PERSPECTIVES: PHYLOGENETICS

Which Mammalian Supertree to Bark Up?

Mark S. Springer and Wilfried W. de Jong

One of the most enduring and contentious disputes in phylogenetics concerns the relationships among the 18 existing orders of placental mammals. Morphological data (from fossil and anatomical studies) and molecular data (from analyses of nuclear and mitochondrial DNA) yield very different mammalian phylogenetic trees. These differences have practical consequences when it comes to comparative and functional genomics. With sequencing of the human genome now complete and the mouse genome sequence on the horizon, we inch closer to understanding the organization and structure of the ancestral mammalian genome as well as the differences that make us unique. A well-resolved phylogeny for the mammalian orders will be essential if appropriate taxa are to be selected for future genome sequencing efforts.

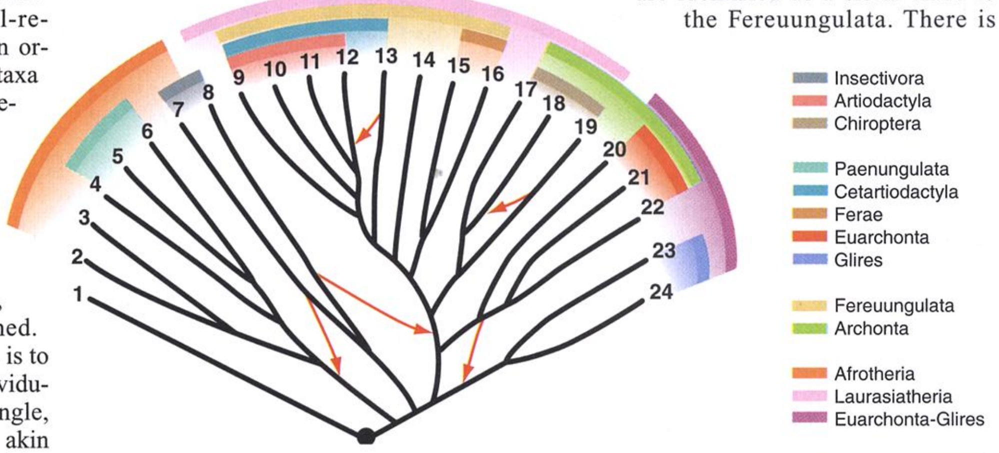
Enhanced online at www.sciencemag.org/cgi/content/full/291/5509/1709

Controversies still rage about the type of data that is most reliable, the method of phylogeny reconstruction that is most appropriate, whether or not to combine morphological and molecular data sets, and how this should be accomplished. One approach for combining data sets is to construct a supermatrix in which individual data matrices are joined into a single, larger matrix. An alternative strategy akin to meta-analysis (*1*), and one adopted by Liu *et al.* (*2*) on page 1786 of this issue, is to construct a supertree from source trees derived from separate morphological and molecular data sets. Weighing in on the classic "molecules versus morphology" debate, the Liu *et al.* work questions recent mammalian classifications based on molecular systematics and instead agrees with a traditional morphology-based classification, at least for some clades of mammals (*3*) (see the figure, this page).

These investigators retrieved and integrated 430 phylogenetic trees from 315 papers published between 1969 and 1999. Of these trees, 314 were derived from molecular data, 89 from morphological data, and 27 from mixed data sets. Adopting Baum/Ragan matrix representation with parsimony (MRP) (*1*), Liu *et al.* constructed morphological, molecular, and combined supertrees, recoding each source tree as a binary matrix and then combining these matrices for parsimony analysis (see the figure, next page).

Combining morphological and molecular source phylogenies, Liu *et al.* challenge the compelling molecular evidence that whales are the sister group of hippopotami, and thus nested within the order Artiodactyla (even-toed ungulates) (*4*); that golden moles and tenrecs cluster with elephant shrews, aardvarks, and paenungulates (elephants, sea cows, and hyraxes) in a clade called Afrotheria, rather than with eulipotyphlan insectivores such as hedgehogs, shrews, and moles (*5*); and that megabats are nested within the microbats, rendering the microbats paraphyletic (that is, including the common ancestor but only a fraction of the descendants) (*6*).

At the interordinal level, the Liu *et al.* morphological and molecular supertrees are remarkably incongruent, only agreeing on the members of the Paenungulata and Ferae (carnivores plus pangolins). When the authors combined their morphological and molecular supertrees, the resulting supertree more closely resembled the supertree derived from molecular data alone. In their combined supertree, cetaceans (whales and dolphins) are grouped together with artiodactyls in the taxon Cetartiodactyla, consistent with molecular systematics. Taxon Fereuungulata encompasses carnivores, pangolins, perissodactyls (horses, rhinos, and tapirs) and cetartiodactyls. Bats (Chiroptera) are identified as a sister clade to the Fereuungulata. There is



What's in a name? A combined morphological and molecular supertree for mammals. To reconcile the combined supertree of Liu *et al.* (*2*) with recent molecular analyses (*4–9*), only a few additional rearrangements (indicated by arrows) are required. Notably, Laurasiatheria and Euarchonta-Glires are each recovered with a single rearrangement. The earliest fossils for both Laurasiatheria and Euarchonta-Glires have been found in the Northern Hemisphere. The name Boreoeutheria has been suggested for the monophyletic clade that includes these taxa. The earliest fossils for the remaining placental mammals, Afrotheria and Xenarthra, are Gondwanan, having been found in Africa and South America, respectively (*8*). The possible apposition of Boreoeutheria with the Southern Hemisphere's Afrotheria and Xenarthra is in keeping with other studies (*12*) that necessitate the reevaluation of the early evolutionary history of placental mammals in the context of plate tectonics. 1. Xenarthra (sloths, armadillos, anteaters), 2. Macroscelidea (elephant shrews), 3. Tubulidentata (aardvarks), 4. Hyracoidea (hyraxes), 5. Sirenia (manatees, dugongs), 6. Proboscidea (elephants), 7. Afrosoricida (tenrecs, golden moles), 8. Eulipotyphla (moles, hedgehogs, shrews), 9. Suidae (pigs), 10. Camelidae (camels), 11. Ruminantia (deer, cows, giraffes), 12. Hippopotamidae (hippos), 13. Cetacea (dolphins, whales), 14. Perissodactyla (horses, rhinos, tapirs), 15. Carnivora (bears, cats, dogs), 16. Pholidota (pangolins), 17. Non-rhinolophoid microbats (New World leaf-nosed bats, vampire bats, vespertilionoid bats), 18. Rhinolophoid microbats (false vampire bats, horseshoe bats, Old World leaf-nosed bats), 19. Megachiroptera (megabats, including flying foxes), 20. Dermoptera (flying lemurs), 21. Scandentia (tree shrews), 22. Primates (humans, monkeys, lemurs), 23. Rodentia (rats, mice, guinea pigs), 24. Lagomorpha (pikas, rabbits). The 18 traditional placental orders are numbers 1–6, 13–16, 20–24, Insectivora (7–8), Artiodactyla (9–12), and Chiroptera (17–19).

M. S. Springer is in the Department of Biology, University of California, Riverside, CA 92521, USA. W. W. de Jong is in the Department of Biochemistry, University of Nijmegen, Nijmegen, Netherlands, and the Institute for Biodiversity and Ecosystems Dynamics, University of Amsterdam, Netherlands.

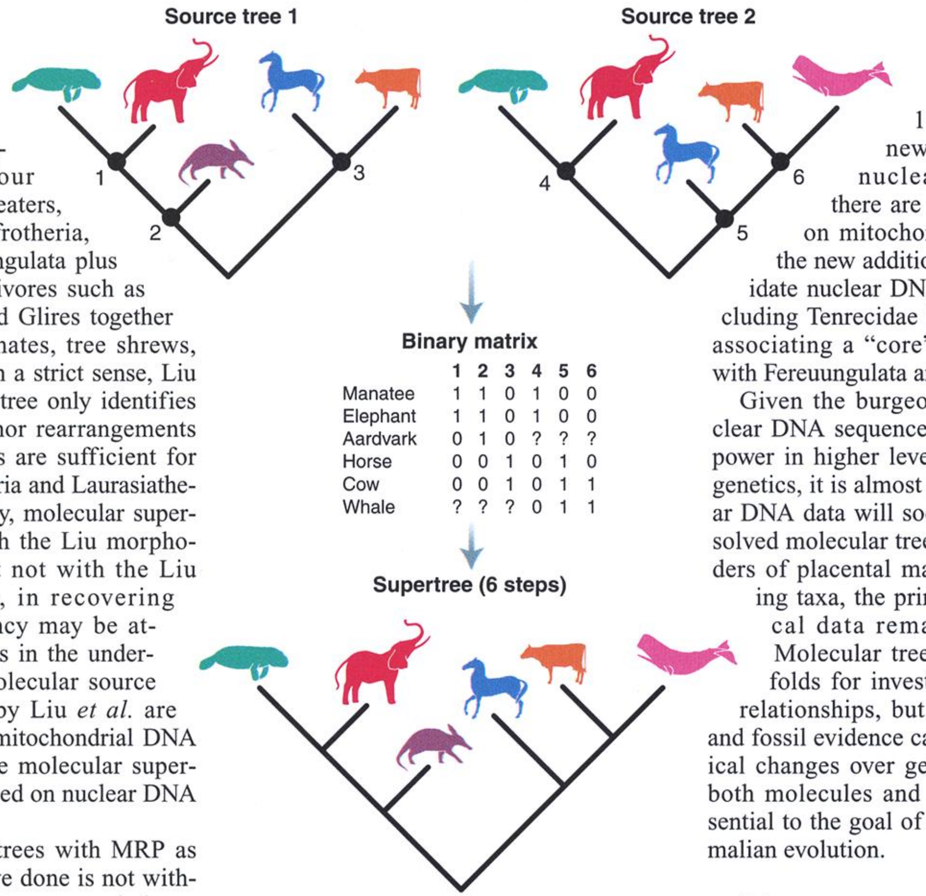
partial recognition of taxon Afrotheria, although the golden moles and tenrecs are left out. Other features of the Liu *et al.* combined supertree reflect input from morphological data—notably, artiodactyls, insectivores, and microbats are maintained in monophyletic groups (which include the common ancestor and all of the descendants); rodents and lagomorphs are clustered in the Glires clade, and elephants and sea cows in the Tethytheria clade.

How do the separate and combined Liu *et al.* supertrees compare with recent supermatrix molecular trees derived by our groups and others from analyses of long concatenated DNA sequences that index multiple genetic loci? As anticipated by Waddell *et al.* (*7*), the molecular supertrees divide placental mammals into four clades: Xenarthra (anteaters, sloths, armadillos), Afrotheria, Laurasiatheria (Fereuungulata plus bats and "core" insectivores such as moles and shrews), and Glires together with Euarchonta (primates, tree shrews, flying lemurs) (*8, 9*). In a strict sense, Liu *et al.*'s combined supertree only identifies the Xenarthra, but minor rearrangements among the insectivores are sufficient for recovery of the Afrotheria and Laurasiatheria as well. Interestingly, molecular supermatrix trees agree with the Liu morphological supertree, but not with the Liu molecular supertree, in recovering Glires. This discrepancy may be attributable to differences in the underlying data sets; the molecular source phylogenies indexed by Liu *et al.* are heavily influenced by mitochondrial DNA sequences, whereas the molecular supermatrices are largely based on nuclear DNA sequences.

Constructing supertrees with MRP as Liu and colleagues have done is not without its problems. First, source phylogenies need to be independent to avoid disproportional representation of certain data. Although Liu and co-workers attempted to minimize nonindependence, inspection of their source phylogenies reveals that nonindependence is pervasive (for example, a single transferrin immunology data set for bats was incorporated into five different source trees). Comparisons by Liu and colleagues of decay indices in a supertree are undermined by this nonindependence. A second problem concerns the recognition and investigation of incompatible source phylogenies. Liu *et al.* did not perform incompatibility tests because of the nonindependence problem, but this only compounds one problem with another. If source phylogenies are incongruent, it may be possible to elucidate the underlying cause and even to rectify the incongruence. Third, Liu *et al.* incorporated source phylogenies that were published when modern methods for phylogeny reconstruction were still in their infancy; included are hand-drawn cladograms that cannot be reconstructed from primary data matrices. A fourth issue concerns weighting of weakly versus strongly supported nodes on individual source phylogenies (*1*), which Liu *et al.* did not attempt. Supertrees are certainly of value for summarizing and combining source phylogenies, but the accuracy of phylogeny reconstruction ultimately depends on primary data, the analyses of these data, and corroboration by independent data sets.

What should we anticipate in future mammalian supertrees? At the interordinal level, Liu *et al.*'s morphological tree is largely congruent with other morphological trees published during the last 20 years. However, changes should be expected; in a recent morphological study by Asher (*10*), some of the analyses failed to recover insectivores as a monophyletic group. The molecular supertree, in turn, is already out of date as a result of the plethora of new molecular trees published after March 1999. In addition to new trees derived from nuclear DNA analyses, there are also new trees based on mitochondrial DNA. Among the new additions are some that validate nuclear DNA phylogenies by including Tenrecidae in Afrotheria and by associating a "core" insectivore (mole) with Fereuungulata and bats (*11*).

Given the burgeoning number of nuclear DNA sequences and their resolving power in higher level mammalian phylogenetics, it is almost inevitable that nuclear DNA data will soon achieve a well-resolved molecular tree for the 18 living orders of placental mammals. Beyond living taxa, the primacy of morphological data remains unchallenged. Molecular trees can serve as scaffolds for investigating evolutionary relationships, but only morphological and fossil evidence can document anatomical changes over geologic time. Clearly both molecules and morphology are essential to the goal of reconstructing mammalian evolution.



	1	2	3	4	5	6
Manatee	1	1	0	1	0	0
Elephant	1	1	0	1	0	0
Aardvark	0	1	0	?	?	?
Horse	0	0	1	0	1	0
Cow	0	0	1	0	1	1
Whale	?	?	?	0	1	1

Constructing supertrees. The Baum/Ragan method for constructing strict supertrees can be used whether the source trees are compatible or incompatible. The minimum requirement for including a source tree in supertree construction is that it shares two or more taxa with at least one other source tree (*1*). Each source tree is recoded as a binary matrix. Binary matrices are then combined and analyzed with parsimony to find the most parsimonious supertree. The combined binary matrix shown above indexes two different source trees. Binary characters 1–3 and 4–6 are from source trees 1 and 2, respectively. Question marks denote taxa that are missing from a source tree.

References

1. M. J. Sanderson, A. Purvis, C. Henze, *Trends Ecol. Evol.* **13**, 105 (1998).
2. F.-G. R. Liu *et al.*, *Science* **291**, 1786 (2001).
3. J. Shoshani, M. C. McKenna, *Mol. Phylogenet. Evol.* **9**, 572 (1998).
4. J. Gatesy, M. Milinkovitch, V. Waddell, M. Stanhope, *Syst. Biol.* **48**, 16 (1999).
5. M. J. Stanhope *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 9967 (1998).
6. J. M. Hutcheon, J. A. W. Kirsch, J. D. Pettigrew, *Philos. Trans. R. Soc. London Ser. B* **353**, 607 (1998).
7. P. J. Waddell, N. Okada, M. Hasegawa, *Syst. Biol.* **48**, 1 (1999).
8. O. Madsen *et al.*, *Nature* **409**, 610 (2001).
9. W. J. Murphy *et al.*, *Nature* **409**, 614 (2001).
10. R. J. Asher, *Cladistics* **15**, 231 (1999).
11. S. K. Mouchaty, A. Gullberg, A. Janke, U. Arnason, *Zool. Scr.* **29**, 307 (2000).
12. S. B. Hedges, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 1 (2001).