; Soltis and Soltis 2009). Phylogenetic analyses including efforts to estimate a comprehensive ToL have helped demonstrate and confirm the bacterial origins of eukaryotic mitochondria and plastids via endosymbioses, as well as subsequent transfer of genes from organellar genomes to the nucleus via endosymbiotic gene transfer (Kurland and Andersson 2000; Martin et al. 2002), a form of LGT. Reticulate evolution via endosymbioses denotes macroevolutionary change giving rise to extensive radiations of new taxa characterized by individual organisms having multiple genomes, and ToL studies have played an important role in developing this new knowledge.

Phylogenetic analyses have also been applied in learning about incomplete lineage sorting within multilocus data sets. The latter is turning out to be a common phenomenon and an important source of incongruence among individual gene trees (Maddison 1997; Pollard et al. 2006; Kuo et al. 2008; Cranston et al. 2009).

Evolution Informs Phylogenetic Methods

As phylogenetic analyses have contributed to increased understanding of evolution, so has knowledge of evolution contributed to refinement of phylogenetic methods. Systematists recognize that all phylogenetic analyses entail some assumptions and systematic biases (Rannala and Yang 2008), and seek to use models that reflect patterns of character evolution that are observed in organismal data sets, to better understand the real world (e.g., Jukes and Cantor 1969; Hasegawa et al. 1985; Yang 1994). Similarly, recognition of frequent incomplete lineage sorting has motivated development of a variety of approaches for estimating species trees from multilocus data or gene trees (see Degnan and Rosenberg 2009; Liu et al. 2009).

Growing understanding of reticulate evolution via LGT, endosymbioses, and interspecific hybridization highlights a pressing need for efficient algorithms and programs for combined analysis of multipartition data sets including the potential to diagnose reticulation (Rannala and Yang 2008; Edwards 2009), and some progress is being made. Concordance among sets of genes for different topologies, potentially distinguishing support for both vertical and horizontal descent, can be assessed in a framework that does not assume a single source of incongruence among gene trees (e.g., Suchard 2005; Ané et al. 2007; Baum 2007; Galtier 2007). Several investigators have focused on detecting hybridization in the presence of incomplete lineage sorting (e.g., Kubatko 2009; Meng and Kubatko 2009; Yu et al. 2011), and Than et al. (2007) develop an approach to detect LGT in the presence of incomplete lineage sorting. Boc et al. (2010) and Boc and Makarenkov (2011) present methods for recovering LGT with a bipartition dissimilarity approach. Alternatively, the PACT program allows reticulated relationships to be found based on potential duplication of entities analyzed (Wojcicki and Brooks 2005; Brooks and van Veller 2008). These efforts use growing knowledge about evolutionary mechanisms to improve accuracy in ToL studies.

Phylogenetic Reticulation Reveals Unsupported Assumptions

Monophyly of evolutionary units.—As a descriptor of life's descent with modification, the ToL necessarily incorporates some core biological concepts. These include genes, species, and other taxa commonly named at the ToL branch tips as putative units of evolution. Identifying reticulation between putative units can show explicitly the lack of support for assumptions: (i) that the units designated are monophyletic or historically singular, (ii) that bifurcation is the primary pattern of diversification across all groups of organisms, and (iii) that genes, species, or other taxa can have only one sister relationship. Exposing lack of support for these assumptions is not new. Hennig (1966) drew a clear distinction between reticulate patterns of breeding within sexual species and bifurcating patterns among species, with the former deemed unresolvable by his methods. Many others have discussed the problems for phylogenetic analyses, recognition of monophyletic groups and classification that are raised by reticulation (e.g., Bremer and Wanntorp 1979; Funk 1981; Nelson 1983; McDade 1990, 1992; Mindell 1992; Doolittle 1999; Goldstein and DeSalle 2000; Mishler and Theriot 2000; Baum 2007; Brooks and van Veller 2008).

We need the concepts of genes and species to discuss and study life's evolution, yet they are not necessarily fundamental units whose components share the same history. Gene definitions have changed from being about units of inheritance of unknown material, to units comprised of a single DNA sequence encoding a single protein, to disjunct DNA or RNA sequences associated with regulatory, transcribed, or other functional regions and corresponding to a component of inheritance. Qualifying adjectives, such as "protein-coding" or "promoter" often precede "gene," and it is widely accepted that no precise definition is possible, as a result of the great variation in mechanisms of molecular biology for acting upon the genetic material. Similarly, no universal definition of species captures the multitude of mechanisms by which groups of organisms differentiate and diversify. Although species within some groups, such as sexually reproducing vertebrates, often appear obvious and readily distinguishable, many more groups of organisms are asexual and less amenable to species delineation. Many accept pluralism in defining "species" and "gene" as a part of studying evolution, a dynamic process, billions of years old, in which change at multiple levels is continual. It is hard to circumscribe moving targets, moveable by many different mechanisms, despite their connections via common descent. Changing meaning for these concepts over time is often driven by ToL studies, and those changes parallel changing use for the ToL itself, as outlined earlier.

Reticulate evolution as revealed or supported by ToL studies has been integrated in the recognition of different forms of homology for organismal and molecular characters. Homologous traits are those that are shared due to common descent, with different forms of homology defined as a result of different evolutionary mechanisms (Mindell and Meyer 2001). Homology for genes is traced not only to processes of speciation giving rise to orthologs in different species and duplication giving rise to paralogs, but also to LGT giving rise to xenologs (Fitch 2000), lineage fusion, as in endosymbioses among prokaryotes and eukaryotes, giving rise to synologs (Gogarten 1994), and recombination yielding chimeric genes, as in exon shuffling, giving rise to partial homologs (Hillis 1994).

Are viruses part of the ToL? Is the ToL monophyletic?— This question of whether viruses are alive has been discussed at length in recent years, and although disagreement remains (e.g., López-García and Moreira 2009; Moreira and López-García 2009), the view of viruses as living, and therefore ToL components, is gaining support (e.g., Bandea 1983; Mindell et al. 2004; Claverie 2006; Forterre 2010; Villarreal and Witzany 2010). This stems from: growing understanding of the broad range of lifestyles among very small organisms, making viruses appear less different; recognizing that viruses satisfy conventional definitions of life involving development, reproduction, and evolution; and finding heritable genes shared with nonviruses. Viruses do require host cells for reproduction, but so do many other organisms, such as, for example, the bacteria Rickettsia and Chlamydia. Recently discovered nucleocytoplasmic large DNA viruses (NCLDVs), including Mimivirus and Marseillevirus, have many genes that are involved in DNA replication, further supporting the idea that viruses are alive and part of the ToL. Some disagree, saying that viruses cannot be linked to the ToL by any genes of their own (Moreira and López-García 2009); that is, because viral genes are all acquired via LGT. Even if that is the case, acquisition of genes from elsewhere, including multiple different sources, seems an arbitrary disqualifier that would lead to disqualification of many nonviruses as well. Recent phylogenetic analyses for 8 homologous DNA processing genes from NCLDVs and from nonviruses representing Eukarya, Bacteria, and Archaea, suggest that NCLDVs have a core genome as old as Eukarya, Bacteria, and Archaea, and raise the idea that they may comprise a fourth domain of life, integral to understanding the ToL (Boyer et al. 2010; Wu et al. 2011; but see Williams et al. 2011b). There is also evidence supporting ancient origins for at least some virus lineages from viral coat protein structures (Burnett 2006).

The heuristic value of the ToL can also be seen in assessment of its root or roots. Evidence for abundant LGT first contributed to doubts about presumed universal common ancestry for life's primary domains, and suggested life may have had multiple independent origins, with independent origins arising from diverse communities of cells in which LGT was more frequent than vertical inheritance (Woese 1998). Recent analyses based on phylogenetic model selection of universal common ancestry versus independent origin scenarios have supported universal common ancestry (Theobald 2010, 2011), although these findings are debated (Koonin and Wolf 2010; Yonezawa and Hasegawa 2010; Martins and Posada 2012). The phylogenetic model selection approach, an aspect of ToL study, appears promising as a heuristic to discriminate between alternative hypotheses about independent origins or universal common ancestry for extant life.

REALISM, INSTRUMENTALISM, AND PRACTICAL EFFECTS

Realism and Instrumentalism

Most systematists pursue phylogenetic analyses to learn about evolutionary patterns and processes in nature, and understand the ToL and evolution to be mutually informing. In that case, current evolutionary understanding and the ToL should seek mutual consistency. Others are less interested in explanations of the natural world and truth, than in making predictions of sister relationships among taxa based on optimal, "empirically adequate" analyses. In this case, the ToL should be evaluated by how effectively it explains available character data, rather than how accurately it explains evolution in nature. These general approaches are known as realism and instrumentalism, respectively (Rosenberg 1994; Rieppel 2007).

Much dissatisfaction with the ToL stems from the lack of correspondence with the growing knowledge of life's complex evolutionary history, which, in turn, stems from its instrumentalist limits as a simplifying model. This includes realization that phylogenetic methods and data sets do not capture many details of prokaryotic evolutionary history. The commitment to strictly bifurcating trees inherent in most software designed for phylogenetic analyses reflects the relative tractability of bifurcation over reticulation, and perhaps the early influence of Hennig (1966, p. 210), who identified the dichotomization of phylogenetic trees as "no more than a methodological principle," necessary for precise, unambiguous character analysis, such that any particular trait is or is not a shared derived character (synapomorphy) for sister taxa. Despite this instrumentalist-type approach, Hennig (1966, p. 99) was a realist, positing, "… the 'beginner' (ancestor) to which each group formation relates is a real reproductive community which has at some time in the past really existed as the ancestral species of the group in question, independently of the mind which conceives it … " and recognized that his system producing strictly dichotomous trees had limitations and would not suit asexually reproducing groups. This is indicative of a long history of mixing instrumentalist and realist approaches in systematics (see Rieppel 2011).

From a realist's perspective, study of the ToL benefits from both approaches. Beginning with little or no knowledge of evolutionary process, an instrumentalist approach relies solely on optimality criteria (e.g., parsimony, similarity) for estimating relationships. As knowledge grows regarding differences in the constraints and rates of evolution for different character types in different lineages, that knowledge can be applied as evolutionary models in phylogenetic analyses, seeking to increase correspondence between the ToL pattern and evolution. When different sequences yield well-supported but incongruent tree topologies, this can support inference of processes, such as LGT and incomplete lineage sorting. This reflects the mutually informing nature of phylogenetic analyses and evolution, as well as the heuristic value of the ToL.

The realist's objective of increasing correspondence between the ToL and evolutionary history requires occasional change in the meanings for terms and concepts. This is not uncommon in biology, particularly, in evolutionary biology where some terms are used in both technical and general discourse (see Keller and Lloyd 1992; Hall and Olson 2003). The alternative of rejecting the ToL as a falsified hypothesis that life's pattern of descent is strictly, or even primarily, bifurcating is unnecessary. There is no evidence that the ToL, as in Darwin's simile, was ever intended as a hypothesis about anything other than common descent, which has overwhelming support.

Practical Effects

Given systematists' general interest in epistemology, I note that the adaptable, broad view of the ToL outlined here is compatible with aspects of the philosophical approaches of realism, as earlier, and with pragmatism. I mention this, not because an external justification is required, but simply to show consilience between epistemology and a pluralistic view of the ToL including reticulation. Pragmatism began with the writings of Charles Sanders Peirce (1878) and has disparate variations (for history see Menand 2001; Haack 2009). However, at its root are 2 related concerns. One is linking theory and practice in science, and the second concern is for meaning, roughly summarized in Peirce's (1878) view that "the meaning of any conception [concept] in the mind is the practical effect it will have in action."

The long-term practical effect of the ToL concept, and ToL studies, has been to better understand life's diversification. This stems from the ToL's heuristic function, described in the previous section. More specifically, the ToL has helped to (i) raise new questions about evolutionary pattern and process, (ii) provide new knowledge about evolution spanning levels of molecular and organismal lineages, (iii) show the need for improved methods in detecting reticulation, and (iv) force us to reconsider our understanding of the origins, forms, and classification of biological diversity.

Starting with putative homologous DNAs (genes) from one or more organisms as the real entities being analyzed, a series of questions arising from our current views of the ToL are as follows.

1. What are the true genealogical relationships (gene trees) for those DNAs?
2. Are trees for multiple genes from the same set of organisms congruent?
3. Are trees for multiple gene regions (e.g., different exons) from the same gene congruent?
4. Are particular gene trees congruent with the species or taxon tree based on additional data?
5. Is the species tree reticulate (web-like); if so, to what degree?
6. Is the true gene tree reticulate; if so, to what degree?

Each of the above has many follow-up questions, regarding the details. These questions include assessment of process underlying the ToL pattern. Phylogenetic discoveries are used continually to refine understanding of the ToL. For example, knowledge of gene tree incongruence discovered in 2 is applied in resolving 5 above, and similarly question 3 is applied in resolving 6, above.

Development of the ToL as a simple but increasingly realistic model of life's diversification is revising understanding of evolution's primary mechanisms. Bapteste and Burian (2010, p. 712) call attention to this in saying, "a systematic account of additional phenomena beyond selection and drift acting on mutation plus vertical transmission … is required to provide an adequate basis of evolutionary change. These phenomena include exchanges of genetic material across cellular lineages, and selection acting independently on the mobile genetic elements that mediate such exchanges, on the exchange processes, and on emergent genetic partnerships bound together, in part, by such exchanges.” Sapp (2009, p. 317) makes related points in saying, “molecular phylogeneticists have transcended Darwinian explanation in revealing evolution’s complementary processes of genetic divergence and integration.” This includes contradicting the idea that evolution does not take leaps—is not saltational—based on abundant reticulation arising from the inheritance of acquired genes and genomes.

Classification

Biologists use classifications, including named species and named higher-level taxa, to communicate with each other and with the public about the scope, ecological functions, and evolution of biodiversity. Classifications also have important societal uses in education, conservation, medicine, and natural resource management. Classifications are increasingly intended to reflect phylogeny, and, thus, the practical effects of the ToL on classification and their applications are considerable. Reticulation is particularly problematic for classification, as it can yield named taxa whose descent cannot be traced to any single phylogenetic node.

Brooks and van Veller (2008, p. 220) have neatly summarized the difficulty in incorporating reticulation into classifications. “There are two choices. Do we classify a tree with reticulations, or do we try to classify a reticulated network? If we choose the former, hybrid species will appear in two different clades, reflecting their phylogenetic origins. This disturbs traditional classifications in which each taxon has a unique place. If we choose to try to classify a network, each species of hybrid origin would be classified separately from either parental species, so the resulting classification would maintain the tradition of each taxon having a unique placement, but could not provide information about the phylogenetic relationships of the hybrid species.” They resolve the situation advocating primacy of phylogeny and having hybrids appear in 2 different clades, saying “If we wish our classifications to reflect what we think we know about evolution, it seems that we will have to opt for the first alternative.”

This is similar to a general approach outlined earlier, allowing multiple sister relationships for taxa, genomes, and genes (Mindell 1992). Where reticulations give rise to chimeric genes, genomes, and species, as well as composite taxa with multiple genomes, they also give rise to multiple sister relationships for the chimera, and if we are to provide a comprehensive view of their relationships, then we should recognize multiple sister relationships for chimera, where the evidence of evolutionary history warrants it. For example, Eukarya would have various sister relationships with (i) proteobacteria based on mitochondrial genes, (ii) cyanobacteria based on chloroplast genes, and (iii) archaebacterial euryarchaeotes as the putative host lineage and nuclear genome source (Pisani et al. 2007).

This follows and recognizes phylogenetic relationships for the heritable molecular components of biological diversity wherever they lead. It is unorthodox, however, in allowing a particular organismal taxon to appear multiple times, in multiple locations, within a single broad classification as a means for accommodating reticulation. This logical approach may work for a small number of reticulation events (e.g., endosymbioses) for a taxon; but the effort becomes increasingly complex and ultimately unworkable as reticulation events accumulate for a given lineage. In some cases with frequent reticulation due to LGT, a tree representing a primary concordance of gene trees or plurality of a genome might be estimated and used for purposes of classification reflecting phylogeny (Baum 2007; see Velasco 2010).

SUMMARY AND OUTLOOK

The ToL has expanded greatly in scope over recent decades, adding increasing numbers of genes, genomes, and new organismal groups to its purview. The kinds of evolutionary lineages (and biodiversity) that the ToL is expected to track through time have expanded as well, growing from organismal lineages alone, to organismal plus molecular sequence lineages. The fact that organismal and sequence lineage histories often differ, and the fact that the ToL has a long history of use by biologists, underlie debate over its current nature (e.g., Doolittle 1999; O’Malley and Boucher 2011). Recognition of widespread reticulate evolution arising from LGT has lead to criticism of the ToL as disproven or restricted in scope, applying to some life forms but not others.

Is the ToL obsolete as either a metaphor or model? Not at all. I argue here that the ToL is thriving based on 3 observations. First, the ToL as a metaphor and model for life’s diversity has a long history of adapting to incorporate new knowledge. The ToL was used by 18th century taxonomists as a metaphor and nonevolutionary model for classifying groups of organisms based on similarities. Starting in the mid-19th century, biologists used the ToL as a simplified model for life’s evolutionary history or common descent. Over time, the ToL has also accommodated changing views of the nature of species, varying from immutable for early taxonomists, to poorly understood units of convenience, as in Darwin’s view, to fundamental units of evolution (based largely on animal studies) during and after the evolutionary synthesis of the 1930s and 1940s, to widespread recognition of the need for pluralism in conceiving species in modern times, reflecting new knowledge of the many different processes of diversification among disparate life forms, including reticulation. Reticulate evolution in the form of hybridization among species and the mating patterns of sexually reproducing plant and animal populations has long been known and accepted as a pattern within the ToL.

Second, the core feature of the ToL, phylogenetic relatedness, remains a valid and useful organizing principle for biological diversity and evolutionary history. There are real historical patterns of common descent to be recovered, despite the inherent difficulties in tracking the particular entities of species, organisms, and genes that we recognize. Indeed, phylogenetic analysis is the primary method for inferring LGT. Furthermore, there seems to be no compelling alternative to the ToL for ordering biological diversity and its history.

Third, ToL studies and hence the ToL, have an important heuristic function. O'Malley and Koonin (2011, p. 8) noted, "the TOL, all its limitations notwithstanding, is necessary as a scaffold for reconstructing scenarios about the evolution of features of organisms (such as various functional systems). Arguably, this is a key goal of evolutionary research, and it is unclear how it can be achieved without using a tree-like framework. This sort of use could be thought of as heuristic."

I advocate for a broad view of the ToL as a metaphor for descent with modification and model for genealogical relationships including reticulations, for both organismal and molecular sequence lineages. This is a realist's approach and is consistent with the heuristic role of the ToL in evolutionary research. The future of the ToL may well depend on its heuristic value. That is, a comprehensive view of the ToL is likely to be sustained if various research communities find it worthwhile to integrate their particular phylogenetic findings with those of the broader community of biologists, and if researchers find it useful to pursue ToL syntheses for research spanning life's disparate forms, as in studies of evolutionary trends, diversification rates, and coevolution. Practical effects of the ToL as a heuristic device include stimulating development of new questions, new analytical tools, and new evolutionary knowledge, particularly for distinguishing and integrating evolution of organismal and molecular lineages.

Doolittle and Bapteste (2007, p. 2048) remarked, "It is important for modern phylogeneticists to remember that reconstructing the TOL was not the goal of Darwin's theory, but rather it was an integral element of his developing model of the evolutionary process … The TOL was thus the ladder that helped the community to climb the wall of acceptance and understanding of evolutionary process. But now that we have climbed it, we do not need this ladder anymore." In the same vein, Doolittle (2010) has noted the problems that a strictly bifurcating ToL has as a metanarrative for evolution, in that it enables creationist attacks when disagreements arise over the nature of the ToL. He states (2010, p. 468), "Tree-making should be part of our evolutionary toolkit …, but not the backbone of the evolutionary metanarrative that we seem to feel obliged to defend from antiscientific attack. We should abandon metanarration, in fact." Doolittle is right: Simplifying models are easy targets for critics, especially, those seeking a straw man to defeat.

However, should we or can we really abandon metanarration or ToL building as suggested? No, not if the ToL as the heuristic "ladder that helped the community" has continuing value. Rather, I think persistent work toward a comprehensive ToL as a metanarrative for life's common descent is necessary, despite the difficulties, precisely because (i) our knowledge of relationships is far from complete and (ii) the ToL is one of our best exploratory tools for first distinguishing then synthesizing evolutionary narratives across the levels of molecular sequences and organisms. ToL studies revealing incongruence among gene trees can teach us about the relative timing and mechanisms of divergence among various genes, populations, and species. Gene tree incongruence can also help us infer timing, abundance, and mechanisms for LGT, as well as the limits of various definitions of genes and species in particular groups, including prokaryotes. If evolution is to remain the unifying concept for biology, and if disparate sub-disciplines are to share concepts and an understanding of complex evolutionary processes, metanarrative and comprehensive heuristic ToL studies are indispensable.

Ultimately, the ToL provides a simplistic presentation of some of life's relationships. No amount of detail from current or future data sets can give a full accounting, due to extinctions and the record of organismal and molecular change being overwritten. ToL components grow, diversify, diminish, and die in every moment, across the planet, now as in the distant past. Though often perceptible to our compartmentalizing minds, the units we have devised to parse DNAs and groups of organisms, allowing us to track them through time and talk about them amongst ourselves and with the public, are also simplifying constructs. A complete, or even satisfactory, ToL will always elude us, though our understanding can grow and become more realistic.

ACKNOWLEDGMENTS

I thank Ron DeBry, Elizabeth Jockusch, Brent Mishler, David Penny, Peter Roopnarine, Joel Velasco, and an anonymous reviewer for helpful review comments.

REFERENCES

Andersson J.O. 2009. Gene transfer and diversification of microbial eukaryotes. Annu. Rev. Microbiol. 63:177–193.

Ané C., Larget B., Baum D.A., Smith S.D., Rokas A. 2007. Bayesian estimation of concordance among gene trees. Mol. Biol. Evol. 24: 412–426.

Archibald J.D. 2009. Edward Hitchcock's pre-Darwinian (1840) "Tree of Life." J. Hist. Biol. 42:561–592.

Arnold M.L. 2006. Evolution through genetic exchange. Oxford: Oxford University Press.

Augier A. 1801. Essai d'une nouvelle Classification des Végétaux conforme à l'Ordre que la Nature paroit avoir suivi dans le Règne Végétal: d'ou Resulte une Méthode qui conduit à la Conaissance des Plantes & de leur Rapports naturels. Lyon (France): Bruyset Ainé.

Bandea C. 1983. A new theory on the origin and the nature of viruses. J. Theor. Biol. 105:591–602.

Bapteste E., Burian R.M. 2010. On the need for integrative phylogenomics, and some steps toward its creation. Biol. Philos. 25:711–736.

Bapteste E., O'Malley M.A., Beiko R.G., Ereshefsky M., Gogarten J.P., Franklin-Hall L., Lapointe F.-J., Dupré J., Dagan T., Boucher Y., Martin W. 2009. Prokaryotic evolution and the tree of life are two different things. Biol. Direct 4:34.

Bapteste E., Susko E., Leigh J., MacLeod D., Charlebois R.L., Doolittle W.F. 2005. Do orthologous gene phylogenies really support tree-thinking? BMC Evol. Biol. 5:33.

Batsch A.J.G.C. 1802. Tabula affinitatum regni vegetabilis, quam delineavit, et nunc ulterius adumbratam…. Vinariae. Weimar (Germany): Landes-Industrie-Comptoir.

Baum D.A. 2007. Concordance trees, concordance factors, and the exploration of reticulate genealogy. Taxon 56:417–426.

Boc A., Makarenkov V. 2011. Towards an accurate identification of mosaic genes and partial horizontal gene transfers. Nucleic Acids Res. 39:e144.

Boc A., Philippe H., Makarenkov V. 2010. Inferring and validating horizontal gene transfer events using bipartition dissimilarity. Syst. Biol. 59:195–211.

Boucher Y., Bapteste E. 2009. Revisiting the concept of lineage in prokaryotes: a phylogenetic perspective. BioEssays 31: 526–536.

Boyer M., Madoui M.-A., Gimenez G., La Scola B., Raoult D. 2010. Phylogenetic and phyletic studies of informational genes in genomes highlight existence of a 4th domain of life including giant viruses. PLoS ONE 5(12):e15530.

Bremer K., Wanntorp H.-E. 1979. Hierarchy and reticulation in systematics. Syst. Zool. 28:624–627.

Brooks D.R., van Veller M.G.P. 2008. Assumption 0 analysis: comparative phylogenetic studies in the age of complexity. Ann. Mo. Bot. Gard. 95:201–223.

Burnett R.M. 2006. More barrels from the viral tree of life. Proc. Natl. Acad. Sci. USA 103:3–4.

Claverie J.M. 2006. Viruses take center stage in cellular evolution. Genome Biol. 7:110.

Cranston K.A., Hurwitz B., Ware D., Stein L., Wing R.A. 2009. Species trees from highly incongruent gene trees in rice. Syst. Biol. 58: 489–500.

Dagan T., Artzy-Randrup Y., Martin W. 2008. Modular networks and cumulative impact of lateral transfer in prokaryote genome evolution. Proc. Natl. Acad. Sci. USA 105:10039–10044.

Darwin C. 1859. On the origin of species. London: Murray.

David L.A., Alm E.J. 2011. Rapid evolutionary innovation during an Archaean genetic expansion. Nature 469:93–96.

Degnan J.H., Rosenberg N.A. 2009. Gene tree discordance, phylogenetic inference and the multispecies coalescent. Trends Ecol. Evol. 24:332–340.

Donati V. 1750. Della Storia Naturale Marina dell' Adriatico. Venice (Italy): Francesco Storti.

Doolittle W.F. 1999. Phylogenetic classification and the universal tree. Science 284:2124–2128.

Doolittle W.F. 2010. The attempt on the life of the Tree of Life: science, philosophy and politics. Biol. Philos. 25:455–473.

Doolittle W.F., Bapteste E. 2007. Pattern pluralism and the tree of life hypothesis. Proc. Natl. Acad. Sci. USA 104:2043–2049.

Edwards S.V. 2009. Is a new and general theory of molecular systematics emerging? Evolution 63(1):1–19.

Eichwald E. 1829. Zoologia specialis quam expositis animalibus tum vivis, tum fossilibus potissimum Rossiae in universum, et Poloniae in species, in usum lectionum publicarum in Universitate Caesarea Vilnensi habendarum. Pars prior. Propaedeuticam zoologiae atque specialem Heterozoorum expositionem continens. Vilnae [Vilnius, Lithuania]: Josephus Zawadzki.

Fitch W.M. 2000. Homology a personal view on some of the problems. Trends Genet. 16:227–231.

Forterre P. 2010. Defining life: the virus viewpoint. Orig. Life Evol. Biosph. 40:151–160.

Franklin-Hall L.R. 2010. Trashing life's tree. Biol. Philos. 25: 689–709.

Fraser C., Hanage W.P., Spratt B.G. 2007. Recombination and the nature of bacterial speciation. Science 315:476–480.

Funk V.A. 1981. Special concerns in estimating plant phylogenies. In: Funk V.A., Brooks D.R., editors. Advances in cladistics, vol. I. New York: New York Botanical Garden. p. 73–86.

Galtier N. 2007. A model of horizontal gene transfer and the bacterial phylogeny problem. Syst. Biol. 56:633–642.

Gladyshev E.A., Meselson M., Arkhipova I.R. 2008. Massive horizontal gene transfer in bdelloid rotifers. Science 320:1210–1213.

Gogarten J.P. 1994. Which is the most conserved group of proteins? Homology-orthology, paralogy, xenology and the fusion of independent lineages. J. Mol. Evol. 39:541–543.

Goldstein P.Z., DeSalle R. 2000. Phylogenetic species, nested hierarchies, and character fixation. Cladistics 16:364–384.

Haack S. 2009. Evidence and inquiry: a pragmatist reconstruction of epistemology. Amherst (NY): Prometheus Books.

Hall B.K., Olson W.M. 2003. Keywords & concepts in evolutionary developmental biology. Cambridge (MA): Harvard University Press.

Hasegawa M., Kishino H., Yano T. 1985. Dating the human-ape splitting by a molecular clock of mitochondrial DNA. J. Mol. Evol. 22: 160–174.

Hennig W. 1966. Phylogenetic systematics. Urbana-Champaign (IL): University of Illinois Press.

Hillis D.M. 1994. Homology in molecular biology. In: Hall B.K., editor. Homology: the hierarchical basis of comparative biology. San Diego (CA): Academic Press. p. 339–368.

Jukes T.H., Cantor C.R. 1969. Evolution of protein molecules. In: Munro H.N., editor. Mammalian protein metabolism. New York: Academic Press. p. 21–123.

Keeling P.J., Palmer J.D. 2008. Horizontal gene transfer in eukaryotic evolution. Nat. Rev. Genet. 9:605–618.

Keller E.F., Lloyd E.A., editors. 1992. Keywords in evolutionary biology. Cambridge (MA): Harvard University Press.

Kondrashov F.A., Koonin E.V., Morgunov I.G., Finogenova T.V., Kondrashova M.N. 2006. Evolution of glyoxylate cycle enzymes in metazoa: evidence of multiple horizontal transfer events and pseudogene formation. Biol. Direct 1:31.

Koonin E.V., Wolf Y.I. 2010. The common ancestry of life. Biol. Direct 5:64.

Kubatko L.S. 2009. Identifying hybridization events in the presence of coalescence via model selection. Syst. Biol. 58:478–488.

Kuo C.-H., Wares J.P., Kissinger J.C. 2008. The Apicomplexan whole genome phylogeny: an analysis of incongruence among gene trees. Mol. Biol. Evol. 25:2689–2698.

Kurland C.G., Andersson S.G. 2000. Origin and evolution of the mitochondrial proteome. Microbiol. Mol. Biol. Rev. 64: 786–820.

Letunic I., Bork P. 2007. Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. Bioinformatics 23: 127–128.

Lewontin R.C. 2001. In the beginning was the word (a book review). Science 291:1263–1264.

Liu L., Yu L., Kubatko L.S., Pearl D.K., Edwards S.V. 2009. Coalescent methods for estimating phylogenetic trees. Mol. Phylogenet. Evol. 53:320–328.

López-García P., Moreira D. 2009. Yet viruses cannot be included in the tree of life. Nat. Rev. Microbiol. 7:615–617.

Maddison W. P. 1997. Gene trees in species trees. Syst. Biol. 46: 523–536.

Mallet J. 2008. Hybridization, ecological races and the nature of species: empirical evidence for the ease of speciation. Phil. Trans. R. Soc. B 363:2971–2986.

Marri P.R., Hao W., Golding G.B. 2007. The role of laterally transferred genes in adaptive evolution. BMC Evol. Biol. 7(Suppl 1):S8.

Martin W., Rujan T., Richly E., Hansen A., Cornelsen S., Lins T., Leister D., Stoebe B., Hasegawa M., Penny D. 2002. Evolutionary analysis of Arabidopsis, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. Proc. Natl. Acad. Sci. USA 99:12246–12251.

Martins L., Posada D. 2012. Proving universal common ancestry with similar sequences. Trends Evol. Biol. 4:e5

McDade L.A. 1990. Hybrids and phylogenetic systematics I. Patterns of character expression in hybrids and their implications for cladistic analysis. Evolution 44:1685–1700.

McDade L.A. 1992. Hybrids and phylogenetic systematics II. The impact of hybrids on cladistic analysis. Evolution 46: 1329–1346.

McInerney J.O., Cotton J.A., Pisani D. 2008. The prokaryotic tree of life: past, present and future? Trends Ecol. Evol. 23:276–281.

Menand L. 2001. The metaphysical club. New York: Farrar, Straus and Girous.

Meng C., Kubatko L.S. 2009. Detecting hybrid speciation in the presence of incomplete lineage sorting using gene tree incongruence: a model. Theor. Popul. Biol. 75:35–45.

Mikulinskii S.R., editor. 1972. History of biology from ancient times to the beginning of the twentieth century (in Russian). Moscow (Russia): Nauka.

Mindell D.P. 1992. Phylogenetic consequences of symbioses: Eukarya and Eubacteria are non-monophyletic taxa. BioSystems 27: 53–62.

Mindell D.P., Meyer A. 2001. Homology evolving. Trends Ecol. Evol. 16:434–440.

Mindell D.P., Rest J.S., Villarreal L.P. 2004. Viruses and the tree of life. In: Cracraft J., Donoghue M.J., editors. Assembling the tree of life. New York: Oxford University Press. p. 107–118.

Mishler B.D., Theriot E.C. 2000. The phylogenetic species concept: monophyly, apomorphy, and phylogenetic species concepts. In: Wheeler Q.D., Meier R., editors. Species concepts and phylogenetic theory. New York: Columbia University Press. p. 119–132.

Moreira D., López-García P. 2009. Ten reasons to exclude viruses from the tree of life. Nat. Rev. Microbiol. 7:306–311.

Nelson G. 1983. Reticulation in cladograms. In: Platnick N.I., Funk V.A., editors. Advances in cladistics. Proceedings of the second meeting of the Willi Hennig Society. New York: Columbia University Press. p. 105–111.

O'Malley M.A., Boucher Y., editors. 2011. Article collection: beyond the tree of life. Biol. Direct. 6. Available at http://www.biology-direct.com/series/tree_of_life.

O'Malley M.A., Koonin E.V. 2011. How stands the Tree of Life a century and a half after The Origin? Biol. Direct 6:32.

Pallas P.S. 1766. Elenchus zoophytorum sistens generum adumbrationes generaliores et specierum cognitarum succintas descriptiones, cum selectis auctorum synonymis. The Hague (The Netherlands): Apud Petrum van Cleef.

Peirce C.S. 1878. How to make our ideas clear. Cambridge (MA): MIT Press / A Bradford Book.

Penny D. 2011. Darwin's theory of descent with modification, versus the biblical tree of life. PLoS Biol. 9:e1001096.

Pisani D., Cotton J.A., McInerney J.O. 2007. Supertrees disentangle the chimerical origin of eukaryotic genomes. Mol. Biol. Evol. 24: 1752–1760.

Pollard D.A., Iyer V.N., Moses A.M., Eisen M.B. 2006. Widespread discordance of gene trees with species tree in Drosophila: evidence for incomplete lineage sorting. PLoS Genet. 2:e173.

Puigbo P., Wolf Y.I., Koonin E.V. 2009. Search for a Tree of Life in the thicket of the phylogenetic forest. J. Biol. 8:59.

Ragan M.A. 2009. Trees and networks before and after Darwin. Biol. Direct 4:43.

Ragan M.A., McInerney J.O., Lake J.A. 2009. The network of life: genome beginnings and evolution. Phil. Trans. R. Soc. Lond. B 364:2169–2185.

Rannala B., Yang Z. 2008. Phylogenetic inference using whole genomes. Annu. Rev. Genomics Hum. Genet. 9:217–231.

Ricard G., McEwan N., Dutilh B., Jouany J.-P., Macheboeuf D., Mitsumori M., McIntosh F., Michalowski T., Nagamine T., Nelson N., Newbold C., Nsabimana E., Takenaka A., Thomas N., Ushida K., Hackstein J., Huynen M. 2006. Horizontal gene transfer from bacteria to rumen ciliates indicates adaptation to their anaerobic, carbohydrates-rich environment. BMC Genomics 7:22.

Richards T.A., Dacks J.B., Jenkinson J.M., Thornton C.R., Talbot N.J. 2006. Evolution of filamentous plant pathogens: gene exchange across eukaryotic kingdoms. Curr. Biol. 16:1857–1864.

Rieppel O. 2007. The nature of parsimony and instrumentalism in systematics. J. Zool. Syst. Evol. Res. 45:177–183.

Rieppel O. 2011. Willi Hennig's dichotomization of nature. Cladistics 27:103–112.

Rosenberg A. 1994. Instrumental biology and the disunity of science. Chicago: University of Chicago Press.

Rühling IoP. 1774. Ordines naturales plantarum commentatio botanica. Göttingen (Germany): Abraham Vandenhoeck.

Sapp J. 2009. The new foundations of evolution. London: Oxford University Press.

Simon-Loriere E., Holmes E.C. 2011. Why do RNA viruses recombine? Nat. Rev. Microbiol. 9:617–626.

Simpson A.G., Perley T.A., Lara E. 2008. Lateral transfer of the gene for a widely used marker, alpha-tubulin, indicated by a milti-protein study of the phylogenetic position of Andalucia (Excavata). Mol. Phylogenet. Evol. 47:366–377.

Slot J.C., Hibbett D.S. 2007. Horizontal transfer of a nitrate assimilation gene cluster and ecological transitions in fungi: a phylogenetic study. PLoS ONE 2:e1097.

Soltis P.S., Soltis D.E. 2009. The role of hybridization in plant speciation. Ann. Rev. Plant Biol. 60:561–588.

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

Suchard M.A. 2005. Stochastic models for horizontal gene transfer: taking a random walk through tree space. Genetics 170: 419–431.

Syvanen M. 2012. Evolutionary implications of horizontal gene transfer. Annu. Rev. Genet. 46:341–358.

Tassy P. 2011. Trees before and after Darwin. J. Zool. Syst. Evol. Res. 49:89–101.

Than C., Ruths D., Innan H., Nakhleh L. 2007. Confounding factors in HGT detection: statistical error, coalescent effects, and multiple solutions. J. Comput. Biol. 14:517–535.

Theobald D.L. 2010. A formal test of the theory of universal common ancestry. Nature 465:219–223.

Theobald D.L. 2011. On universal common ancestry, sequence similarity, and phylogenetic structure: the sins of P-values and the virtues of Bayesian evidence. Biol. Direct 6:60.

Velasco J.D. 2010. Species, genes, and the tree of life. Brit. J. Phil. Sci. 61:599–619.

Villarreal L.P., Witzany G. 2010. Viruses are essential agents within the roots and stem of the tree of life. J. Theor. Biol. 262:698–710.

Williams D., Fournier G.P., Lapierre P., Swithers K.S., Green A.G., Andam C.P., Gogarten J.P. 2011a. A rooted net of life. Biol. Direct 6:45.

Williams T.A., Embley T.M., Heinz E. 2011b. Informational gene phylogenies do not support a fourth domain of life for nucleocytoplasmic large DNA viruses. PLoS ONE 6(6):e21080.

Woese C.R. 1987. Bacterial evolution. Microbiol. Rev. 51:221–271.

Woese C.R. 1998. The universal ancestor. Proc. Natl. Acad. Sci. USA 95:6854–6859.

Wojcicki M., Brooks D.R. 2005. PACT: a simple and efficient algorithm for generating area cladograms. J. Biogeogr. 32:755–774.

Wu D., Wu M., Halpern A., Rusch D.B., Yooseph S., Frazier M., Venter J.C., Eisen J.A. 2011. Stalking the fourth domain in metagenomic data: searching for, discovering, and interpreting novel, deep branches in marker gene phylogenetic trees. PLoS ONE 6:e18011.

Yang Z. 1994. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. J. Mol. Evol. 39:306–314.

Yonezawa T., Hasegawa M. 2010. Was the universal common ancestry proved? Nature 468: E9–E9.

Yu Y., Than C., Degnan J.H., Nakhleh L. 2011. Coalescent histories on phylogenetic networks and detection of hybridization despite incomplete lineage sorting. Syst. Biol. 60:138–149.

Zhaxybayeva O., Lapierre P., Gogarten J.P. 2004. Genome mosaicism and organismal lineages. Trends Genet. 20:254–260.