# NEW USES FOR OLD PHYLOGENIES

## *An introduction to the volume*

Olaf R. P. Bininda-Emonds

## 1. Introduction

> "This is a paper with an attitude problem. This may sound facetious, but is meant in all seriousness. It has in my opinion entirely the wrong attitude to phylogenetic reconstruction and indeed to the entire scientific process."
>
> From an anonymous review of the carnivore supertree of Bininda-Emonds *et al.* (1999)

What are supertrees and what is all the fuss about?

These are two of the questions that this volume will attempt to answer. A brief answer to the former is that supertree construction is a phylogenetic approach that combines tree topologies instead of the primary character data that they are based on. It differs from traditional consensus techniques, which also combine tree topologies, in that the constituent (or "source") trees need only be overlapping, and not identical, with respect to the terminal taxa they contain. As such, the resulting supertree can be, and usually is, larger than any of the source trees contributing to it. Supertrees thus represent an exciting opportunity to build more comprehensive phylogenies: in essence, new uses for old phylogenies (with apologies to Harvey *et al.*, 1996). However, the use of tree topologies and not primary character data as the source data has attracted much criticism (e.g., Rodrigo, 1993; Slowinski and Page, 1999; Novacek, 2001; Springer and de Jong, 2001; Gatesy *et al.*,

*Bininda-Emonds, O. R. P. (ed.) Phylogenetic Supertrees: Combining Information to Reveal the Tree of Life, pp. 3–14. Computational Biology, volume 3 (Dress, A., series ed.).*

2002), such that supertree construction is an increasingly popular, but highly controversial approach in phylogenetic systematics.

Although supertree construction has attracted increasing attention only in the past few years, the fundamental idea behind it — that of combining numerous source trees to yield a single, more inclusive tree — has a longer, if unrecognized, history. The process of synthesizing systematic knowledge by cutting and pasting together evolutionary trees as "informal" supertrees is probably nearly as old as systematics itself. Even today, any detailed depiction of the single Tree of Life (e.g., the Tree of Life Web Project; http://tolweb.org/tree/phylogeny.html), if not any conception that we have of it, can be achieved only using supertrees; the largest phylogenies based on primary character data are on the order of thousands of species only (e.g., Källersjö *et al.*, 1998; Johnson, 2001). Although informal supertrees continue to be constructed (e.g., Garland *et al.*, 1993; Kennedy *et al.*, 1996; Ortolani, 1999; Webb, 2000; Cardillo and Bromham, 2001; Hall and Harvey, 2002), this volume deals exclusively with the more formal supertree construction techniques.

The formalization of supertree methodology, and the term supertree itself, stem from Allan Gordon's (1986) seminal paper. In this paper, Gordon described the supertree equivalent of strict consensus, whereby the supertree contained only those groups found on or implied jointly by all the source trees. However, it did not have much of an immediate impact for several reasons. First, the paper was published in a mathematical journal, whereas the current popularity of supertrees arguably derives largely from the biological community. Second, the current interest in supertrees derives in large measure from their ability to build very large phylogenies of hundreds of species, something that has become computationally feasible within only the past decade at best. Finally, Gordon's method was limited to overlapping source trees that were compatible: they could differ from one another, but not actually conflict. Thus, the method was of limited utility. As most systematists know, phylogenies usually conflict with one another.

The breakthrough for supertrees came in 1992, when Bernard Baum, Jeff Doyle, and Mark Ragan independently described the supertree technique know known as simply MRP (matrix representation with parsimony; Baum, 1992; Doyle, 1992; Ragan, 1992). MRP, like Brook's Parsimony Analysis (Brooks, 1981), makes use of additive binary coding (Farris *et al.*, 1970) to represent a given tree in matrix format. The "matrix representations" of the different source trees are then combined into a single matrix that can be analyzed using any desired optimization criterion (but usually parsimony). This procedure removed the fundamental limitation of Gordon's strict supertree method: all overlapping source trees could now be combined as a supertree, regardless of how much they conflicted with one another. At the same time, these trees could be derived from all possible data types (including no data whatsoever!), overcoming the limitation of combined-data ("total evidence" or "supermatrix") approaches (*sensu* Kluge, 1989; Sanderson *et al.*, 1998; respectively) that the data types be analyzable using a single optimization criterion.

It was Andy Purvis who perhaps first realized the tremendous potential of supertrees to biology. His MRP supertree of all 203 extant species of primate represented the first, complete (super)tree of a significant clade that was based on an objective methodology (Purvis, 1995a). Its large size, its unprecedented completeness, and its high amount of resolution demonstrated what supertree construction could achieve. Moreover, Purvis showed immediately that supertrees have biological utility beyond their obvious systematic value when he used his primate supertree to answer numerous macroevolutionary questions in a phylogenetic framework (Purvis *et al.*, 1995). The primate supertree has gone on to become perhaps the reference standard for supertrees. It has been updated twice (Purvis and Webster, 1999; Vos and Mooers, in prep.), and is often the tree against which new methodologies are tested (e.g., Moore *et al*., 2004; Vos and Mooers, 2004).

However, in some ways, Purvis caught the phylogenetic community unawares. The next major supertree, that of the mammalian order Carnivora that I published (Bininda-Emonds *et al.*, 1999), took another four years to be published, largely as a result of hostile reviews (see above). An evolutionary journal thought that the carnivore supertree was "too taxonomic". A taxonomic journal thought the reverse: it was "too evolutionary". The supertree was published eventually in a review journal, despite arguably containing very little explicit review material. Many other supertrees studies have faced equally difficult routes to publication and critiques of the supertree approach are appearing more frequently (see above).

Today, supertree construction is an active field of theoretical, practical, and applied research in mathematics, algorithmics, computer sciences, and biology. This multidisciplinary, bioinformatic nature to supertree research has produced numerous advances and developments in a short time: supertrees are being constructed at an increasing pace, and supertree methods continue to be developed and improved. Supertree construction is also mentioned increasingly as perhaps being a key element in our efforts to reconstruct the Tree of Life (e.g., Soltis and Soltis, 2001; Pennisi, 2003). It has certainly illuminated some of the largest portions of the Tree to date, and will continue to do so for some time to come. However, in so doing, the role of supertrees might change from simply combining existing information to being an important analytical tool to search large character matrices efficiently (Bininda-Emonds *et al.*, 2002). Supertrees, like the species they depict, are also continuing to evolve.

## 2. Structure of the volume

This volume is divided roughly, if not somewhat arbitrarily, into the following five sections: existing supertree methods, new supertree methods, methodological considerations, a critical look at supertrees, and supertrees and their applications. Here, I examine each of these sections, and introduce briefly the chapters within them, in turn.

### 2.1 Existing supertree methods

Together with the following section, this is the largest component of the book, reflecting the plethora of supertree methods that have been and continue to be developed. More than a dozen major methods, and numerous variants on these methods, exist currently (Table 1). This section provides reviews of most of the current methods. Only reviews of Gordon's strict supertrees (Gordon, 1986), semi-strict supertrees (Goloboff and Pol, 2002), and (modified) MINCUTSUPERTREE (Semple and Steel, 2000; Page, 2002) are lacking, although the latter does make an appearance in several chapters in the book.

The first chapter, perhaps fittingly, is about MRP, by far the most popular of the many supertrees techniques. In this chapter, Bernard Baum and Mark Ragan discuss their motivations in developing MRP, review several MRP-related issues, and argue strongly for a continuing role of MRP in large-scale phylogeny reconstruction. But, as Baum and Ragan (1993) themselves have pointed out, parsimony is not the only option for analyzing representations of source trees.

The two following chapters review alternative methods for analyzing matrix representations of source trees. First, Howard Ross and Allen Rodrigo explore an idea first raised by Purvis (1995b) nearly a decade ago (and followed up by Rodrigo, 1996; Pisani, 2002): using compatibility instead of parsimony to analyze the matrix representations of the source trees. Gordon Burleigh and colleagues then review the concept of (minimum) flip supertrees, in which analysis proceeds not by optimizing the combined matrix representations, but by altering ("flipping") individual cells in the matrix so as to remove any conflict between them.

The section concludes with two chapters about long-existing, but perhaps unappreciated supertree methods: the average consensus and gene tree parsimony. François-Joseph Lapointe and Claudine Levasseur discuss the natural extension of the average consensus (Lapointe and Cucumel, 1997) to the supertree setting. This method is notable because it uses an alternative form of matrix representation based on path-length distance matrices rather than an MRP-like membership criterion. As such, it can maintain and utilize

*Table 1.* The major supertree methods and their variants. The methods are subdivided according to whether they produce a supertree that either summarizes common structure among the source tree ("agreement supertrees") or maximizes the fit to the set of source trees according to some objective function ("optimization supertrees"). Methods in bold face are either reviewed or introduced in this volume.

| Agreement supertrees | Optimization supertrees |
|---|---|
| Gordon's strict | **Average consensus (also known as matrix representation with distances, MRD)** |
| MinCutSupertree, including:<br>modified MinCutSupertree | **Bayesian supertrees** |
| **RankedTree** | **Gene tree parsimony** |
| **Semi-labelled- and ancestralBuild** | **Matrix representation with compatibility (MRC)** |
| Semi-strict | **Matrix representation with flipping (MRF; also known as MinFlip supertrees)** |
| **Strict consensus merger** | **Matrix representation with parsimony (MRP)**, including:<br>Purvis sister-group coding<br>Irreversible MRP |
|  | **Quartet supertrees** |

branch-length information in the source trees unlike most supertree methods. Likewise, the extension to the supertree setting of gene tree parsimony — a technique developed originally to reconcile conflicting phylogenies in co-evolutionary, host-parasite, biogeography, or gene family evolution studies (Slowinski and Page, 1999) — is described by James Cotton and Rod Page.

## 2.2 New supertree methods

An active area in supertrees is the continuing development of different methods. Of the methods listed in Table 1, the majority date from the past few years only. This includes four new methods described for the first time in this volume.

Three existing supertree methods — Gordon's strict method, MinCutSupertree, and modified MinCutSupertree — all rely on the Build algorithm of Aho *et al.* (1981), which, interestingly, was developed for other purposes entirely (namely, relational databases!) and actually predates the formalization of the supertree approach. Two of the new methods in this section derive from further modifications of Build. First, David Bryant and colleagues describe RankedTree, an algorithm that is able to incorporate both relative and absolute dating information from the source trees so as to directly produce a supertree with divergence-date estimates. Then, Philip Daniel and Charles Semple provide a solution to a problem raised by Rod Page in a chapter that appears later in the book, and introduce modifications to BUILD that allow sets of source trees with nested terminal taxa to be combined. Their algorithms make use of the fact that not only the terminals, but also the nodes in a phylogeny might be labeled.

Although the idea of supertree methods based on quartets has been raised previously (e.g., Thorley and Page, 2000; Pisani and Wilkinson, 2002), Raul Piaggio-Talice and colleagues provide one of the first working descriptions of a quartet-based supertree method. In their chapter, they explore the performance of two quartet-supertree methods that they based on the character-based quartet methods of Stephen Willson (1999, 2001). The final chapter by Fredrik Ronquist *et al.* continues the expansion of Bayesian methodology into evolutionary biology by introducing Bayesian supertrees. Interestingly, the heart of Bayesian supertrees is the same matrix representation of supertrees used by MRP and several other related methods. As it turns out, the matrix representations provide an excellent summary of the structure of a tree (in the form of taxon bipartitions) that translates well to the Bayesian framework.

## 2.3 Methodological considerations

In supertree construction, as in conventional phylogenetics, there are always questions about how to apply or expand upon those methods that do exist. Some of these issues are dealt with directly when developing a supertree method, but many more general ones still exist. This section examines a handful of some of these many issues.

The previous two sections in the book and Table 1 together indicate that supertree methods, each with slightly different properties, abound. But, what properties should a (good) supertree method have? Taking their cue from the axiomatic approach that is common in the mathematical literature (e.g., McMorris and Neumann, 1983; Barthélemy *et al*., 1995; Steel *et al*., 2000), Mark Wilkinson and colleagues propose a list of desirable features ("desiderata") that are based on the goals of accuracy and practicality. They then attempt to characterize the many liberal (or optimization) supertree techniques for these desiderata.

There then follow two chapters that look more closely at the raw data of a supertree analysis, the source trees themselves. In the first chapter, Rod Page poses numerous questions and challenges for supertree researchers that derive from taxonomic considerations. As mentioned, one problem was solved subsequently by Daniel and Semple (but appears in a previous chapter). Page also presents several answers himself, in particular using the concept of a classification graph to potentially increase the degree of taxonomic overlap between source trees. Then, together with several members of the "Mammal SuperTeam", I outline the protocol for source-tree collection and manipulation that we established as part of our efforts to construct a supertree of most extant mammalian species. We hope that our protocol, suitably adjusted for the supertree project in question, will help other researchers in constructing their supertrees.

An active area of research in conventional, data-based phylogenetics is the development of methods to infer divergence times for phylogenies from DNA sequence information (e.g., Rambaut and Bromham, 1998; Thorne *et al.*, 1998; Huelsenbeck *et al.*, 2000a; Yoder and Yang, 2000; Sanderson, 2002; Thorne and Kishino, 2002). Previous efforts in this area for supertrees have been less formal (e.g., Purvis, 1995a; Bininda-Emonds *et al.*, 1999). Rutger Vos and Arne Mooers address this deficiency and describe a method to fit available DNA sequence data to a supertree topology. Using their method of "gene shopping" and "taxon shopping", they infer divergence dates for a new update to the primate supertree, and compare their estimates to those from Purvis's (1995a) original supertree of the Primates.

Usman Roshan and colleagues then provide the exception to the title of this introduction (if not the book), and describe what might well be the future of supertree research. Rather than use supertrees to piece together previously existing information, Roshan *et al.* explore the idea of supertrees as part of a divide-and-conquer strategy, in which a large data matrix is broken down and analyzed as easier subproblems, which are then recombined using supertrees to give the global answer. A similar strategy is provided also by Gordon Burleigh and colleagues in their earlier chapter on MRF supertrees, in which their biclique approach is used to decompose a large, incomplete data matrix into smaller, complete submatrices, the results of which can be combined with supertrees.

## 2.4 A critical look at supertrees

As I indicated above, supertrees are not an uncontroversial area of research. Interestingly, however, the harshest criticisms of the supertree approach originate almost exclusively from the phylogenetic systematics community (see above), a group for which supertrees might be thought to hold the greatest immediate benefit. Perhaps connected with this fact, it remains that few major supertree analyses have been published in and of themselves in one of the leading systematic or taxonomic journals. The exception is the genus-level grass supertree of Salamin *et al.* (2002), which was published in *Systematic Biology*, but had a strong methodological component to it. This section casts a critical eye at supertrees, indicating perhaps that there is still room for improvement in this young field.

The first chapter in this section is not so much a critique of the supertree approach, but rather serves to point out some fundamental limitations when attempting to build unrooted supertrees. These problems are perhaps underappreciated by biologists, who tend to work with rooted trees; building rooted supertrees, as it turns out, faces far fewer such problems. In his chapter, Sebastian Böcker reviews some of these known limitations, but also reveals instances under which it is possible to build an unrooted supertree efficiently.

By virtue of being the most popular supertree method, MRP is also the one that has attracted the most direct criticism. (In fact, many of the alternative supertree methods have been developed to address perceived shortcomings with MRP.) The remaining chapters in this section continue this critical examination of MRP, each from a slightly different perspective.

In the biological systematics community, parsimony is linked intimately with cladistics. Consequently, there is a natural tendency to associate MRP and cladistics, and several critiques of MRP derive from attempting to interpret the method and its results in a cladistic framework. Harold Bryant considers this relationship in more detail, and examines explicitly how well MRP meets the assumptions of cladistic analysis.

Supertree construction is but one approach to combine existing phylogenetic information to derive more comprehensive phylogenies. Another is the direct combination of the primary character data, known variously as the supermatrix (*sensu* Sanderson *et al.*, 1998) or total evidence approach (*sensu* Kluge, 1989). In their chapter, John Gatesy and Mark Springer contrast these two approaches to expand on their previous criticisms of both MRP and the supertree approach as a whole (Springer and de Jong, 2001; Gatesy *et al.*, 2002).

Finally, David Williams examines the representation of source tree topologies, not as binary characters as is usually done, but as three-item statements encoding relationships within the tree. Intriguingly, he suggests that any undesirable aspects to MRP might derive not only from its use of what he holds to be an inferior form of matrix representation (i.e., binary coding), but, more importantly, from a fundamental shortcoming in parsimony optimization itself.

## 2.5 Supertrees and their applications

The final section departs from the methodological focus of the previous chapters to examine (biological) applications of supertrees. In recent years, the importance of examining biological questions in a phylogenetic framework has become increasingly appreciated. As phylogenetic hypotheses, supertrees can be used like any other conventionally derived phylogeny. However, the ability of supertree construction to yield large complete phylogenetic estimates for a given clade allows biologists to potentially examine questions of greater scope and with more statistical power than would be possible with conventionally derived phylogenies.

The section begins with a large-scale supertree, that of 171 species of mammalian order Artiodactyla by Annette Mahon. Although the supertree is not complete — it is missing the Cetacea (now agreed widely to cluster within Artiodactyla) and several artiodactyl species — this was a result of conscious decisions on her part. This illustrates that supertrees need not be complete, but, like any phylogeny, can be tailor-made to suit a given objective and is also subject to the same issues of data quality and availability.

John Gittleman and colleagues then provide a general, wide-ranging review of how supertrees as large, complete phylogenetic estimates have been and perhaps could be used in biological research beyond their obvious utility for descriptive systematics. Although bigger might be better, they strike a note of caution as well, and point out limitations to supertrees that might occasionally make them inappropriate as a basis upon which to draw biological inferences.

The final two chapters in the volume use supertrees to examine macroevolutionary patterns in three diverse clades (grasses, angiosperms, and primates). All three analyses are possible because of the high level of taxonomic completeness that is potentially achievable through supertree construction. First, Nicolas Salamin and Jon Davies test numerous key-innovation hypotheses from the literature that attempt to relate species richness in grasses and angiosperms with certain morphological adaptations. Finally, Brian Moore and colleagues expand on their previous work with topology-based methods to investigate diversification rates (e.g., Chan and Moore, 2002), and develop a suite of shift statistics that can pinpoint where significant changes in rate have occurred on a supertree. They then apply their statistics to the dated primate supertree of Purvis (1995a), and compare their whole-tree macroevolutionary inferences to those of Purvis *et al.* (1995), which were derived using statistics that require divergence date estimates.

## 3. A last word

This volume hopefully reflects the current diversity of supertree research, with contributions from the different communities of mathematics, algorithmics, computer sciences, and biology. It is important to note that the writing conventions and even the “languages” of these communities differ greatly (particularly between the "extremes" of mathematics and biology). I have made a conscious decision to retain these differences, so as to maximize the impact and accuracy of each chapter for their respective community. As such, the appeal and accessibility of the different chapters will undoubtedly vary greatly depending on the background of the reader. However, in the end, I hope that the book serves both to provide an introduction to supertree construction and to highlight current research areas and issues.

## Acknowledgements

Financial support from the BMBF, through the "Bioinformatics for the Functional Analysis of Mammalian Genomes" (BFAM) project is gratefully acknowledged.

## References

AHO, A. V., SAGIV, Y., SZYMANSKI, T. G., AND ULLMAN, J. D. 1981. Inferring a tree from lowest common ancestors with an application to the optimization of relational expressions. *SIAM Journal on Computing* 10:405–421.

BARTHÉLEMY, J.-P., MCMORRIS, F. R., AND POWERS, R. C. 1995. Stability conditions for consensus functions defined on *n*-trees. *Mathematical Computer Modeling* 22:79–87.

BAUM, B. R. 1992. Combining trees as a way of combining data sets for phylogenetic inference, and the desirability of combining gene trees. *Taxon* 41:3–10.

BAUM, B. R. AND RAGAN, M. A. 1993. Reply to A.G. Rodrigo's "A comment on Baum's method for combining phylogenetic trees". *Taxon* 42:637–640.

BININDA-EMONDS, O. R. P., GITTLEMAN, J. L., AND PURVIS, A. 1999. Building large trees by combining phylogenetic information: a complete phylogeny of the extant Carnivora (Mammalia). *Biological Reviews* 74:143–175.

BININDA-EMONDS, O. R. P., GITTLEMAN, J. L., AND STEEL, M. A. 2002. The (super)tree of life: procedures, problems, and prospects. *Annual Review of Ecology and Systematics* 33:265–289.

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

CARDILLO, M. AND BROMHAM, L. 2001. Body size and risk of extinction in Australian mammals. *Conservation Biology* 15:1435–1440.

CHAN, K. M. A. AND MOORE, B. R. 2002. Whole-tree methods for detecting differential diversification rates. *Systematic Biology* 51:855–865.

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

FARRIS, J. S., KLUGE, A. G., AND ECKHARDT, M. J. 1970. A numerical approach to phylogenetic systematics. *Systematic Zoology* 19:172–191.

GARLAND, T., JR, DICKERMAN, A. W., JANIS, C. M., AND JONES, J. A. 1993. Phylogenetic analysis of covariance by computer simulation. *Systematic Biology* 42:265–292.

GATESY, J., MATTHEE, C., DESALLE, R., AND HAYASHI, C. 2002. Resolution of a supertree / supermatrix paradox. *Systematic Biology* 51:652–664.

GOLOBOFF, P. A. AND POL, D. 2002. Semi-strict supertrees. *Cladistics* 18:514–525.

GORDON, A. D. 1986. Consensus supertrees: the synthesis of rooted trees containing overlapping sets of labeled leaves. *Journal of Classification* 3:31–39.

HALL, J. P. W. AND HARVEY, D. J. 2002. Basal subtribes of the Nymphidiini (Lepidoptera: Riodinidae): phylogeny and myrmecophily. *Cladistics* 18:539–569.

HARVEY, P. H., LEIGH BROWN, A. J., MAYNARD SMITH, J., AND NEE, S. (eds) 1996. *New Uses for New Phylogenies*. Oxford University Press, Oxford.

HUELSENBECK, J. P., LARGET, B., AND SWOFFORD, D. 2000. A compound Poisson process for relaxing the molecular clock. *Genetics* 154:1879–1892.

JOHNSON, K. P. 2001. Taxon sampling and the phylogenetic position of Passeriformes: evidence from 916 avian cytochrome *b* sequences. *Systematic Biology* 50:128–136.

KÄLLERSJÖ, M., FARRIS, J. S., CHASE, M. W., BREMER, B., FAY, M. F., HUMPHRIES, C. J., PETERSEN, G., SEBERG, O., AND BREMER, K. 1998. Simultaneous parsimony jackknife analysis of 2538 *rbc*L DNA sequences reveals support for major clades of green plants, land plants, seed plants and flowering plants. *Plant Systematics and Evolution* 213:259–287.

KENNEDY, M., SPENCER, H. G., AND GRAY, R. D. 1996. Hop, step and gape: do the social displays of the Pelecaniformes reflect their phylogeny? *Animal Behaviour* 51:273–291.

KLUGE, A. G. 1989. A concern for evidence and a phylogenetic hypothesis of relationships among *Epicrates* (Boidae, Serpentes). *Systematic Zoology* 38:7–25.

LAPOINTE, F.-J. AND CUCUMEL, G. 1997. The average consensus procedure: combination of weighted trees containing identical or overlapping sets of taxa. *Systematic Biology* 46:306–312.

MCMORRIS, F. R. AND NEUMANN, D. 1983. Consensus functions defined on trees. *Mathematical Social Sciences* 4:131–136.

MOORE, B. R., CHAN, K. M. A., AND DONOGHUE, M. J. 2004. Detecting diversification rate variation in supertrees. In O. R. P. Bininda-Emonds (ed.), *Phylogenetic Supertrees: Combining Information to Reveal the Tree of Life*, pp. 487–533. Kluwer Academic, Dordrecht, the Netherlands.

NOVACEK, M. J. 2001. Mammalian phylogeny: genes and supertrees. *Current Biology* 11:R573-R575.

ORTOLANI, A. 1999. Spots, stripes, tail tips and dark eyes: predicting the function of carnivore colour patterns using the comparative method. *Biological Journal of the Linnean Society* 67:433–476.

PAGE, R. D. M. 2002. Modified mincut supertrees. In R. Guigó and D. Gusfield (eds), *Algorithms in Bioinformatics, Second International Workshop, WABI 2002, Rome, Italy, September 17–21, 2002, Proceedings*, pp. 537–552. Springer, Berlin.

PENNISI, E. 2003. Modernizing the Tree of Life. *Science* 300:1692–1697.

PISANI, D. 2002. *Comparing and Combining Data and Trees in Phylogenetic Analysis*. Ph.D. dissertation, Department of Earth Sciences, University of Bristol, United Kingdom.

PISANI, D. AND WILKINSON, M. 2002. Matrix representation with parsimony, taxonomic congruence, and total evidence. *Systematic Biology* 51:151–155.

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

PURVIS, A. 1995b. A modification to Baum and Ragan's method for combining phylogenetic trees. *Systematic Biology* 44:251–255.

PURVIS, A., NEE, S., AND HARVEY, P. H. 1995. Macroevolutionary inferences from primate phylogeny. *Proceedings of the Royal Society of London B* 260:329–333.

PURVIS, A. AND WEBSTER, A. J. 1999. Phylogenetically independent comparisons and primate phylogeny. In P. C. Lee (ed.), *Comparative Primate Socioecology*, pp. 44–70. Cambridge University Press, Cambridge.

RAGAN, M. A. 1992. Phylogenetic inference based on matrix representation of trees. *Molecular Phylogenetics and Evolution* 1:53–58.

RAMBAUT, A. AND BROMHAM, L. 1998. Estimating divergence dates from molecular sequences. *Molecular Biology and Evolution* 15:442–448.

RODRIGO, A. G. 1996. On combining cladograms. *Taxon* 45:267–274.

SALAMIN, N., HODKINSON, T. R., AND SAVOLAINEN, V. 2002. Building supertrees: an empirical assessment using the grass family (Poaceae). *Systematic Biology* 51:136–150.

SANDERSON, M. J. 2002. Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. *Molecular Biology and Evolution* 19:101–109.

SANDERSON, M. J., PURVIS, A., AND HENZE, C. 1998. Phylogenetic supertrees: assembling the trees of life. *Trends in Ecology and Evolution* 13:105–109.

SEMPLE, C. AND STEEL, M. 2000. A supertree method for rooted trees. *Discrete Applied Mathematics* 105:147–158.

SLOWINSKI, J. B. AND PAGE, R. D. M. 1999. How should species phylogenies be inferred from sequence data? *Systematic Biology* 48:814–825.

SOLTIS, P. S. AND SOLTIS, D. E. 2001. Molecular systematics: assembling and using the Tree of Life. *Taxon* 50:663–677.

SPRINGER, M. S. AND DE JONG, W. W. 2001. Phylogenetics. Which mammalian supertree to bark up? *Science* 291:1709–1711.

STEEL, M., DRESS, A. W. M., AND BÖCKER, S. 2000. Simple but fundamental limitations on supertree and consensus tree methods. *Systematic Biology* 49:363–368.

THORLEY, J. L. AND PAGE, R. D. 2000. RadCon: phylogenetic tree comparison and consensus. *Bioinformatics* 16:486-7.

THORNE, J. L., KISHINO, H, AND PAINTER, I. S. 1998. Estimating the rate of evolution of the rate of molecular evolution. *Molecular Biology and Evolution* 15:1647–1657.

THORNE, J. L. AND KISHINO, H. 2002. Divergence time and evolutionary rate estimation with multilocus data. *Systematic Biology* 51:689–702.

VOS, R. A. AND MOOERS, A. Ø. 2004. Reconstructing divergence times for supertrees. In O. R. P. Bininda-Emonds (ed.), *Phylogenetic Supertrees: Combining Information to Reveal the Tree of Life*, pp. 281–299. Kluwer Academic, Dordrecht, the Netherlands.

WEBB, C. O. 2000. Exploring the phylogenetic structure of ecological communities: an example for rain forest trees. *American Naturalist* 156:145–155.

WILLSON, S. J. 1999. Building phylogenetic trees from quartets by using local inconsistency measures. *Molecular Biology and Evolution* 16:685–693.

WILLSON, S. J. 2001. An error correcting map for quartets can improve the signals for phylogenetic trees. *Molecular Biology and Evolution* 18:344–351.

YODER, A. D. AND YANG, Z. H. 2000. Estimation of speciation dates using local molecular clocks. *Molecular Biology and Evolution* 17:1081–1090.