Is the Tree of Life the Best Metaphor, Model, or Heuristic for Phylogenetics?

DAVID A. MORRISON

Section for Parasitology, Swedish University of Agricultural Sciences, 751 89 Uppsala, Sweden
Correspondence to be sent to: E-mail: david.morrison@slu.se

Mindell (2013) [henceforth simply "Mindell"] has provided a claim that the Tree of Life (ToL) is still useful in phylogenetics as a model, a metaphor, and a heuristic. Here I examine all three of these claims. Mindell noted that what biologists have long discussed as the ToL is in fact reticulate, and always has been. He therefore objected to simplistic declarations that there is no ToL, in the sense of no recoverable phylogenetic history for recognized taxa. Mindell argued that a "tree with reticulations" is still basically a tree, and that the ToL therefore continues to be a useful metaphor, model, and heuristic for phylogenetics.

Here, I argue an alternative point of view, in which a "tree with reticulations" is a network, and that therefore a network will be a *better* metaphor, model, and heuristic for phylogenetics, in the sense that it will be more inclusive and more powerful. This distinction between tree and network in the face of reticulations is not a semantic one. The tree metaphor/model/heuristic pre-supposes tree-like data, whereas the network allows the data to determine the tree-likeness of the metaphor/model/heuristic—some networks are more tree-like than are others. So, the network view does not deny the importance of the ToL, but simply makes it a special case of something much more general.

In this Point of View I will point out, first, that the tree as a metaphor is actually pre-dated by the network as a phylogenetic metaphor, and that many of the current debates in phylogenetics are not new and actually date back to the beginnings of phylogenetic analysis. We can learn from this history. Second, I point out that the use of a tree as a model pre-supposes a particular approach to data modeling that has serious limitations for phylogenetic analysis. Modeling from complex to simple, rather than simple to complex, has benefits for the analysis. Finally, I note that, as a heuristic, an unrooted network has many advantages over the use of a rooted tree. Data exploration and display is a sorely under-valued part of phylogenetics.

Note that I am not arguing that a tree cannot be used as a model, heuristic, or metaphor for phylogenetics, because it surely can, but I am arguing instead that it is not the best model, heuristic, or metaphor—a network is better. Thus, one does not need "a pluralistic view of the ToL" (Mindell, p. 479) but a rather simplistic view of a network, instead. Mindell sees the ToL as basically a tree even though he recognizes that parts of it are not tree-like; but instead we should see an evolutionary history in which some parts are more tree-like than are others.

"DESTROYING" THE TOL?

I will start by addressing what appears to me to be the basic intent of Mindell's paper. Mindell's main thesis is that reticulation does not invalidate the concept of a ToL, either in theory or in practice. In many ways, this parallels the argument of Greenhill et al. (2009) for anthropology, who concluded from simulations that tree topologies (representing ancestor–descendant cultural histories) can be robust to "realistic levels of borrowing" among cultural groups. Similarly, Andam and Gogarten (2013) have noted that horizontal gene transfer (HGT) can in fact be used to provide information for the ToL, by finding the tree that minimizes the number of transfer events; and furthermore that a transferred gene can also be regarded as a shared derived character, so that preferential modes of gene transfer will result in tree-like patterns of evolution.

Therefore, Mindell's paper can be seen as a response to those commentators who have decided metaphorically to "destroy" the ToL. There are those who see the Tree as having been "uprooted" (Pennisi 1999; Doolittle 2000), "axed" (Lawton 2009), "annihilated" (Michael Syvanen, quoted in *New Scientist* magazine 2009), "trashed" (Franklin-Hall 2010), or "politely buried" (Michael Rose, quoted in *The Guardian* newspaper 2009).

This seems to be an unnecessarily nihilistic approach to recent changes in our understanding of the root of the ToL. There are many evolutionary processes involving the transfer of genetic information, including vertical descent from parent to offspring, but also including so-called horizontal transfer, such as

occurs via hybridization, recombination, introgression, gene transfer, and genome fusion. Vertical transfer in combination with speciation creates a tree-like organismal history, whereas the other processes create reticulations in that tree. All of these processes have presumably been occurring continuously throughout the history of life, and so the phylogeny of life has been more or less network-like in different parts, depending on the balance between vertical and horizontal descent.

So, the ToL needs to suffer no fatality. What is needed is a change of viewpoint, rather than wholesale destruction. With a move to a network-based perspective in phylogenetics, the only fate the Tree suffers is to become a special case of a more general phenomenon—a network simplifies to a tree if there are no reticulation signals in the data. So, metaphorically, the current ToL is not destroyed but becomes more like a forest instead (Dickerman 1998).

A tree is basically a set of inter-connected chains (Baum and Smith 2012), and a network is a set of inter-connected trees. Evolutionarily, the chain represents descent with modification, tracing a single path of ancestor–descendant relationships between a taxon and its forebears. A tree represents the inter-connections of those chains, as divergent speciation creates a multitude of possible paths forward in time from any given ancestor (or, if you prefer, the chains coalesce backward in time). A network then represents the inter-connections of those trees, because each gene fragment has its own genealogical tree, and horizontal evolutionary processes (hybridization, recombination, gene transfer, and introgression) create a multitude of possible paths through the trees within any one genome. Thus, all phylogenies are networks, but if there has been no reticulation then any given phylogeny will also be a tree, and if there has been no speciation then it will also be a chain. Both a chain and a tree are special cases of a network. [Note that a network can also be seen as a set of inter-linked chains, but a tree has more restrictions on the way that the chains can be inter-connected.]

PHYLOGENETIC METAPHORS

The image that I have just used of chains, trees, and networks is, of course, a set of metaphors. Mindell provides a brief history of the tree metaphor in phylogenetics, but this is a slightly Whiggish history, because it suggests that the current preference for the tree metaphor is a logical culmination of history. This is far from the truth. Mindell is, however, right to stress the importance of the metaphor we choose for phylogenetics.

Here I will update the history of phylogenetic metaphors, and point out that the network metaphor precedes the tree in both theory and practice. It has, however, been marginalized for some time, and is only now starting to move back into the spotlight. So, I will argue that the tree is a historical artifact: it was first introduced as a simplification of a network, and Darwin later re-introduced it as a rhetorical device. The tree metaphor has been a digression.

The Origin of Trees and Networks as Metaphors

It is sometimes claimed that our metaphors (or models) for biological relationships have progressed from a chain (or ladder) to a tree to a network (e.g., Kull 2003). However, it has also been noted that the network (or web or map) was the predominant metaphor during the 1800s (Stevens 1984; Ragan 2009), and that the chain is still alive and well in the minds of many non-specialists (see Morrison 2013a for references).

If we take "modern" metaphors for biological relationships to be those that post-date the Great Chain of Being (e.g., Bonnet 1745), when the search for "natural" groups of organisms became more widespread, then the earliest metaphor appears to be that of Donati (1750): "In addition, the links of the chain are joined in such a way with the links of another chain, that the natural progressions should have to be compared more to a net than to a chain, that net being, so to speak, woven with various threads which show, between them, changing communications, connections, and unions." [Translated from the Italian by Ragan 2009.] This was followed immediately by a very similar idea from Linnaeus (1751), in aphorism 77: "All plants show affinities on either side, like territories in a geographical map." [Translated from the Latin.]

The tree metaphor first comes from Pallas (1776), who explicitly acknowledged the earlier ideas: "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, [which remain] variously contiguous as they advance up the trunk, Animals and Vegetables." [Translated from the Latin by Ragan 2009.]

The network metaphor thus not only pre-dates the tree metaphor (by 16 years), but more importantly the tree was explicitly stated to be a simplification of the previously proposed network metaphor. That is, the tree was a metaphoric simplification of admittedly reticulate relationships. This has important implications for modeling, as I will explain in a later section. The first published networks appeared in 1755 (Georges-Louis Leclerc, Comte de Buffon) and 1766 (Antoine Nicolas Duchesne), while the first map appeared in 1792 (Linnaeus) and the first tree in 1801 (Augustin Augier). Thus, appropriately enough, practice followed theory through time.

It is important for my purpose to note that both the map and the tree were originally based on affinity rather than the genealogy expressed by the two networks. Affinity and genealogy are two quite different ways of looking at natural relationships, and confusion often arises because they can be depicted using the same iconography, even though they have such different
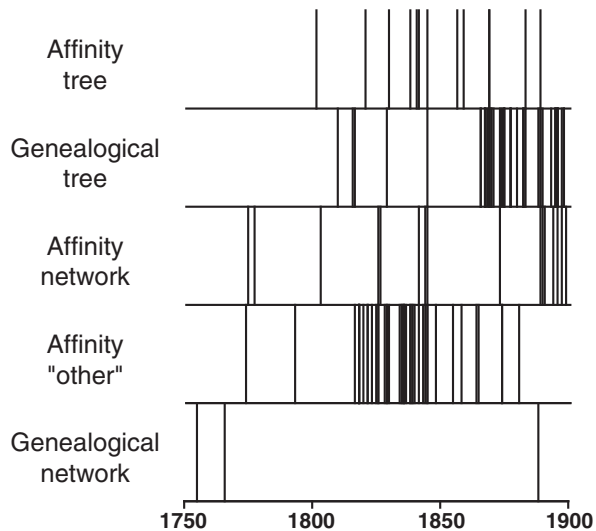
FIGURE 1. Time of publication of 124 papers or books with systematic diagrams, categorized as displaying either affinity of genealogical relationships and as being either a tree, a network, or another form of reticulating figure. "Genealogical trees" includes transformational trees, and "Affinity other" includes (mostly) maps and quinarian diagrams. The data were compiled from citations in Lam (1936), Nelson and Platnick (1981), Gaffney (1984), Stevens (1984, 1994), O'Hara (1988, 1991, 1996), Barsanti (1992), Gould (1997), Coggon (2002), Willmann (2003), Archibald (2009), Ragan (2009), Bigoni and Barsanti (2011), Gontier (2011), Tassy (2011), and Pietsch (2012).

interpretations. Indeed, Mindell perpetuates this same confusion in his "Historical Overview", failing to distinguish affinity trees from genealogical trees.

Affinity refers to a natural (rather than artificial) overall group resemblance, usually quantified (in modern terminology) by some sort of weighted similarity of characters. (In modern terminology, it is phenetics.) It was the underlying concept of the Natural System of classification, which was such an important part of 18th and 19th century systematics (Stevens 1994). In practice, patterns of affinity may, indeed, result from evolutionary relationships, but affinity is a much broader concept than genealogy—in particular, affinity relationships are usually multi-directional rather than nested. This distinction between affinity (phenetics) and genealogy (phylogenetics) runs throughout the history of depictions of biological relationships, and continues to this day, especially in the use of networks to display relationships (see below).

What is equally important is that the subsequent history of relationship images shows the continuing dominance of reticulating diagrams over trees until the publication of Darwin's (1859) major work. If we draw the line at 1900, then Figure 1 illustrates the publication date of 124 systematic works with diagrams. This list is reasonably comprehensive but certainly not exhaustive, and it serves to illustrate the general trends. (The data are included as Online Supplementary Appendix 1, http://dx.doi.org/10.5061/dryad.8p351.) Reticulating affinity diagrams (networks, maps, quinarian; see Pietsch 2012) dominated before 1859, and putatively

genealogical trees dominated thereafter. However, affinity maps began to be replaced by affinity networks late in the 1800s.

The graph shows that systematists clearly disagreed regarding the metaphor needed to describe natural order, but they did not doubt that an order existed, and that it was their role to uncover that underlying order (akin to the search in chemistry for what we now recognize as the periodic table). This approach changed with the gradual acceptance of Darwin's evolutionary ideas.

The Darwinian Digression

Darwin had two effects that are important for the discussion of metaphor. First, he replaced the idea of an inherent order with a less ordered view of biodiversity as resulting from the contingencies of natural selection. This meant that the previous metaphors that allowed for multiple relationships among taxa (required to express the observed complexity of biodiversity), and hence the documented preference for reticulating diagrams (networks, maps, circles, cones, etc.) was no longer needed. Darwin focused attention solely on genealogical relationships, to the exclusion of all other relationships.

Second, Darwin championed the tree as the appropriate metaphor. This was possible because descent with modification, which was one of Darwin's main theses (along with natural selection), can easily be expressed in a tree, provided that we focus (as he did) on vertical genealogical relationships rather than horizontal ones. Darwin knew about horizontal evolutionary events like hybridization, but he did not really integrate them into his metaphor. (Darwin did not use the word "network", but he did use the word "web" with regard to affinity).

An important point here is that affinity relationships can be expressed by an *unrooted* network but genealogical relationships require a *rooted* tree (since this is necessary to depict the ancestor–descendant relationships that were of interest to Darwin). So, in addition to the change from network to tree, there was an equally important change from undirected to directed metaphors. Indeed, Darwin (1859, p. 129) actually introduced the leap from affinity to genealogy via the tree metaphor quoted by Mindell: "The affinities of all the beings of the same class have sometimes been represented by a great tree ... The green and budding twigs may represent existing species), and those produced during each former year may represent the long succession of extinct species." Note that the quote starts with affinity trees but ends with genealogical trees.

These two Darwinian effects were thus combined to reduce, for more than a century, the requirement for (unrooted) networks in systematics and promote the use of (rooted) trees. It has only been in the last 30–40 years that systematists have started seriously questioning this change of metaphor in systematics. For example, during the "cladistic revolution" of the 1970s Sneath (1975)

explicitly questioned the use of trees in bacteriology (where HGT is widespread), and Bremer and Wanntorp (1979) asked about the usefulness of trees for the study of evolution via hybridization (which is widespread among plants).

So, in one sense the modern interest in networks is actually a return to the original metaphor of 250 years ago (a set of inter-linked trees). However, it is important to recognize that it is also a new metaphor, in the sense that what is now required is a *rooted* genealogical network (as used by Buffon and Duchesne) not an unrooted affinity network (as used by most people since then). Figure 1 shows that these were as scarce as hen's teeth before 1900, and they also remained rare during the 1900s. From 1900 to 1990, I know of only nine published genealogical networks, all of them in botany. (This seems to reflect the common perception that zoologists think about hybridization much less than do botanists. The primacy of the tree metaphor may be due to the powerful influence of zoologists on theory and practice in systematics.)

This distinction between rooted and unrooted networks seems to be poorly understood (Morrison 2013b), as the current phylogenetics literature is replete with unrooted networks (most commonly splits graphs), which express affinity relationships rather than genealogical ones. These networks certainly can provide insight into evolutionary history, as originally suggested by Bandelt and Dress (1992), but they are solely heuristic tools. Indeed, as heuristic devices they are likely to be far superior to a phylogenetic tree (discussed further below).

PHYLOGENETIC MODELING

The main importance of metaphors in science is that they serve as the basis for models. That is, we try to express the conceptual metaphor as a mathematical model, which then serves as the basis for quantitative description and prediction. This use of explicit models lies at the heart of making phylogenetics a rigorous science.

Trees, Networks, and Graphs

Mathematically, a tree is an acyclic, leaf-labeled, and connected graph (Semple and Steel 2003). Trees are a well-understood part of graph theory, so that this geometric structure provides a mathematically tractable framework for quantifying the biological concept of a phylogenetic tree. Most importantly for my purpose here, networks have essentially the same mathematical definition as a tree, in that they differ only in the presence or absence of reticulation nodes. Thus, much of what Mindell says about the use of a tree as a model can be made general by applying it to a network, instead.

The graphical relationship between trees and networks is illustrated in Figure 2. Affinity trees and networks (left column) are undirected graphs, which
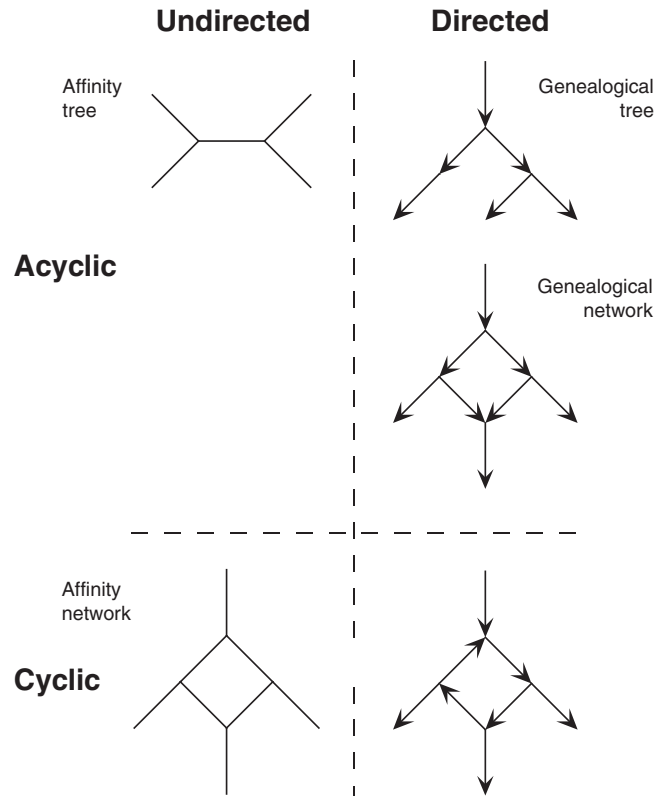

FIGURE 2. The relationship between direction and cycles in a connected graph. In all of the graphs there are four (unlabeled) leaves, although in the directed graphs (right column) one of them is the common ancestor. The number of internal nodes and edges varies depending on whether there are cycles (bottom row: 4 nodes, 4 edges) or not (top row: 2 nodes, 1 edge; middle row: 4 nodes, 4 edges).

biologists call unrooted trees and unrooted networks. The difference between a tree and a network in this case is whether there are cyclic relationships (= loops) or not. Genealogical trees and networks are both acyclic directed graphs, as this combination (directed+acyclic) gives them a root. A directed cyclic graph (bottom-right of Fig. 2) cannot represent a realistic evolutionary history, because at one of the nodes in the cycle (= circuit) an inferred ancestor is also its own descendant (or one of the inferred descendants is also its own ancestor). The difference between a tree and a network in this case is whether there are so-called reticulation nodes, where two directed edges intersect.

From the viewpoint of mathematical modeling, therefore, the generalization from a genealogical tree to a genealogical network is conceptually straightforward— we simply allow reticulation nodes in the acyclic, leaf-labeled, and connected graph. This makes a genealogical tree a special case of the more general notion of a genealogical network.

Current tree-building algorithms involve a model in which character mutations are given relative costs; for example, nucleotide data are modeled based on the probability of substitutions and indels. Presence or absence of particular branches in the tree is then based on the probabilities of character mutations along

those branches. This system can be generalized by adding reticulation events to the model, which are also associated with some cost. Under these circumstances, a tree has an implicit assumption of infinite cost for reticulations (Dickerman 1998), so that they cannot occur, whereas networks apply some smaller non-zero cost.

There are, unfortunately, a number of issues that make this generalization difficult in practice. This has hindered the development of methods by computational scientists, and therefore delayed the widespread use of networks by evolutionary biologists. First, when searching for optimal networks (instead of optimal trees), the space of rooted networks is vastly larger than the space of rooted trees. Searching this space even heuristically is a daunting computational challenge (Huson et al. 2011). Second, adding reticulations to a tree or network will monotonically improve the value of any optimality criterion based on the character data, because a more complex network can never fit the data worse than a simpler one (Makarenkov and Legendre 2000; Jin et al. 2007). This overestimates the amount of reticulation in the data; and so there needs to be a separate optimality criterion for the number of reticulations, as well. Third, the usual strategy used for calculating a rooted phylogenetic tree, where we first produce an unrooted tree and then root it (e.g., with an outgroup), does not work for rooted networks. When we add a root to any of the currently available methods for unrooted networks, the resulting diagram either (i) has internal nodes that do not represent inferred ancestors or (ii) not all of the edges have a unique direction, or both (Morrison 2011).

Nevertheless, work is actively proceeding to address these issues (Bloomquist and Suchard 2010; Nakhleh 2011; Bapteste et al. 2013). The desideratum is the same as for the phylogenetic analysis of trees—that we can feed data into a single analysis and be given a plausible phylogenetic history in return. We are not near this goal yet, but progress is definitely being made.

Strategy for Modeling

This brings us to the issue of the modeling strategy that is currently used in phylogenetics, and how this might change in response to the use of a rooted genealogical network as the model of choice for phylogenetic analysis. The selection of an optimal model for data analysis is important, but there are a large number of models that could be used. This creates a combinatorial problem that cannot be solved by exhaustively examining all possible models, and so we need a search strategy.

In statistical modeling (Anderson 2008), model selection is seen as a tradeoff between bias (due to non-random variation) and variance (due to random variation), and there is no necessarily best procedure for finding the balance between them. Model selection can be tested by starting with the simplest model and progressively adding model complexity until the

desired level of model fit to the data is achieved. Alternatively, one can start with the most complex model and progressively delete unnecessary components while maintaining the desired level of model fit to the data. The first approach is constructive, in the sense that the model is constructed piece by piece (stepwise addition), while the second approach is reductive, in the sense that the full model is pared down to its simplest form (stepwise deletion).

This distinction in approaches to model selection is relevant to the difference between using trees and networks as phylogenetic models. At the moment, the most common approach to phylogenetic analysis is the constructive one (Felsenstein 2004). One starts with the simplest acceptable model, a bifurcating tree (since a chain is not likely to be adequate, in practice), and assesses the degree to which it fits the data. If the fit is poor, as it often is with multi-gene data, especially if the different gene data are concatenated, then complexity is added. For example, one might include incomplete lineage sorting (ILS) in the model, which allows the different genes to fit different trees, while still maintaining the need for a single dichotomous species tree (Knowles and Kubatko 2010). Alternatively, one might consider gene duplication–loss as a possible addition to the model, this being another major source of incompatibility between multi-gene data and a single species tree (Bansal et al. 2012). Only if these additional complexities also fail to attain the desired degree of model fit does one consider adding components of reticulate evolution to the model, such as hybridization or HGT (e.g., Kubatko 2009; Yu et al. 2011).

Thus, if we approach phylogenetics from the tree perspective then our only choice is to consider reticulations as additional (and unusual) occurrences. The argument is that a tree is the most parsimonious graph because it involves the minimum number of evolutionary connections (Nelson 1983); and we can then simply assume increased homoplasy on a tree, rather than postulating extra events outside of strict descent with modification.

The reductive (or simplification) approach, however, proceeds the other way (Crawley 2007). A general network model would be used as the starting point. The various components of this model would include a dichotomous tree as a special case, along with ILS, duplication–loss, hybridization, recombination, and HGT (or introgression) as individual components. These special cases would then be evaluated simultaneously, and each one dropped if it is contributing nothing worthwhile to the model fit (i.e., its contribution is below some specified threshold). The final model would then consist of the simplest combination of components that still maintains the specified fit of data and model; this may indeed be a simple tree.

The main advantage of the latter approach is that all of the components of the model are evaluated simultaneously, so that their potential interactions can be quantitatively assessed. Components are dropped from the model only if they contribute little to the

model, either independently or in synergy with the other components. That is, they are dropped only if they can be shown to be redundant for practical purposes.

This does not happen with the constructive approach to modeling. Here, the components are evaluated in some specified order, and components that are later in the order will not be evaluated unless the earlier components prove to be inadequate. This order-dependence means that these later components are potentially excluded from statistical consideration, so that their possible contribution to biological explanation may never be quantitatively assessed. This is important because in current phylogenetic analyses evolutionary reticulation is effectively considered to be a "last resort"—it is considered as a possible biological explanation only if all else has already failed.

A specific example is provided by various published analyses of the data set of Rokas et al. (2003), which initially showed an unexpected amount of incongruence among the 106 gene trees analysed. Several authors later questioned the details of the phylogenetic tree model used by Rokas et al., including Phillips et al. (2004), Ren et al. (2005), and Shapiro et al. (2006), all of whom expressed doubt about the adequacy of the particular tree model used. However, Bloomquist and Suchard (2010) and Yu et al. (2012) questioned the use of a tree model in the first place. These authors used a network model, instead, and came to the conclusion that hybridization is a more likely explanation for the incongruent results observed by Rokas et al.

The constructive philosophy seems to be as much a historical artifact as anything else. As detailed above, the first phylogenetic diagrams (by Buffon and Duchesne) were networks not trees, but they were replaced a century later by the tree model advocated by Darwin; and the tree has retained its primacy since that time. This leads naturally to the constructive approach to modeling that is so prevalent in the current literature.

There is no necessary statistical superiority of the constructive approach to modeling. Indeed, statisticians seem to consider forward and backward selection of model components to be essentially equivalent, although they may lead to different models for any given data set (this is called selection bias; Anderson 2008). The most commonly specified advantage of the constructive approach to modeling is that it is likely to avoid possible problems arising from having too many components in the model (e.g., over-parameterization, lack of identifiability); and the biggest disadvantage is the possible order-dependence of the final model, which has been severely criticized, particularly when used for linear regression analysis (e.g., Hocking 1983).

In terms of model fit, a reticulated network cannot fit the data any less well than can a tree, although it may fit no better. Furthermore, it is likely that even a more complicated model (a reticulated network vs. a tree) will not produce improbable evidence (i.e., evidence that could not easily have arisen by chance), whereas a more constrained model (a tree vs. a reticulated network) may produce such improbabilities. On these grounds, a network is expected to be a better first approximation to the true phylogeny than is a dichotomous tree.

So, the reductive approach has the distinct advantage of simultaneously evaluating all possible special cases of a network, and thus does not exclude any possible biological explanation that might apply to the observed data. This may provide more biological insight than does the constructive approach to phylogenetic modeling. As noted by Nakhleh (2013): "phylogenetic incongruence should not be viewed as a problem to be masked or despite which inference should be made; rather, it should be viewed as a powerful character with a rich set of states to reconstruct and understand evolutionary phenomena."

Heuristics

The previous section discussed the use of rooted genealogical networks as models for phylogenetic analysis. Here, I consider the use of unrooted affinity networks as heuristic tools for the exploration of phylogenetic data. Mindell focuses on genealogical relationships, whereas the phylogenetic metaphor and heuristic have historically included affinity as well.

One of Mindell's strongest points is that, irrespective of whether the history of life is tree-like or not, a tree model (or metaphor) is still a valuable heuristic for exploring data that are the product of an evolutionary history. I do not dispute this point, since any epistemic tool that provides scientific insight is a valuable heuristic, even if its ontological status is doubtful.

My point in this section is to emphasize how much *more* valuable a network can be. Doolittle and Bapteste (2007) have suggested that the ToL has out-lived its heuristic usefulness, just as advocates of phylogenetic trees have claimed that the Great Chain of Being has out-lived its usefulness. Since a tree is a set of inter-linked chains, and a network is a set of inter-linked trees, then chains and trees have heuristic value only for rather more basic evolutionary scenarios than apply to networks. Networks provide a better heuristic, with broader implications for the interpretation of evolution. As noted by Huson and Bryant (2006): "Even when evolution proceeds in a tree-like manner, analysis of the data may not be best served by forcing the data onto a tree or assuming a tree like model."

There are at least five heuristic uses of affinity networks in phylogenetics: (i) Exploratory data analysis; (ii) Displaying data patterns; (iii) Displaying data conflicts; (iv) Summarizing analysis results; and (v) Testing phylogenetic hypotheses. It is worth looking at each of these in turn.

Heuristic Use 1: Exploratory Data Analysis

Exploratory data analysis is a seriously under-valued concept in biology in general, not just in phylogenetics; and yet computational scientists

emphasize it very strongly. All data analyses rest on certain basic assumptions, and the data need to be evaluated for their relationship to these assumptions *before* the formal analysis begins. Exploratory analysis traditionally involves both graphical displays and numerical summaries of the data (Tukey 1977); and several examples of the use of affinity networks for graphically exploring phylogenetic data are provided by Wägele and Mayer (2007), Wägele et al. (2009), and Morrison (2010).

Affinity networks are designed to display any character (or tree) conflict that might exist in a data set, without prior assumptions about the causes of those conflicts. These conflicts appear as incompatible character distributions in the data. They might be caused by any one or more of a number of things: (a) Estimation errors, including (i) incorrect data (e.g., inadequate data-collection protocol, poor laboratory/museum/herbarium technique, lack of quality control after data collection, misadventure); (ii) inappropriate sampling (e.g., distant outgroup, rapid evolutionary rates, short internal branches); (iii) model mis-specification (e.g., wrong assessment of primary homology, wrong substitution model, different optimality criteria); and (b) Biological conflict, including (iv) analogy (i.e., parallelism, convergence, reversal); and (v) homology (i.e., hybridization, recombination, introgression, gene transfer, genome fusion, ILS, duplication–loss). Specific examples of the use of affinity networks to explore each of these concepts are provided by Morrison (2011, Appendix 1).

Heuristic Use 2: Displaying Data Patterns

The use of networks to display data patterns is based on the idea that, if the pattern being investigated is still clear even when all of the conflicting data are being displayed, then that must be a strong pattern. This is probably the most common use in the recent literature of splits networks, particularly NeighborNet (Huson and Bryant 2006). These networks are now widespread in biology, and they often accompany phylogenetic trees, as a means of visually showing that the authors' conclusions are robust to the data conflict. Indeed, this was the use intended by Bandelt and Dress (1992) when they first introduced splits networks.

Of considerable potential future importance is the use of network methods for extracting evolutionary signals from whole genomes, the early phase of which often requires heuristic data analyses. The within-organism processes that assemble and shuffle functional motifs, subunits, genes, and gene complexes do not usually produce strictly tree-like patterns, and therefore much of the genomic information is not very effectively summarized using trees (Bapteste et al. 2012). Networks, on the other hand, have enormous potential in this regard.

Heuristic Use 3: Displaying Data Conflicts

The use of networks to display data conflicts is based on the idea that insight will be gained by highlighting the locations of the conflict. Perhaps the most obvious source of potential conflict is between data sets based on different characters, such as different morphological features or different genomic regions. This heuristic use differs from (2) in that here the focus is on the conflict itself rather than on the patterns that appear in spite of the potential conflict; these are two quite different purposes (although in practice they may often be two sides of the same coin).

Morrison (2013b) illustrates use (3) with a SuperNetwork showing the areas of disagreement between the two primate trees published by Mivart (1865), based on characters of the spinal column, and Mivart (1867), based on characters of the limbs. It turns out that much of the conflict is created by the fact that the root location is quite different in the two trees.

This is an interesting example, because it demonstrates that many of the current controversies in the literature (Franklin-Hall 2010; Rieppel 2010; O'Malley and Koonin 2011) are not new. As discussed above, much of the contemporary argument is about whether the ubiquity of HGT (and introgression) entails abandoning the ToL. However, systematists have always known about character conflicts, and this did not lead Darwin or his contemporaries to abandon the tree model. If we consider genes to be the "characters" of genomes (and for convenience we can, although non-recombining sequence blocks would be a better choice), then between-gene conflict is simply the genotype equivalent of the phenotype character conflict widely noted by 18th and 19th century systematists (Stevens 1994). Indeed, when asked about this by Darwin in 1870, Mivart noted: "The diagram in the Pro. Z. Soc. expresses what I believe to be the degree of resemblance as regards the spinal column only. The diagram in the Phil. Trans. expresses what I believe to be the degree of resemblance as regards the appendicular skeleton only" (Darwin Correspondence Project letter 7170). In other words, the trees are character trees not species trees, just as we now recognize a distinction between gene trees and species trees. Mindell's defense of the tree model is thus simply a part of a long historical debate in systematics.

Heuristic Use 4: Summarizing Analysis Results

There are two major aspects to the use of networks for summarizing phylogenetic analyses, whether they be single trees, reconciled trees, consensus trees, or supertrees. One involves bootstrap support values and the other involves Bayesian posterior probabilities. In both cases, the "robustness" of tree results is currently presented most commonly using numbers drawn on the tree branches, but this information could be much better visualized using network reticulations.

Bootstrapping, for example, does not directly evaluate the incompatibilities among the characters, but does so

via a tree-based analysis, whereas a network analysis can evaluate the character incompatibilities directly. Therefore, bootstrap values do not always agree with the equivalent network assessment of data patterns (Wägele and Mayer 2007; Wägele et al. 2009; Morrison 2010, 2013b; Kelchner and BPG 2013). An affinity network such as a splits graph evaluates the character (or distance) data independently of any tree, whereas a bootstrap analysis evaluates the data *solely* in terms of a tree. A bootstrap analysis records the trees at each iteration (or pseudoreplicate) rather than recording the bootstrapped character set itself, and so a bootstrap analysis does not directly assess the character support for the tree. A splits graph *does* assess the characters directly.

The importance of this distinction for phylogenetics is that a tree analysis forces the data into a tree irrespective of how well the data fit that tree. All that is required is that the tree be the optimal one based on a particular criterion (distance, parsimony, likelihood, etc.), while the degree of fit of the data and tree is effectively treated as immaterial to the analysis—the tree-likeness of the data is assumed rather than evaluated.

The consequence of this is that it is entirely possible, indeed quite likely, that there will be characters supporting patterns that never appear in the bootstrap consensus tree. That is, highly supported patterns may get a bootstrap value of zero. Morrison (2013b) illustrates an empirical example of this, where the second-best supported character pattern has a 0% bootstrap value in the Neighbor-joining tree analysis; it does, however, play a large part in the equivalent NeighborNet network analysis. (Note: 0% support can also occur as a result of inconsistency in the tree-building method.)

Bayesian posterior probabilities are currently dealt with in phylogenetics in rather an odd manner. Bayesian methods differ from other forms of probabilistic analysis in that they are concerned with estimating a probability distribution, rather than a single estimate of the maximum probability. That is, Bayesian analysis is not about identifying the most likely outcome, it is about estimating and comparing the likelihood of all possible outcomes (expressed as a posterior probability distribution, combining the likelihoods with the prior probabilities). This creates a potentially confusing situation in a phylogenetic analysis, as the result of most Bayesian analyses is published as a single tree, rather than showing the probability distribution of all trees.

The Bayesian summary tree is basically a Consensus Tree (in practice, usually an extended 50% majority consensus tree), but use of a Consensus Network would be considerably more logical (Holland et al. 2005). The tree produced by the current form of Bayesian phylogenetic analysis is built from the best-supported branches of the set of trees sampled by the Markov Chain Monte Carlo (MCMC) algorithm, but only a subset of compatible branches can be included in the Consensus Tree (the greedily compatible subset). Any well-supported but incompatible or weakly compatible branches will not be shown, and it is the absence of these branches that causes the phylogenetic tree

to deviate from the standard Bayesian philosophy of presenting a probability distribution. A Consensus Network alleviates this problem because it is specifically designed to present a specified percentage of the incompatible branches, as well as the compatible ones (Holland et al. 2004).

Some of the information from the probability distribution is used in the Bayesian tree, usually the posterior probabilities that are attached to each of the tree branches, but this is a poor visual summary of the available information. Morrison (2013b) provides an example showing just how much more informative the network is, as it makes it visually clear just where there are major incompatibilities among the MCMC trees.

Heuristic Use 5: Testing Phylogenetic Hypotheses

Any phylogenetic hypothesis that can be stated *a priori* can potentially be tested using an affinity network. In this approach, the hypothesis is framed in terms of predicting the existence of a particular character pattern in the data, which is then assessed using the network. That is, the hypothesized process is predicted to create a particular pattern *a priori*, and it is the *a posteriori* existence of the pattern that constitutes the test. This would not necessarily be a very strong test (Morrison 2011), because many other possible processes could predict exactly the same character pattern, but it may nevertheless be an important component of a phylogenetic study.

Bosch et al. (2004) provide an explicit example of such hypothesis testing. The authors sequenced segments of the same gene from three primate species, and predicted that: "If gene conversion has been evolutionarily persistent, we expect to see clustering of paralogs, rather than orthologs, in interspecific comparisons". This pattern was tested with a Split Decomposition network, which revealed little evidence for the clustering of orthologs.

AFTERWORD

My argument in this article has been that we would be better off if we simply admit that there is no such thing as a "species tree"—there are only "species networks", some of which are more tree-like than are others. As Mindell noted, the ToL is reticulate and always has been, but I contend that if a set of gene trees is incompatible (i.e., the genes have different histories) then the associated species genealogy should be seen as a network not a tree. A "tree (possibly with reticulations)" is a less useful idea than a "network (which will be more or less tree-like)".

This is a less extreme position than the currently accepted one, which effectively ignores reticulation and thus treats it as equivalent to homoplasy in a tree. Reticulation events (non-vertical descent) should not be relegated to the role of "nuisance parameters" in a model or "noise" in a metaphor, but should be quantified

as potentially important components of evolutionary history. Phylogeneticists do not need to choose *a priori* between a tree and a network, but instead they can let the data decide.

There seems, however, to be some resistance to this idea, and Mindell (whether intentionally or not) has compiled many of the arguments in favor of this resistance. In some sense, then, species trees are still seen as central to phylogenetics. This has many far-reaching effects on the thought processes of phylogeneticists.

For example, Losos et al. (2012) have gone so far as to state: "When two phylogenies are fundamentally discordant, at least one data set must be misleading." This represents an extreme tree-centric viewpoint, as the apparent discordance refers only to a strict tree model. Other researchers have sought to resolve all discordance using ILS as an extra component in the species-tree model, as this has been demonstrated to lead inevitably to incongruent gene trees (see the collection edited by Knowles and Kubatko 2010). Yet others have added duplication–loss as an extra component in the basic species-tree model, in another attempt at gene-tree reconciliation; and Stolzer et al. (2012) have combined ILS and D–L in their tree-reconciliation model. Similarly, Williams et al. (2011) have tried to build a "Rooted Net of Life" by first constructing a tree and then adding reticulations as necessary. So, the tree is apparently still the standard by which phylogenetic results will be judged — it is the *a priori* expectation, rather than the *a posteriori* outcome of the analysis.

As one final example, Teeling and Hedges (2013) noted: "Untangling the root of the evolutionary tree of placental mammals has been nearly an impossible task. The good news is that only three possibilities are seriously considered ... Now, two groups of researchers have scrutinized the largest available genomic data sets bearing on the question and have come to opposite conclusions ... Needless to say, more research is needed." No! These analyses are still based on tree models, no matter how sophisticated those models are (see also the several other papers cited by Teeling and Hedges 2013) and no matter how many data are involved. As noted by Hallström and Janke (2010): "Mammalian evolution may not be strictly bifurcating"—a network analysis of retroposon insertion data provides an alternative hypothesis for the history of placentals: the early divergences involved ILS and hybridization.

We might therefore see the suggested move from trees to networks as a form of Kuhnian paradigm shift. In Kuhn's historical model, during the period of "normal science" the failures of results to conform to the current paradigm are seen not as falsifying the paradigm but are seen as errors by researchers instead (e.g., inadequate models, unreliable data). However, as anomalous results accumulate, a new paradigm emerges that subsumes the old results along with the anomalous results, forming a single new framework or paradigm.

Non-tree-like phylogenetic results are currently *not* seen by many, if not most, phylogeneticists as falsifying the paradigm of a phylogenetic tree, but instead are the result of inadequate phylogenetic tree models (as exemplified by Salichos and Rokas 2013). Nevertheless, these results can also be seen as falsifying that paradigm; and a shift to network-thinking would embrace all of the tree results as well as the non-tree ones. We should not call this a Kuhnian "revolution", of course, since tree-thinking and network-thinking are not incompatible, but rather the one is an extension of the other. A tree with reticulations is a network, and we should call it that rather than calling it a tree.

REFERENCES

Andam C.P., Gogarten J.P. 2013. Biased gene transfer contributes to maintaining the Tree of Life. In: Gophna U., editor. Lateral gene transfer in evolution. New York: Springer. p. 263–274.

Anderson D.R. 2008. Model based inference in the life sciences: a primer on evidence. New York: Springer.

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

Bandelt H.-J., Dress A.W.M. 1992. Split decomposition: a new and useful approach to phylogenetic analysis of distance data. Mol. Phylogenet. Evol. 1:242–252.

Bansal M.S., Alm E.J., Kellis M. 2012. Efficient algorithms for the reconciliation problem with gene duplication, horizontal transfer and loss. Bioinformatics 28:i283–i291.

Bapteste E., Bouchard F., Burian R.M. 2012. Philosophy and evolution: minding the gap between evolutionary patterns and tree-like patterns. In: Anisimova M., editor. Evolutionary genomics: statistical and computational methods. Vol. 2. New York: Humana Press. p. 81–110.

Bapteste E., van Iersel L., Janke A., Kelchner S., Kelk S., McInerney J.O., Morrison D.A., Nakhleh L., Steel M., Stougie L., Whitfield J. 2013. Networks: expanding evolutionary thinking. Trends Genet. 29:439–441.

Barsanti G. 1992. La scala, la mappa, l'albero: immagini e classificazioni della natura fra sei e ottocento. Firenze: Sansoni Editore.

Baum D.A., Smith S.D. 2012. Tree thinking: an introduction to phylogenetic biology. Greenwood Village (CO): Roberts & Company Publishers.

Bigoni F., Barsanti G. 2011. Evolutionary trees and the rise of modern primatology: the forgotten contribution of St. George Mivart. J. Anthrop. Sci. 89:93–107.

Bloomquist E.W., Suchard M.A. 2010. Unifying vertical and nonvertical evolution: a stochastic ARG-based framework. Syst. Biol. 59:27–41.

Bonnet C. 1745. Traité d'insectologie, ou observations sur les pucerons, premiere partie. Paris: Durand.

Bosch E., Hurles M.E., Navarro A., Jobling M.A. 2004. Dynamics of a human interparalog gene conversion hotspot. Genome Res. 14:835–844.

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

Coggon J. 2002. Quinarianism after Darwin's *Origin*: the circular system of William Hincks. J. Hist. Biol. 35:5–42.

Crawley M.J. 2007. The R book. Chichester: John Wiley & Sons.

Darwin C. 1859. On the origin of species by means of natural selection. London: John Murray.

Dickerman A.W. 1998. Generalizing phylogenetic parsimony from the tree to the forest. Syst. Biol. 47:414–426.

Donati V. 1750. Della storia naturale marina dell' Adriatico. Venezia: Francesco Storti.

Doolittle W.F. 2000. Uprooting the Tree of Life. Sci. Am. 282(2):90–95.

Felsenstein J. 2004. Inferring phylogenies. Sunderland (MA): Sinauer Associates.

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

Gaffney E.S. 1984. Historical analysis of theories of chelonian relationship. Syst. Zool. 33:283–301.

Gontier N. 2011. Depicting the Tree of Life: the philosophical and historical roots of evolutionary tree diagrams. Evol. Edu. Outreach 4:515–538.

Gould S.J. 1997. Redrafting the Tree of Life. Proc. Am. Philos. Soc. 141:30–54.

Greenhill S.J., Currie T.E., Gray R.D. 2009. Does horizontal transmission invalidate cultural phylogenies? Proc. Roy. Soc. B, Biol. Sci. 276:2299–2306.

Hallström B.M., Janke A. 2010. Mammalian evolution may not be strictly bifurcating. Mol. Biol. Evol. 27:2804–2816.

Hocking R.R. 1983. Developments in linear regression methodology: 1959–82. Technometrics 25:219–230.

Holland B.R., Delsuc F., Moulton V. 2005. Visualizing conflicting evolutionary hypotheses in large collections of trees: using consensus networks to study the origins of placentals and hexapods. Syst. Biol. 54:66–76.

Holland B.R., Huber K.T., Moulton V., Lockhart P.J. 2004. Using consensus networks to visualize contradictory evidence for species phylogeny. Mol. Biol. Evol. 21:1459–1461.

Huson D.H., Bryant D. 2006. Application of phylogenetic networks in evolutionary studies. Mol. Biol. Evol. 23:254–267.

Huson D.H., Rupp R., Scornavacca C. 2011. Phylogenetic networks: concepts, algorithms and applications. Cambridge: Cambridge University Press.

Jin G., Nakhleh L., Snir S., Tuller T. 2007. Inferring phylogenetic networks by the maximum parsimony criterion: a case study. Mol. Biol. Evol. 24:324–337.

Kelchner S.A., Bamboo Phylogeny Group. 2013. Higher level phylogenetic relationships within the bamboos (Poaceae: Bambusoideae) based on five plastid markers. Mol. Phylogen. Evol. 67:404–413.

Knowles L.L., Kubatko L.S., editors. 2010. Estimating species trees: practical and theoretical aspects. Hoboken (NJ): Wiley-Blackwell.

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

Kull K. 2003. Ladder, tree, web: the ages of biological understanding. Sign Syst. Stud. 31:589–603.

Lam H.J. 1936. Phylogenetic symbols, past and present (being an apology for genealogical trees). Acta Biotheor. 2:153–194.

Lawton G. 2009. Why Darwin was wrong about the Tree of Life. New Sci. 201(2692):34–39.

Linnaeus C. 1751. Philosophia botanica. Amsterdam: Z. Chatelain.

Losos J.B., Hillis D.M., Greene H.W. 2012. Who speaks with a forked tongue? Science 338:1428–1429.

Makarenkov V., Legendre P. 2000. Improving the additive tree representation of a given dissimilarity matrix using reticulation. In: Kiers H.A.L., Rasson J.P., Groenen P.J.F., Schade M., editors. Data analysis, classification, and related methods. Berlin: Springer. p. 35–46.

Mindell D.P. 2013. The Tree of Life: metaphor, model, and heuristic device. Syst. Biol. 62:479–489.

Morrison D.A. 2010. Using data-display networks for exploratory data analysis in phylogenetic studies. Mol. Biol. Evol. 27:1044–1057.

Morrison D.A. 2011. Introduction to phylogenetic networks. Uppsala: RJR Productions.

Morrison D.A. 2013a. [Book review of] "Tree thinking: an introduction to phylogenetic biology". Syst. Biol. 62:634–637.

Morrison D.A. 2013b. Phylogenetic networks are fundamentally different from other kinds of biological network. In: Zhang W.J., editor. Network biology: theories and methods. New York: Nova Science Publishers. p. 23–68.

Nakhleh L. 2011. Evolutionary phylogenetic networks: models and issues. In: Heath L.S., Ramakrishnan N., editors. Problem solving handbook in computational biology and bioinformatics. New York: Springer. p. 125–158.

Nakleh L. 2013. Computational approaches to species phylogeny inference and gene tree reconciliation. Trends Ecol. Evol. (in press).

Nelson G.J. 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 Uni. Press. p. 105–111.

Nelson G., Platnick N. 1981. Systematics and biogeography: cladistics and vicariance. New York: Columbia University Press.

O'Hara R.J. 1988. Diagrammatic classifications of birds, 1819–1901: views of the natural system in 19th–century British ornithology. In: Ouellet H., editor. Acta XIX Congressus Internationalis Ornithologici. Ottawa: National Museum of Natural Sciences. p. 2746–2759.

O'Hara R.J. 1991. Representations of the natural system in the nineteenth century. Biol. Philos. 6:255–274.

O'Hara R.J. 1996. Trees of history in systematics and philology. Mem. Soc. Ital. Sci. Nat. Mus. Civ. Stor. Nat. Milano 27:81–88.

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. 1776. Elenchus zoophytorum. The Hague: Petrum van Cleef.

Pennisi E. 1999. Is it time to uproot the Tree of Life? Science 284:1305–1307.

Phillips M.J., Delsuc F., Penny D. 2004. Genome-scale phylogeny and the detection of systematic biases. Mol. Biol. Evol. 21:1455–1458.

Pietsch T.W. 2012. Trees of Life: a visual history of evolution. Baltimore: Johns Hopkins University Press.

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

Ren F., Tanaka H., Yang Z. 2005. An empirical examination of the utility of codon-substitution models in phylogeny reconstruction. Syst. Biol. 54:808–818.

Rieppel O. 2010. The series, the network, and the tree: changing metaphors of order in nature. Biol. Philos. 25:475–496.

Rokas A., Williams B.L., King N., Carroll S.B. 2003. Genome-scale approaches to resolving incongruence in molecular phylogenies. Nature 425:798–804.

Salichos L., Rokas A. 2013. Inferring ancient divergences requires genes with strong phylogenetic signals. Nature 497:327–331.

Semple C., Steel M. 2003. Phylogenetics. Oxford: Oxford University Press.

Shapiro B., Rambaut A., Drummond A.J. 2006. Choosing appropriate substitution models for the phylogenetic analysis of protein-coding sequences. Mol. Biol. Evol. 23:7–9.

Sneath P.H.A. 1975. Cladistic representation of reticulate evolution. Syst. Zool. 24:360–368.

St. Mivart G. 1865. Contributions towards a more complete knowledge of the axial skeleton in the primates. Proc. Zool. Soc. Lond. 33:545–592.

St. Mivart G. 1867. On the appendicular skeleton of the primates. Philos. Trans. R. Soc. Lond. 157:299–429.

Stevens P.F. 1984. Metaphors and typology in the development of botanical systematics 1690–1960, or the art of putting new wine in old bottles. Taxon 33:169–211.

Stevens P.F. 1994. The development of biological systematics: Antoine–Laurent de Jussieu, nature, and the natural system. New York: Columbia University Press.

Stolzer M, Lai H, Xu M, Sathaye D, Vernot B, Durand D. 2012. Inferring duplications, losses, transfers and incomplete lineage sorting with nonbinary species trees. Bioinformatics 28:i409–i415.

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

Teeling E.C., Hedges S.B. 2013. Making the impossible possible: rooting the tree of placental mammals. Mol. Biol. Evol. 30:1999–2000.

Tukey J.W. 1977. Exploratory data analysis. Reading (MA): Addison-Wesley.

Wägele J.W., Letsch H., Klussmann-Kolb A., Mayer C., Misof B., Wägele H. 2009. Phylogenetic support values are not necessarily

informative: the case of the Serialia hypothesis (a mollusk phylogeny). Frontiers Zool. 6:12.

Wägele J.W., Mayer C. 2007. Visualizing differences in phylogenetic information content of alignments and distinction of three classes of long-branch effects. BMC Evol. Biol. 7:147.

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

Willmann R. 2003. From Haeckel to Hennig: the early development of phylogenetics in German-speaking Europe. Cladistics 19:449–479.

Yu Y., Degnan J.H., Nakhleh L. 2012. The probability of a gene tree topology within a phylogenetic network with applications to hybridization detection. PLoS Genet. 8:e1002660.

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.