Meta-trees: grafting for a global perspective

V. A. Funk* and Chelsea D. Specht

(VAF) Department of Botany, National Museum of Natural History, Smithsonian Institution,
MRC 166, Washington, D.C. 20013-7012, U.S.A., e-mail: funkv@si.edu;
(CDS) Department of Plant and Microbial Biology, 111 Koshland Hall MC-3102,
University of California, Berkeley, California 94720-3102, U.S.A.,
e-mail: cdspecht@nature.berkeley.edu

Abstract.—Assembling phylogenetic trees for groups of organisms with thousands of taxa is problematic because of the large amount of data and trees that must be generated and analyzed. Several attempts have been made to develop a better way to handle the problems of large trees. Meta-trees (Meta Supertrees) are suggested as a way to generate a phylogeny for groups where you have many different data sets of overlapping but not identical sets of taxa and without a common set of markers. Meta-trees graft phylogenies onto a fixed base tree and avoid the problems of missing or redundant data and misplaced taxa that plague other types of supertrees.

Supertrees are constructed from novel analyses of multiple independently derived tree topologies or data matrices with overlapping but not identical taxa. They can provide an excellent summary of past work and a useful framework for examining character evolution in a broader historic and phylogenetic framework. For all types of supertrees, the major concern is how reliable they are at providing an estimate of phylogenetic relationships, or how much confidence can be placed in each of the nodes and the taxonomic associations they indicate. As with all types of analyses, the worth of the product is based on the quality of the data used and the reliability of the method.

It seems that the term was first used by Gordon (1986) but then languished until it was made popular by Sanderson et al. (1998) in conjunction with developing taxonomically broad ''trees of life.'' In the last five years, with the advent of large-scale phylogenetic programs, using a variety of sampling techniques to provide increasingly broad (but not always deep) taxonomic coverage, and the rapid acquisition of large amounts of data for sampled taxa (e.g., multiple molecular markers spanning entire genomes), the topic has taken on a new urgency. The use of supertrees is necessary to combine all available phylogenetic information on any particular group into a single framework for testing hypotheses of character evolution and biogeography among major clades.

Defining supertrees as we have above, there are three main groups: 1) those that summarize different tree topologies using clades in those topologies as coded characters for the construction of novel topological associations (*consensus trees* or just *supertrees*); 2) those that are generated from the phylogenetic analysis of a data set made up of many smaller data sets that contain overlapping but not identical sets of taxa, with a certain percentage of missing data inherent to the combined matrix (*supermatrix trees* or *supermatrices*); and 3) those that graft

---

* Corresponding author.

phylogenies onto a fixed base tree (*meta-trees*). There has been a great deal of discussion recently about the pros and cons of the first two types of supertrees (i.e., Steel et al. 2000, Gatesy et al. 2002, Goloboff & Pol 2002, Bininda-Emonds et al. 2003, etc.), and there is a relatively recent book about the construction and use of these trees edited by Bininda-Emonds (2004c). In our research we have used the newly developed meta-tree approach. There has been little discussion in print on what we are calling meta-trees and little understanding of precisely how they are constructed. Therefore, while this commentary mentions other types of supertrees in a comparative context, it focuses on the meta-trees. It is not our intention to discuss algorithms or to provide detailed examples or discussions of the various types of trees. Rather, we wish to draw attention to what we think is a problem in combining total available phylogenetic information and provide a possible solution. There is one relatively new term used here: *meta-supertree* (or *meta-tree*) which was first used by Funk et al. (2005) and describes a supertree formed by a mosaic of trees grafted onto a fixed base tree (see discussion below).

Consensus based supertrees combine topological data. They have the advantage of summarizing large numbers of phylogenies but have problems with redundant data as well as difficulty in resolving conflict among the source trees. The methods used in consensus tree construction are quite basic to tree topology examination and go back to the early 1980's. The ability to reduce trees to a matrix was first developed by Brooks (1981, 1985) and was later called Brooks Parsimony Analysis (BPA). It was further developed by Funk & Brooks (1990), using the famous fish cladograms prepared by Rosen (1978, 1979). The taxa in these trees were mostly identical.

Another type of tree produced by the combining of trees is called simply a supertree. Gordon (1986) and later workers constructed consensus trees based on the collection of supertrees that were built from partly overlapping input trees, but they were also built by consensus methods in the conventional way. In all recent literature on supertrees, the term consensus methods (or ''consensus'' setting) is used to distinguish the case of completely overlapping input trees from consensus supertree methods that consider partially overlapping input trees. This method has been applied to a number of groups, including angiosperms (Davies et al. 2004), and the method is discussed at length in regard to the mammal work by Gatsey et al. (2002, with a reply by Bininda-Emonds et al. 2003; Springer & DeJong 2001). In other words, there has been a distinction drawn between methods that combine tree topologies and are consensus-like (agreement or conservative supertrees) and those that are not (optimization or liberal supertrees) (for a full discussion see Bininda-Emonds 2004b, Wilkinson et al. 2004). Various types of consensus methods have been used to form supertrees (see Goloboff & Pol 2002 for a discussion) and many more have been recently described (ca. 20 according to Bininda-Emonds 2004b). A recent paper by Wilkinson et al. (2005) examines the future of supertrees and many papers on supertree construction and evaluation are described in Bininda-Emonds (2004c).

The sheer size of current projects both in the number of trees (e. g., 430 trees, Liu et al. 2001) and in the number of terminal taxa (over 500) compounds the problems involved in tree construction but provides primary and topological data that can be incorporated into a consensus analyses. Page (2004) does a good job of discussing this type of supertree and its practicality in contributing to our understanding of the tree of life.

Super-matrix trees are a form of supertree, although the 'super' really refers to the large and combined data matrix used

to create the final phylogenetic tree. They fit the broad definition of a supertree since they are combining data from different analyses that contain overlapping but not identical taxa.

Super-matrix trees have the advantage of being produced by primary data (rather than topological information) but they often suffer from a large percentage of missing information. Super-matrix trees such as those used by Gatsey et al. (1999, 2002, and critiqued by Bininda-Emonds 2004a) have been used for years; they are essentially a means of performing a phylogenetic analysis for a group of organisms where all of the data are available for some of the taxa and some of the data are available for all of the taxa (or an overlapping subset thereof). The methods used to deal with large amounts of missing data were discussed extensively in the 1980's because the problems were (and still are) common in paleobiology and morphological datasets. Super-matrix analyses are conventional character-based phylogenetic analyses on a more taxonomically inclusive scale and with varying amounts of missing data.

Once again the size of current matrices is often large and many of the problems inherent to such data sets are revisited (Goldstein & Specht 1998) even with the advent of increasingly rapid heuristic searches. Recent studies in animals (e.g., mammals, Gatsey et al. 1999, 2002) and plants (i.e., basal angiosperms, Qiu et al. 2005) use 17 data sets with 1400 informative characters and 9 genes for 100 taxa, respectively.

Super-matrix trees have been criticized for ignoring the effects of large sections of missing character data, although simulation studies have been successful in investigating the effect of missing data under a variety of circumstances (reviewed in Sanderson & Driskell 2003). The effects of missing character data on cladistic analyses have been examined at various times in the history of phylogenetic tree building as characters not applicable to certain taxa become prominent for phylogenetic reconstruction (Dunn et al. 2003, Nixon 1996, Nixon & Davis 1991, Platnick et al. 1991; Wilkinson 1995a, 1995b; see Wiens 2006 for review) and such issues are not insurmountable.

Meta-trees (tree of trees) or meta-supertrees, are trees for which "branches" or groups of branches (i.e., clades) are grafted onto a "base tree" or "backbone" phylogeny, where the base tree comprises a subset of all taxa to be included in the meta-tree. In essence, one is taking shallow, densely sampled trees and grafting them onto deep, sparsely sampled ones. They are different from other supertrees both in what they are and how they are built. The base tree is constructed from primary character data, analyzed in an appropriate phylogenetic context, with taxon sampling that fully spans the breadth of the taxonomic question to be answered by the meta-tree and includes any taxa that are known a priori to be of uncertain phylogenetic affiliation. The trees grafted on to the base tree are more densely sampled and use character data appropriate to the level of the phylogenetic question, often at the species level. Character data for one grafted tree may also be used in a second grafted tree, but statements of homology across the sampled characters for the taxa in each individual graft may be difficult due to the large span of phylogenetic distance.

Why are meta-trees necessary? Would it not be better to have a large super-matrix to link all taxa studied into a single analysis? In an ideal world we would have all the data for all the taxa, or at least most of the data for all the taxa, and be able to produce the definitive tree for the group. Unfortunately, we are unlikely to have that any time soon. Despite the increased speed and decreased cost of molecular work, an overlapping dataset

for large groups such as the flowering plant family Compositae (25,000 species and 1700 genera) will be out of reach for super-matrix trees for some time, for reasons of sequencing costs and time as well as the difficulty of acquiring complete data sets that are both informative and alignable (with discernable homology) over such a large evolutionary scale. Combining topologies is also out of the question for many groups because there are few studies that share overlapping taxa. In order to assemble the tree of life, we must develop more immediate solutions that will accurately reflect the phylogeny of such groups.

Although the concept had been discussed before, the development of methods to produce meta-trees as well as their actual use are rather new because it is only recently that we have had access to large numbers of well-resolved trees for different parts of the same group. Meta-trees are constructed in the following manner. First, establish a stable and comprehensive base tree, one that includes several exemplar taxa to represent each well-known (and well-supported) clade, a strong sampling of taxa from the basal clades, and any individual taxa that are traditionally difficult to place (problem taxa). The problem taxa are usually genera with 1–few species that have been kicked out of all known monophyletic groups. The base tree is usually a phylogeny with primary character data and, in the best scenario, with little to no missing data. The validity of the meta-tree depends on the accuracy of the topology of the base tree. Typically, the data used to produce a base tree would involve more conservative regions of the genome. While developing the base tree it is necessary to continue to add molecular markers until the tree is well resolved.

The problem taxa are those that have been removed from existing well supported clades based on molecular data and not picked up in any other higher taxon. For instance, a genus might be removed from a tribe so that the tribe could be monophyletic; however, after it was removed it may not have been included in any other tribe. Often, in the past, these problem taxa have been placed in two or more locations (using morphological data) and rejected from all (using morphological and molecular data). It is hoped that in the comprehensive "base tree" analysis they will find a home. At the very least they should associate with certain areas of the base tree so that they can be included in detailed analyses. Since nothing in the base tree has a 'fixed' location the problem taxa are free to associate with any clade.

Figure 1 provides a hypothetical example of a base tree with 26 terminal taxa. For the purpose of this discussion let us say that this cladogram represents a family. Clades A–F are recognized sub-family groups, such as tribes, and P1–P4 are problem taxa. The analysis of the molecular data shows that tribes A, B, D, and F remain monophyletic. Tribe C contains one of the problem taxa (P3) and tribe D has one of the problem taxa (P2) as a sister group to the clade. Problem taxon P4 is the sister group to tribes A–D. Tribe E, which was previously thought to be a clade, is paraphyletic. In most meta-tree analyses, the individual clades are each examined by separate groups of researchers. In this example the scientists working on tribes C and D would be informed that one of the problem taxa had been placed in or near their group, and they would be asked to rerun their analyses to accommodate the taxon. In the case of group E the development of the base tree will have to include as many taxa as possible from tribe E in order to determine how many clades are involved and how they are related to one another. As with all cladistic analyses, the building of meta-trees involves reciprocal illumination (Hennig 1966).
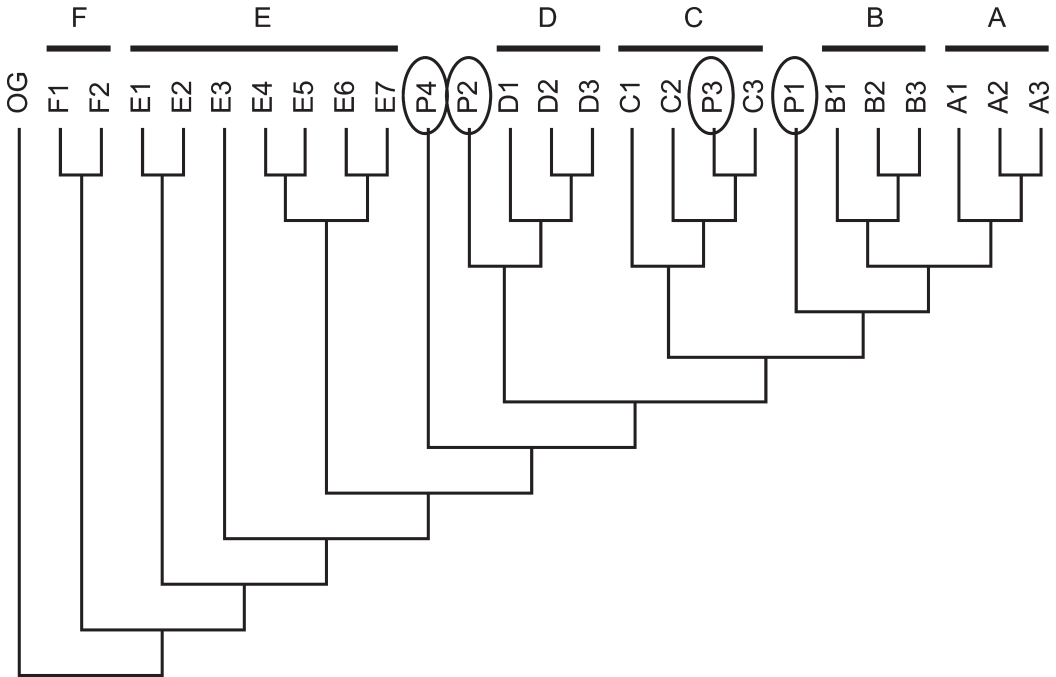
Fig. 1. Hypothetical base tree for a family of organisms. Letters indicate sub-family groups, such as tribes, that were believed, at the start of the study, to be monophyletic. To form a meta-tree, or super meta-tree, each tribe shown to be monophyletic would be replaced with cladogram from an analysis with many terminal taxa. P = problem taxa.

Once the base tree has been constructed, branches representing well-studied and well-supported clades are grafted onto the base tree in place of each clade of exemplars. How one selects the tree to be grafted onto the base tree must be explicitly stated in any publication. These grafts are individual tree topologies that represent more restricted taxonomic groups than the entire meta-tree analysis but with more detailed sampling than can be undertaken during the base tree construction. Each terminal taxon should be included in only one analysis unless it has to be divided because it is not monophyletic (e.g., a genus that has members in two different clades). Often, rapidly evolving genes or gene regions used for reconstructing the topology of the grafted branches (which are typically genus or species-level analyses) are difficult to align across the scale of questions addressed in the meta-tree, while the genes used in constructing the base tree would give little resolution at the level of the grafted analyses. The use of grafting reduces the need to make difficult, or impossible, homology statements for rapidly evolving genes across large-scale evolutionary distances.

In Figure 1, the exemplar taxa holding the position of monophyletic groups (clades A–D and F) would be removed and the larger, more detailed cladograms of the independent study would be inserted. For instance, Figure 1 could represent the flowering plant family Compositae and the three taxa of tribe A could be replaced with a cladogram of 200 taxa that have been analyzed with numerous outgroups and is therefore believed to be monophyletic. Individual clades of group E would be treated the same way if detailed analyses were available for the individual clades. The final meta-tree would have the structure of the

base tree with the details of the various analyses of the monophyletic clades (see Funk et al. 2005 for an example).

The base tree and each of the clades that is grafted onto the base tree will have available tree statistics (CI, RI, bootstrap, etc.) and will be constructed with the character data that best answers the question at hand. Using the grafting approach enables the researcher to use phylogenetically relevant information at different scales, increasing the potential to recover well-supported phylogenies of each monophyletic lineage, and to combine the well-supported results into a single topology that can be used to investigate broad evolutionary questions. Each grafted branch maintains the original support values and phylogenetic character data, unlike the construction of consensus trees and supertrees where support values and branch lengths are lost with the construction of a topology-coding matrix. It is important to note that the branch length information from the individual clades, once combined into a meta-tree, is not comparable among lineages or across the entire meta-tree. In fact, meta-trees should be drawn as regular cladograms and not phylograms. However, one can retain the information within each lineage in separate small diagrams in order to show relative support for the individual topologies and for identifying areas where more detailed character or taxon sampling could aid in strengthening evolutionary questions tested by the meta-tree topology. Branch lengths, while not completely comparable, can also begin to provide an idea of where (topologically) and when (temporally) rates of speciation may have shifted across the topology, and symmetry of branching patterns can be used to investigate locations of overall shifts in diversification rates at the topological level (Moore et al. 2004).

When a clade has multiple trees with overlapping but not identical taxa, a branch of exemplars may be replaced with a super-matrix tree or a consensus supertree, depending on the data that are available.

The resulting tree is actually a 'tree of trees' and therefore a *meta-tree*; however, because it can contain overlapping but not identical taxa, it may also be called a meta-supertree. Meta-trees allow one to produce a reliable phylogeny for many more taxa than is possible with super-matrix trees and potentially develop a phylogenetic hypothesis with greater nodal support across the topology. Funk et al. (2005) made use of a base tree that was a standard phylogeny with ca. 125 taxa and 14,000 base pairs (from Panero & Funk 2002), and the meta-tree had nearly 600 taxa (an updated version of the tree has over 700 taxa). Specht and Kress have used a base tree with ca. 2000 base pairs for 150 taxa (unpublished data) to examine the relationships among 750 taxa. The increased taxonomic breath clearly benefits studies of character evolution, interspecies interactions, and biogeography, among other areas of inquiry.

Although not as commonly discussed as the other types of supertrees, the meta-supertree does have some history. Mishler (1994) described 'compartmentalization' as a way to help analyze large clades with small amounts of data by creating hypothetical ancestors as place holders for well-defined clades and using them as terminal taxa. The compartmentalization method has been used recently by Sanchez-Baracaldo et al. (2005). Sanderson et al. (1998) also mention 'compartmentalization,' but they never take the next step of re-attaching the original cladograms to the resulting tree of the higher-level analysis, nor do they actually construct and use these trees. In addition, 'compartmentalization' makes use of hypothetical ancestors as place holders, and the base tree for a meta-tree uses sampled exemplars. The term 'composite trees' has been used for trees similar to meta-trees;

Weiblen et al. (2000) used such assemblages to examine dioecy in monocotyledons. However, we prefer the term meta-tree because it better describes the method and because one us (VAF) works on the flowering plant family Compositae (i.e., 'composites,' and a composite composite tree is an awkward term at best; also, considering that one-tenth of all flowering plants belong to the composites, the term ''composite trees'' is botanically confusing). Perhaps more important scientifically, our methods differ from those of Weiblen et al. (2000) in that we selected trees to graft onto the base tree that were all constructed using molecular data, and these data were analyzed using similar if not identical methods of phylogenetic reconstruction. But both the Weiblen et al. (2000) and the Funk et al. (2005) analyses had explicitly stated methods that were used to develop the described supertree. Finally, meta-trees as proposed here are different from the large-scale tree of Purvis (1995), which was composed of hierarchically nested trees formed by combining topologies (supertrees). That same holds for many other large-scale mammalian supertrees (Bininda-Emonds et al. 1999, Jones et al. 2002, Grenyer & Purvis 2003) that use combined methods different from those proposed here.

Gatesy et al. (2002) dismissed 'composite trees' as a class by listing them under ''Appeals to Authority'' and by stating that they were ''…not assembled according to specific rules.'' This classification is inaccurate; Weiblen et al. (2000), as well as the latter publication by Funk et al. (2005), had a specific way of constructing their trees that could be easily understood and repeated if desired. One should not condemn a method in general because of the way one particular author constructed the tree (the criticism by Gatesy et al. 2002 of Sanderson et al. 1998). As meta-trees develop there will no doubt be many different types just as there are different types of consensus trees and super-matrix trees. As with any combined or consensus analysis, it is imperative that the authors of new meta-trees be specific about the source of the data and the method(s) of analysis.

Meta-trees provide effective and efficient summaries of phylogenetic data for large groups where it is currently difficult or impossible to look at the whole taxon using traditional super-matrix or consensus methods. However, each type of supertree has its strengths and weaknesses. What type of supertree to use will be driven by the type of data that are available or that can be collected and the type of question driving the collection of such data. As with all science, one considers the options by checking the pros and cons of various methods and then selects the one (or more) that is most appropriate for the data. Certainly as we move toward building a tree of life, many such methods will have to be attempted and evaluated as to their reliability and predictability. One advantage of the meta-tree method that we propose here is that it allows systematic biologists to develop large scale trees without getting buried in algorithms or often inappropriately-applied statistics, utilizing maximally their knowledge of the data at hand and the taxa in question. They also solve several problems that we believe plague large scale studies. First, meta-trees allow us to get around the different rates of change within different markers; we need faster markers for species and generic level analyses and more conservative ones for among genera, tribes, and subfamilies. Mixing the two in the same analysis often means you can not align the fast evolving sequences on the super-generic level while simultaneously much sequencing effort and money is wasted acquiring identical or nearly identical sequences of the slowly evolving sequences at the species level. Second, meta-trees allow us to avoid the problem of thousands of equally parsimonious trees that result from a dataset containing many closely related taxa where the noise of

sequences with low divergence can confound the phylogenetic signal. Finally, the use of meta-trees allows the use of many programs that simply will not run (at least in any reasonable time frame) for 700 taxa. Rather than running a substandard analysis that can be done in a reasonable time frame, the meta-trees approach enables the combination of analyses that center around appropriate data collection and analyses at each level of the supertree. We feel that the use of meta-trees ultimately allows the scientist to focus her/his efforts on evaluating topics of fundamental evolutionary importance, such as character evolution, biogeography, and plant-animal interactions.

Acknowledgments

Literature Cited

Bininda-Emonds, O. R. P., J. Gittleman, & A. Purvis. 1999. Building large trees by combining phylogenetic information: A complete phylogeny of the extant Carnivora (Mammalia).—Biological Review 74:143–175.

———. et al. 2003. Supertrees are a necessary not so evil: A comment of Gatesy et al.—Systematic Biology 52:724–729.

———. 2004a. Trees versus characters and the supertree/supermatrix "Paradox".—Systematic Biology 53:356–359.

———. 2004b. The evolution of supertrees.—Trends in Ecology and Evolution 19:315–322.

——— (ed.). 2004c. Phylogenetic Supertrees: Combining Information to Reveal the Tree of Life. Kluwer Academic, Dordrecht, the Netherlands, 550 pp.

Brooks, D. R. 1981. Hennig's parasitological method: a proposed solution.—Systematic Zoology 30:229–249.

———. 1985. Historical ecology: A new approach to studying the evolution of ecological associations.—Annals of the Missouri Botanical Garden 72:660–80.

Davies, T. J., T. G. Barraclough, M. W. Chase, P. S. Soltis, D. E. Soltis, & V. Savolainen. 2004. Darwin's abominable mystery: insights from a supertree of the angiosperms.—Proceedings of the National Academy of Science 101:1904–1909.

Dunn, K. A., J. D. McEachran, & R. L. Honeycutt. 2003. Molecular phylogenetics of myliobatiform fishes (Chondrichthyes: Myliobatiformes) with comments on the effects of missing data on parsimony and likelihood.—Molecular Phylogenetics and Evolution 27:259–270.

Funk, V. A., & D. R. Brooks. 1990. Phylogenetic systematics as the basis of comparative biology.—Smithsonian Contributions to Botany 73:1–45.

———. et al. 2005. Everywhere but Antarctica: Using a supertree to understand the diversity and distribution of the Compositae. Pp. 343–373 in I. Friis and H. Balslev, eds., Proceedings of a Symposium on Plant Diversity and Complexity Patterns - Local, Regional and Global Dimensions. —Biologiske Skrifter, vol. 55.

Gatsey, J., M. Milinkovitch, V. Waddell, & M. Stanhope. 1999. Stability of cladistic relationships between Cetacea and higher-level artiodacty taxa.—Systematic Biology 48:6–20.

———, C. Matthee, R. DeSalle, & C. Hayashi. 2002. Resolution of a supertree/supermatrix paradox.—Systematic Biology 51:652–664.

Goldstein, P. Z., & C. D. Specht. 1998. Pitfalls in phylogenetic analysis of large molecular data sets. Pp. 253–274 in R. DeSalle and B. Schierwater, eds., Molecular Approaches to Ecology and Evolution. Birkhauser, Verlag, Basel.

Goloboff, P. A., & D. Pol. 2002. Semi-strict supertrees.—Cladistics 18:514–525.

Gordon, A. G. 1986. Consensus supertrees: the synthesis of rooted trees containing overlapping sets of labeled leaves.—Journal of Classification 3:335–348.

Grenyer, R., & A. Purvis. 2003. A composite species-level phylogeny of the 'Insectivora' (Mammalia: Order Lipotyphla Haeckel, 1866).—Journal of Zoology 260:245–257.

Hennig, W. 1966. Phylogenetic Systematics. University of Illinois Press, Champaign-Urbana.

Jones, K. E., A. Purvis, A. MacLarnon, O. R. P. Bininda-Emonds, & N. B. Simmons. 2002. A phylogenetic supertree of the bats (Mammalia: Chiroptera).—Biological Review 77:223–259.

Liu, F. et al. 2001. Molecular and morphological supertrees for eutherian (placental) mammals.—Science 291:1786–1789.

Mishler, B. D. 1994. Cladistic analysis of molecular and morphological data.—American Journal of Physical Anthropology 94:143–156.

Moore, B. R., K. M. A. Chan, & M. J. Donoghue. 2004. Detecting diversification rate variation in supertrees. Pp. 487–533 in O. Bininda-Emonds, ed., Phylogenetic Supertrees: Combining Information to Reveal the Tree of Life. Kluwer Academic, Dordrecht, The Netherlands.

Nixon, K. C. 1996. Paleobotany in cladistics and cladistics in paleobotany: enlightenment and uncertainty.—Review of Palaeobotany and Palynology 90:361–373.

———, & J. I. Davis. 1991. Polymorphic taxa, missing values and cladistic analysis.—Cladistics 7:233–241.

Page, R. 2004. Taxonomy, supertrees, and the tree of life. Pp. 247–265 in O. R. P. Bininda-Emonds, ed., Phylogenetic Supertrees: Combining Information to Reveal the Tree of Life. Kluwer Academic Publishers, The Neatherlands.

Platnick, N. I., C. E. Griswold, & J. A. Coddington. 1991. On missing entries in cladistic analysis.—Cladistics 7:337–343.

Panero, J., & V. Funk. 2002. Toward a phylogenetic subfamilial classification for the Compositae (Asteraceae).—Proceedings of the Biological Society of Washington 115:909–922.

Purvis, A. 1995. A composite estimate of primate phylogeny.—Philosophical Transactions of the Royal Society of London B 348:405–421.

Qui, Y.-L. et al. 2005. Phylogenetic analyses of basal angiosperms based on nine plastid, mitochondrial, and nuclear genes.—International Journal of Plant Sciences 166:815–842.

Rosen, D. E. 1978. Vicariant patterns and historical explanation in biogeography.—Systematic Zoology 27:159–188.

———. 1979. Fishes from the uplands and intermontane basins of Guatemala: Revisionary studies and comparative biogeography.—Bulletin of the American Museum of Natural History 162:267–376.

Sanchez-Baracaldo, P., P. K. Hayes, & C. E. Blank. 2005. Morphological and habitat evolution in the Cyanobacteria using a compartmentalization approach.—Geobiology 3:145–165.

Sanderson, M. J., A. Purvis, & C. Henze. 1998. Phylogenetic supertrees: assembling the trees of life.—Trends in Ecology and Evolution 13:105–109.

———, & A. C. Driskell. 2003. The challenge of constructing large phylogenetic trees.—Trends in Plant Science 8:374–379.

Springer, M. S., & W. W. DeJong. 2001. Which mammalian supertree to bark up?—Science 291:1709–1711.

Steel, M., A. W. M. Dress, & S. Böcker. 2000. Simple but fundamental limitations on supertree and consensus tree methods.—Systematic Biology 49:363–368.

Weiblen, G. D., R. K. Oyama, & M. J. Donoghue. 2000. Phylogenetic analysis of dioecy in monocotyledons.—The American Naturalist 155:46–58.

Wiens, J. J. 2006. Missing data and the design of phylogenetic analyses.—Journal of Biomedical Informatics 39:34–42.

Wilkinson, M. 1995a. Coping with abundant missing entries in phylogenetic inference using parsimony.—Systematic Biology 44:501–514.

———. 1995b. Arbitrary resolutions, missing entries, and the problem of zero-length branches in parsimony analysis.—Systematic Biology 44:108–111.

———, J. L. Thorley, D. Pisani, F.-J. Lapointe, & J. O. McInerney. 2004. Some desiderata for liberal supertrees. Pp. 227–246 in O. R. P. Bininda-Emonds, ed., Phylogenetic Supertrees: Combining Information to Reveal the Tree of Life. Kluwer Academic, Dordrecht, The Netherlands.

———. et al. 2005. The shape of supertrees to come: tree shape related properties of fourteen supertree methods.—Systematic Biology 54:419–431.

Associate Editor: Carol Hotton