

Molecular and Morphological Supertrees for Eutherian (Placental) Mammals

Fu-Guo Robert Liu,¹ Michael M. Miyamoto,^{1*} Nicole P. Freire,¹
Phong Q. Ong,¹ Michele R. Tennant,² Timothy S. Young,¹
Kikumi F. Gugel¹

A large body of diverse comparative data now exists for a major phylogenetic synthesis of the higher-level relationships among eutherian (placental) mammals. We present such a phylogenetic synthesis using the composite trees or supertrees from the combined and separate analyses of their published molecular and morphological source phylogenies. Our combined and separate supertrees largely support the same suprafamilial taxa and orders, but different interordinal clades. These similarities and differences reinforce the continuing contributions of morphological studies, while highlighting the growing influence of molecular information on the field. As current summaries of past research, our supertrees emphasize opportunities for future work, while providing a step toward the eventual integration of the data and characters themselves.

Eutherian (placental) mammals are of great importance to humans as sources of food, materials, labor, companionship, and as test subjects in both basic and applied research (1). As a result, a large body of diverse comparative data currently exists for the phylogenetic synthesis of their higher level relationships (2–4). Ideally, such a phylogenetic synthesis will include the direct incorporation of all relevant information at the level of the individual datum and character (5). However, such integration remains presently unfeasible for the infraclass, because of several practical and analytical limitations, particularly with respect to its incomplete taxon sampling of characters (6).

In light of these problems, the next best strategy is to focus instead on the published (source) phylogenies of different studies, rather than on their separate data sets (6–8). This alternative strategy for phylogenetic synthesis is best embodied by the “matrix representation with parsimony” (MRP) approach. In this approach, the available source phylogenies are encoded as a series of binary characters that capture the branching patterns of the original trees. The final data matrix for the source phylogenies is then analyzed by the parsimony method to obtain a composite tree, or “supertree,” that summarizes the different conclusions within the field (9). Theoretical justification for this approach is founded on recent

evolutionary simulations that underscore its potential utility for phylogenetic synthesis (10).

In this study, 430 source phylogenies from 315 research articles were individually encoded by the MRP approach (11). The final data matrix of presence/absence characters for the 430 phylogenies consisted of 1965 binary features for the 90 recognized families and orders of the eutherian study group and one hypothetical outgroup (operational taxonomic units or OTUs) (12). The numbers of scored binary characters for the eutherian OTUs ranged from 342 for Solenodontidae of Insectivora to 1520 for Bovidae of Artiodactyla (Fig. 1). In addition to Bovidae, only nine other OTUs of the study group were scored for >900 (>45%) of the 1965 characters (Balaenopteridae and Delphinidae of Cetacea, Carnivora, Caviidae and Muridae of Rodentia, Equidae of Perissodactyla, Leporidae of Lagomorpha, Primates, and Suidae of Artiodactyla). Not surprisingly, this uneven coverage of the infraclass was heavily biased toward those eutherian families and orders with species of obvious economic, scientific, and aesthetic importance to humans (1).

Five most parsimonious (MP) solutions were recovered for the entire (combined) matrix using a 4:1-weighting scheme in favor of those source phylogenies that were derived from the direct phylogenetic study of multiple data sets and/or large numbers of taxa and characters (13). This weighting strategy was based on recent empirical and theoretical work in support of the greater power and accuracy of multiple data sets and increased character and taxon sampling in phylogenetic estimation (14, 15). The strict consensus tree for these five MP solutions [i.e., the supertree for the combined data matrix (9)] supported the monophyly of all nine eutherian orders with multiple represented families (Fig. 1) (16). Indeed, except for Artiodactyla and Insectivora, the monophyly of these orders was defined by ≥ 37 Bremer steps (17).

Except for Artiodactyla and Insectivora, interfamilial clusters within the other five orders with ≥ 3 families were also largely consistent with their traditional arrangements (Fig. 1). These intraordinal groupings included Mysticeti and Odontoceti of Cetacea, Microchiroptera (all families minus Pteropodidae) of Chiroptera, Ceratomorpha of Perissodactyla, Hystricognathi and Sciurognathi (excluding the poorly known Anomaluridae) of Rodentia, and Pilosa of Xenarthra (1, 18, 19). These polytypic suprafamilial taxa were corroborated by ≥ 21 Bremer steps, except for the two rodent suborders and Pilosa.

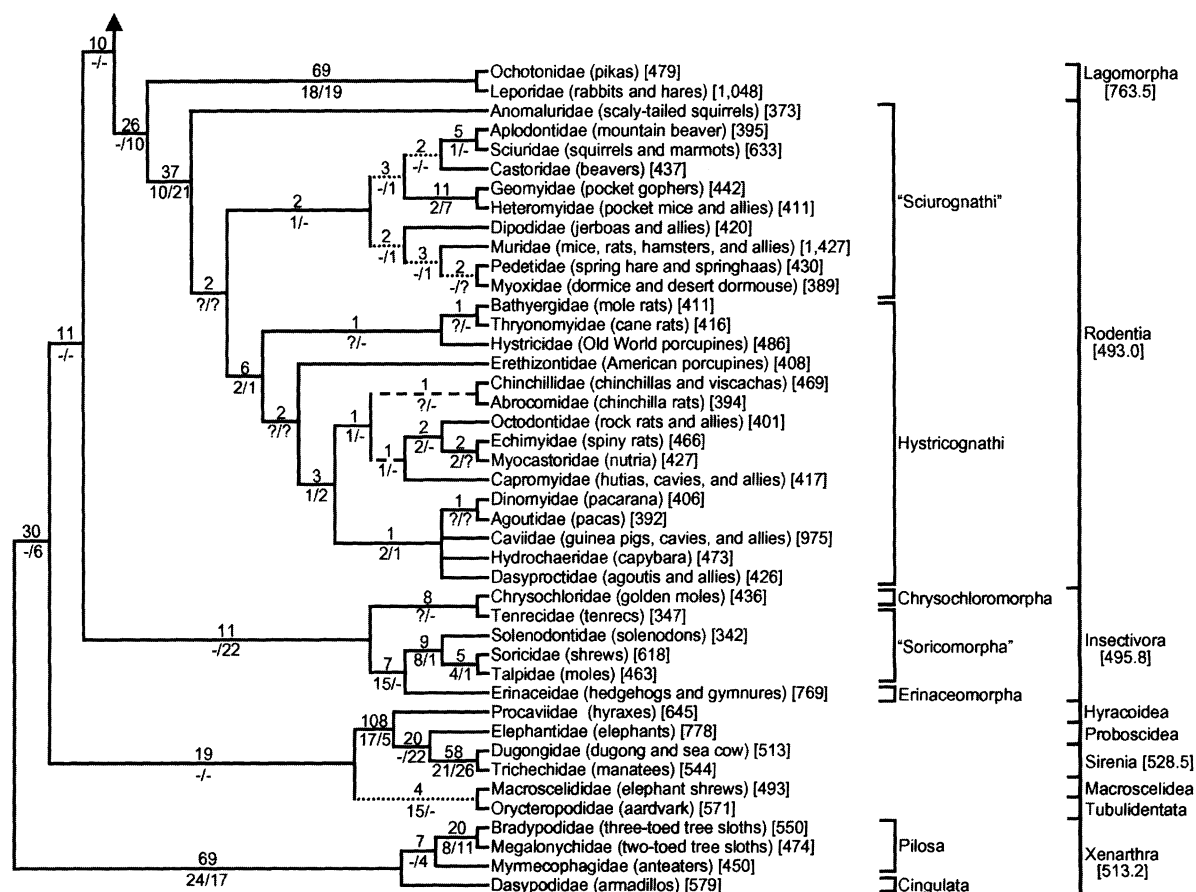
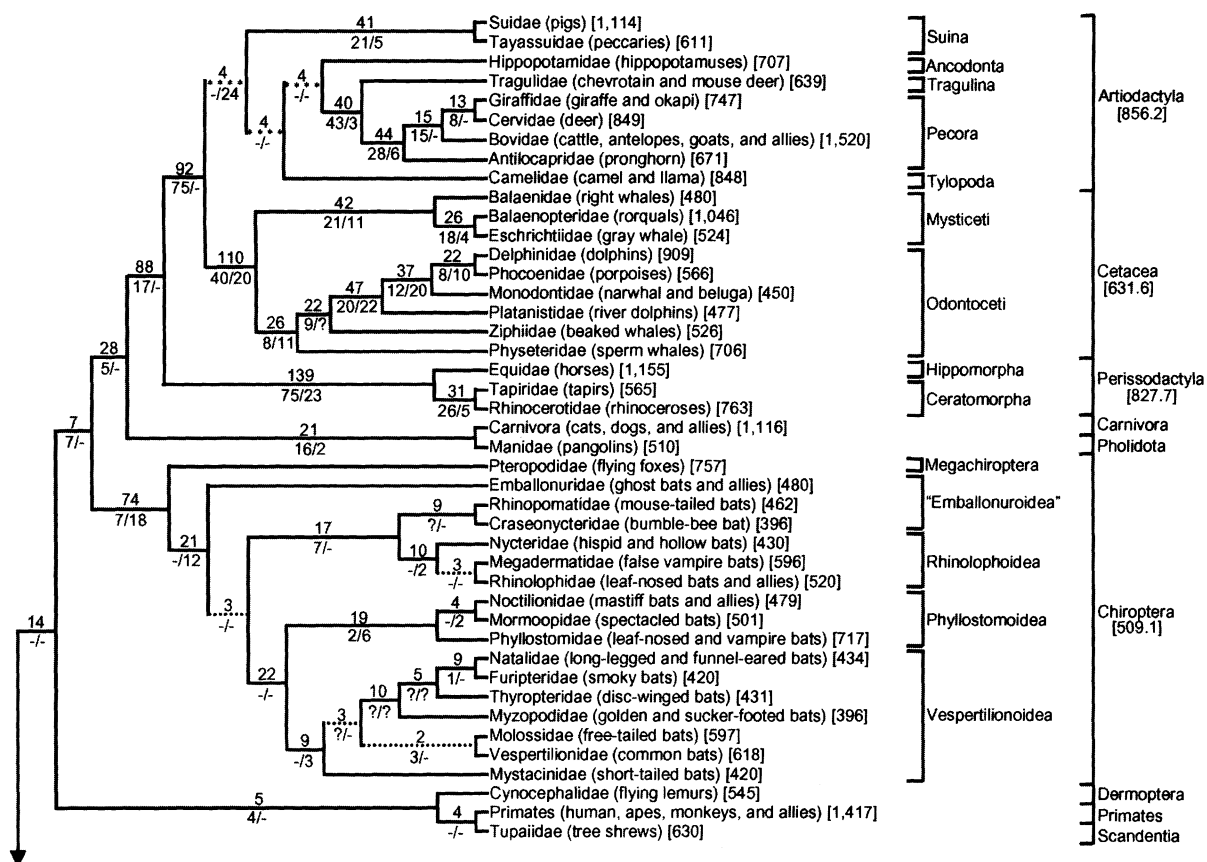
Above the ordinal level, the following nine interordinal clades were supported by ≥ 19 Bremer steps (Fig. 1): (i) Artiodactyla and Cetacea (i.e., the superorder Cetartiodactyla); (ii) Cetartiodactyla and Perissodactyla; (iii) Carnivora and Pholidota (Ferae); (iv) Cetartiodactyla, Ferae, and Perissodactyla (Fereuungulata); (v) Lagomorpha and Rodentia (Glires); (vi) Proboscidea and Sirenia (Tethytheria); (vii) Hyracoidea and Tethytheria (Paenungulata); (viii) Macroscelidea, Paenungulata, and Tubulidentata; and (ix) Placentalia minus Xenarthra (Epitheria) (3, 4, 20, 21). Six other interordinal groups were recovered by all four weighting schemes of the combined matrix, including Chiroptera and Fereuungulata (Scrotifera), Primates and Scandentia, and

Fig. 1. (Opposite) Strict consensus tree of the five MP solutions for the 4:1-weighted analysis of the combined data. This supertree is rooted against a hypothetical outgroup representing the noneutherian mammals, Marsupialia and Monotremata (not shown) (18, 19). Solid lines refer to clades that are stable across all four weighting schemes of this study. In contrast, dashed, dotted, and asterisked lines correspond to those that are not recovered in the 1:1-, 1:1- and 2:1-, or 8:1-weighting schemes, respectively. Quotes identify those suprafamilial taxa that are not monophyletic, whereas brackets give the total numbers of scored binary characters per OTU, along with the means for those orders with multiple represented families. Numbers above and below the internal branches correspond to Bremer scores for this combined analysis versus the separate molecular and morphological ones, respectively. Hyphens and question marks highlight those groups that are not recovered or that become “nonapplicable” by the exclusion of one or more families with extensive missing data in these separate analyses, respectively. Arrows connect the two parts of this supertree.

¹Department of Zoology, Box 118525, University of Florida, Gainesville, FL 32611–8525, USA. ²Health Science Center Libraries, Box 100206, University of Florida, Gainesville, FL 32610–0206, USA.

*To whom correspondence should be addressed. E-mail: miyamoto@zoo.ufl.edu

REPORTS



Dermoptera, Primates, and Scandentia (Euarchonta). However, none of these latter clades was defined by >14 Bremer steps.

At and below the ordinal level, the molecular and morphological supertrees both largely recovered the same interfamilial taxa, as emphasized above for the combined analysis (Figs. 1 and 2). Of five exceptions at these hierarchical levels, three coincided with the recent challenges from molecular sources for the nonmonophyly of Artiodactyla, Insectivora, and Microchiroptera. In contrast to the combined and morphological analyses, the molecular supertree favored a deep nesting of Cetacea within Artiodactyla and its close relationship to Ancodonta (22, 23). This alternative arrangement was also obtained by the combined data when weighted 8:1 in favor of phylogenetic syntheses that largely emphasized multiple genes. The molecular supertree also supported the removal of Chrysomorpha from Insectivora, its alternative placement with Macroscelidea, Paen-

ungulata, and Tubulidentata (Afrotheria), and its corresponding designation as a new order (Afrosoricida) (24–26). Finally, the failure of the molecular supertree to recover a monophyletic Microchiroptera was suggestive of the recent hypothesis for a nesting of Megachiroptera within this suborder (27). However, in contrast to Artiodactyla and Insectivora, this discrepancy was defined by one Bremer step, by some (but not all) weighting schemes for the molecular data, and by the unique grouping of Megachiroptera with Emballonuridae rather than with Rhinolophoidea. An analogous situation applied to the remaining two discrepancies for a nonmonophyletic Pilosa and Sciurognathi.

Above the ordinal level, only Ferae and Paenungulata from the combined analysis were corroborated by both the molecular and morphological supertrees (Figs. 1 and 2). Otherwise, the various interordinal clades of the combined analysis were recovered by one or neither of the separate

supertrees, but not both (3, 20, 21). The combined and molecular supertrees agreed on Cetartiodactyla, Cetartiodactyla/Perisodactyla, Fereuungulata, Scrotifera, Euarchonta, and Macroscelidea/Tubulidentata. Furthermore, both favored the concept of Afrotheria, albeit without Afrosoricida according to the combined data (24, 25). In contrast, only Epitheria, Glires, and Tethytheria were unique to both the combined and morphological analyses (20, 21).

The combined and separate supertrees testify to the expanding influence of molecular studies on the accepted higher level relationships among eutherian mammals (3, 4, 28). The current hypotheses of a nonmonophyletic Artiodactyla and Insectivora and of a monophyletic Afrotheria, Cetartiodactyla, Euarchonta, Fereuungulata, and Scrotifera are derived from molecular investigations (21–25, 29). Furthermore, the monophyly of Ferae and Paenungulata receives growing and convincing support, respectively, from molecular sources (2, 4,

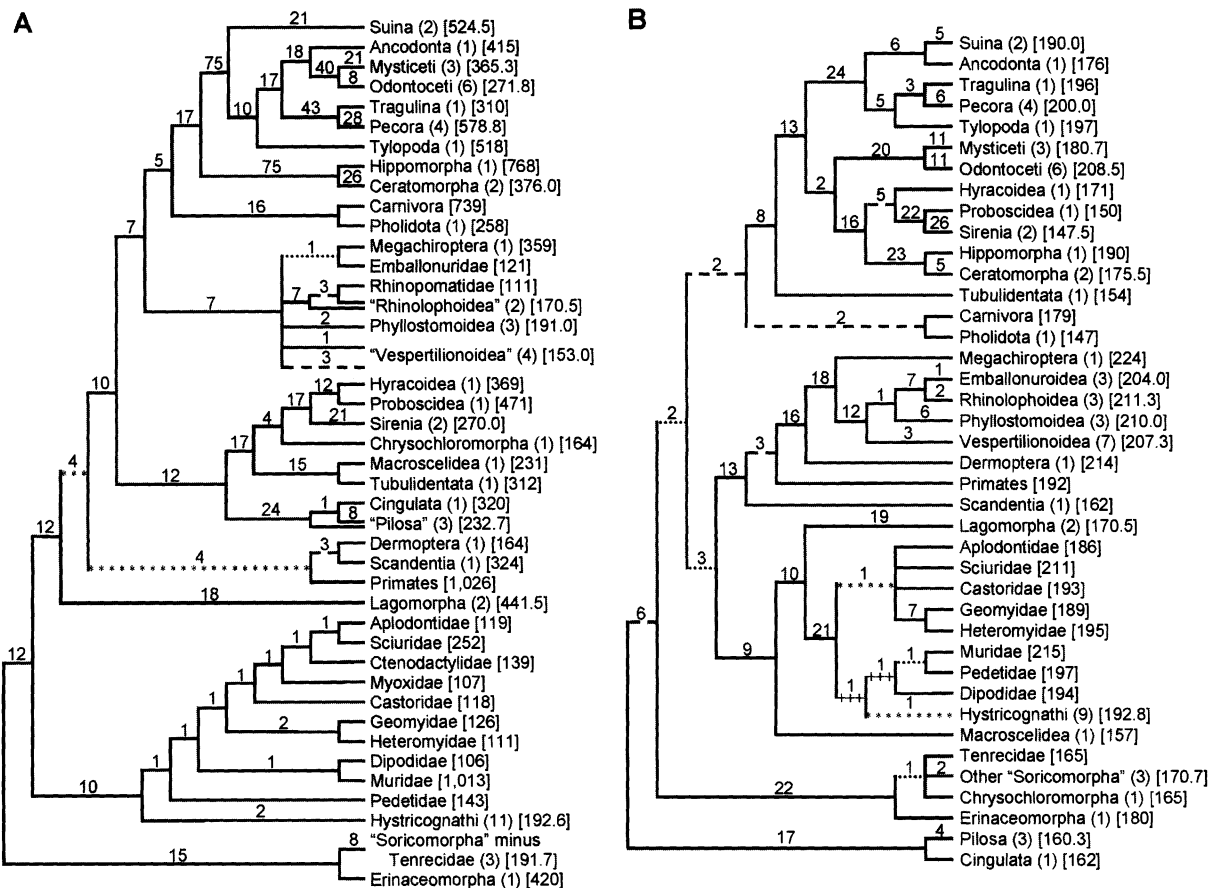


Fig. 2. Strict consensus trees of the 8 and 64 MP solutions from the separate 4:1-weighted analyses of the molecular (A) and morphological (B) source phylogenies. These molecular and morphological supertrees are based on 314 and 89 source phylogenies, respectively, and are summarized at the levels of eutherian families, suprafamilial taxa, and orders (Fig. 1). An additional 27 source phylogenies of the combined

matrix are not included in these separate analyses because they are based on both molecular and morphological data. Values in parentheses indicate the number of represented families in each interfamilial taxon, whereas "+" lines identify those groups that are not recovered in any of the alternative weighting schemes. All other numbering and conventions follow Fig. 1.

30). Conversely, this conclusion does not imply that morphological studies are becoming irrelevant for the resolution of the higher level relationships among eutherian mammals (20, 21). Rather, the combined and separate supertrees document that both lines of evidence agree on the majority of their traditional suprafamilial taxa and orders and that the recent molecular challenges for a nonmonophyletic Artiodactyla, Insectivora, and Microchiroptera are exceptions to this rule (Figs. 1 and 2). Above the ordinal level, the contributions of morphological studies remain most obvious in their corroboration of Ferae and Paenungulata and in their convincing support of Glires and Tethytheria, two superorders whose relationships are ambiguous according to molecular sources (20, 21, 24, 29). Both molecular and morphological studies remain critical for the continued resolution of higher level eutherian phylogeny.

Our complete list of 315 research articles, final data matrix and associated summary for 430 source phylogenies, and combined and separate supertrees (Figs. 1 and 2) are available as supplementary Web information (31). Our final data matrix and associated summary provide the opportunity to evaluate weighting strategies different from the one emphasized here (13). For example, an obvious alternative to assess is that of the carnivore and primate supertree studies with its greater emphasis on different methods of phylogenetic analysis (7, 8). Furthermore, these supplementary materials offer the chance to review the historical impact of systematics on the recognition of particular eutherian taxa (32). In turn, our combined and separate supertrees can serve as working hypotheses to study the origins and significance of biological diversity within the infraclass and to highlight those questions that are important for the greater understanding of eutherian phylogeny, character homology, and organismal evolution (8). For example, resolution of the molecular versus morphological conflict for Insectivora (minus Afrosoricida) versus Xenarthra at the base of the eutherian tree is necessary for the identification of a more closely related outgroup than Marsupialia and Monotremata for phylogenetic and comparative studies within the infraclass (3, 20). Our supertrees advance the field one step closer toward its ultimate goal of a phylogenetic synthesis for eutherian mammals at the level of their individual datum and character (5, 21).

References and Notes

1. R. M. Nowak, *Walker's Mammals of the World* (Johns Hopkins Univ. Press, Baltimore, MD, ed. 6, 1999).
2. F. S. Szalay, M. J. Novacek, M. C. McKenna, Eds., *Mammal Phylogeny: Placentals* (Springer-Verlag, New York, 1993).
3. W. W. de Jong, *Trends Ecol. Evol.* **13**, 270 (1998).
4. P. J. Waddell, N. Okada, M. Hasegawa, *Syst. Biol.* **48**, 1 (1999).
5. A. G. Kluge, *Syst. Zool.* **38**, 7 (1989).
6. M. J. Sanderson, A. Purvis, C. Henze, *Trends Ecol. Evol.* **13**, 105 (1998).
7. A. Purvis, *Philos. Trans. R. Soc. London Ser. B* **348**, 405 (1995).
8. O. R. P. Bininda-Emonds, J. L. Gittleman, A. Purvis, *Biol. Rev.* **74**, 143 (1999).
9. In cases of more than one MP solution, a strict consensus tree can serve to summarize the agreement among these alternatives (8, 32).
10. O. R. P. Bininda-Emonds, M. J. Sanderson, *Syst. Biol.*, in press.
11. The following electronic literature databases were searched over the time period from the year in parentheses to March 1998 for publications with eutherian interfamilial phylogenies: Biological Abstracts (Silverplatter) (1980), Biological Abstracts: Reports, Records, and Monographs (Silverplatter) (1989), Current Contents Connect (1998), MEDLINE (Silverplatter) (1966), and Web of Science (1992). These electronic surveys were supplemented until March 1999 with more recent articles from the primary journals in mammalogy, systematics, and evolutionary biology (e.g., *Systematic Biology*). These searches focused on extant eutherian families and orders, given our interests in both their molecular and morphological data and their higher level relationships. They resulted in a starting list of 16,102 titles and abstracts, from which 1477 were kept on the basis of their summaries for further review. These retained articles were screened for source phylogenies that were based on different character sets and/or were unique by >30% of their OTUs. Only the most recent and complete source phylogenies were saved from those articles with more broadly overlapping OTUs for the same character matrix. Source phylogenies that were derived from combined information were selected in the same way for the different published combinations of their data partitions. In contrast, in the absence of any justified preference, the alternative solutions for the same data set were reduced to a single source phylogeny by Adams consensus (33). These reductions were limited to 26 source phylogenies, and no OTU was repeatedly affected by them. In these ways, redundancy and nonindependence were lessened as the 430 source phylogenies were compiled for supertree analysis. However, because some nonindependence remained, bootstrapping and homogeneity partition tests were not conducted to assess group support and molecular-versus-morphological conflict, respectively (7, 8, 17).
12. The 430 source phylogenies were encoded as a series of presence/absence characters by the MRP methods both of Baum (34) and Ragan (35) and of Purvis (36). However, only the results using the Baum and Ragan method were presented, because the two coding schemes supported similar conclusions. The 90 eutherian OTUs included all recognized extant families of the infraclass, except for those of Carnivora and Primates (37). These two were represented in the data matrix as separate orders, rather than by their extant families, because of the availability of their published intraordinal supertrees (7, 8).
13. Supertrees were constructed from the combined and separate molecular and morphological data by parsimony analyses with PAUP* 4.0 (33). These parsimony analyses relied on heuristic searches with 1000 random additions and TBR (tree bisection-reconnection) branch swapping. Four, 13, and 10 eutherian families were excluded from these combined and separate molecular and morphological analyses, respectively, because of their extensive missing data that destabilized the heuristic searches. In the 4:1-weighting scheme, "large" was defined for the molecular and morphological source phylogenies as >1500 characters and 20 terminal taxa, and >40 features from different anatomical systems (e.g., hard versus soft anatomy), respectively. These analyses were then repeated with other unequal (2:1 and 8:1), as well as equal (1:1), weighting schemes to check the robustness of the initial 4:1 results. This series of weights followed that of the carnivore and primate supertree studies (7, 8).
14. A. Graybeal, *Syst. Biol.* **47**, 9 (1998).
15. A. Mitchell, C. Mitter, J. C. Regier, *Syst. Biol.* **49**, 202 (2000).
16. For 21 of the 89 morphological source phylogenies, the original OTUs consisted of suprafamilial taxa or orders, rather than individual species, genera, or families. This use of interfamilial OTUs necessitated the adoption and subsequent scoring of these taxa as monophyletic in our final data matrix, as tacitly accepted in their original studies (8). Thus, our morphological support for certain suprafamilial taxa and orders was based to some degree on a priori assumptions, rather than on new character analyses. This concern was much less of an issue for the molecular support, because only 6 of its 314 source phylogenies were based on interfamilial OTUs.
17. The strength of grouping for an individual clade of the combined and separate supertrees was measured by its Bremer support (38). This coefficient summarized the number of extra steps necessary for the removal of a group from its MP solution. A Bremer score of only one or two steps was indicative of a group with minimum stability, because of conflicting evidence and/or missing data (8). In contrast, a count of, for example, 19 or 20 steps was representative of a clade that was relatively well supported. In combination with the alternative weighting schemes, these Bremer scores provided important checks on the robustness of our 4:1-weighted supertrees.
18. G. G. Simpson, *Bull. Am. Mus. Nat. Hist.* **85**, 1 (1945).
19. M. C. McKenna, S. K. Bell, *Classification of Mammals Above the Species Level* (Columbia Univ. Press, New York, 1997).
20. J. Shoshani, M. C. McKenna, *Mol. Phylogenet. Evol.* **9**, 572 (1998).
21. F.-G. R. Liu, M. M. Miyamoto, *Syst. Biol.* **48**, 54 (1999).
22. J. Gatesy, M. Milinkovitch, V. Waddell, M. Stanhope, *Syst. Biol.* **48**, 6 (1999).
23. M. Nikaido, A. P. Rooney, N. Okada, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 10261 (1999).
24. M. J. Stanhope et al., *Mol. Phylogenet. Evol.* **9**, 501 (1998).
25. M. J. Stanhope et al., *Proc. Natl. Acad. Sci. U.S.A.* **95**, 9967 (1998).
26. The other representative of Afrosoricida (Tenrecidae) was excluded from the molecular analysis because of its extensive missing data (only 100 scored binary characters for this partition).
27. E. C. Teeling et al., *Nature* **403**, 188 (2000).
28. S. J. O'Brien et al., *Science* **286**, 463 (1999).
29. P. J. Waddell, Y. Cao, J. Hauf, M. Hasegawa, *Syst. Biol.* **48**, 31 (1999).
30. M. S. Springer, H. M. Amrine, A. Burk, M. J. Stanhope, *Syst. Biol.* **48**, 65 (1999).
31. Supplementary material is available at www.sciencemag.org/cgi/content/full/291/5509/1786/DC1
32. O. R. P. Bininda-Emonds, *Mol. Phylogenet. Evol.* **16**, 113 (2000).
33. D. L. Swofford, *PAUP*: Phylogenetic Analysis Using Parsimony (* and Other Methods)*, version 4 (Sinauer, Sunderland, MA, 1998).
34. B. R. Baum, *Taxon* **41**, 3 (1992).
35. M. A. Ragan, *Mol. Phylogenet. Evol.* **1**, 53 (1992).
36. A. Purvis, *Syst. Biol.* **44**, 251 (1995).
37. D. E. Wilson, D. M. Reeder, Eds., *Mammal Species of the World: A Taxonomic and Geographic Reference* (Smithsonian Institution Press, Washington, DC, 1993).
38. K. Bremer, *Evolution* **42**, 795 (1988).
39. We thank M. W. Allard, O. R. P. Bininda-Emonds, J. F. Eisenberg, A. Purvis, and M. J. Sanderson for comments about our research, K. Park, J. M. Piascik, and R. Tseng for help with manuscript preparation, and the Department of Zoology, University of Florida, for financial assistance.

4 October 2000; accepted 9 January 2001