

Mammalian phylogeny: Genes and supertrees

Michael J. Novacek

A massive effort to sample mammals for genes has yielded new proposals for the branching architecture of the great radiation of placental mammals. Some of these are notably discrepant with morphologically based analyses, but they suggest new research that should address several major outstanding issues.

Address: The Division of Paleontology, American Museum of Natural History, New York, New York 10024, USA.

Current Biology 2001, 11:R573–R575

0960-9822/01/\$ – see front matter

© 2001 Elsevier Science Ltd. All rights reserved.

Despite a general recognition of certain key branches of the tree of life — the schema of relationships among taxa — many sectors of this tree remain either unresolved or the focus of much debate. Notable among these is the diverse group the placental mammals that includes humans and other primates, bats, cats, rats and sundry taxa. Besides representing the community of organisms within which are our own evolutionary roots, placental mammals offer a distinctly enriched fossil record [1], and are an excellent subject for studies of evolutionary radiations. They are also highly advantageous as a target for many studies that have applications in medicine, pharmaceuticals and food production. Accordingly, placental mammals have attracted a flourish of comparative molecular work catalyzed by major advances in DNA sequencing technology, data retrieval and computation.

The capacity of molecular studies to capture a much greater level of character and taxon information than has been obtained in earlier analyses is illustrated in a recent study of four nuclear genes and mitochondrial RNA genes in 26 taxa of placentals conducted by Madsen *et al.* [2]. Trees constructed from these data using various techniques, such as maximum likelihood and parsimony based algorithms, reveal a geometry both consistent and markedly discrepant with those generated from morphological studies. For example, a grouping of elephants, hyraxes and sirenians (sea cows and manatees) faithfully mirrors the morphologically based Paenungulata. This grouping, however, is nested within a clade labeled Afrotheria which is not supported by morphological results. In addition to paenungulates, Afrotheria also comprises the Madagascar tenrecs, as well as golden moles, elephant shrews and armadillos (Figure 1). As their name implies, afrotherians putatively represent a radiation of placentals rooted in the African continent and nearby island of Madagascar. From a morphological standpoint, this aggregation is extremely provocative. Tenrecs and golden moles are generally referred to

the insectivoran groups, along with shrews moles and hedgehogs. Elephant shrews and armadillos have had varying assignments, but rarely ones that emphatically associate them with other afrotherians [3].

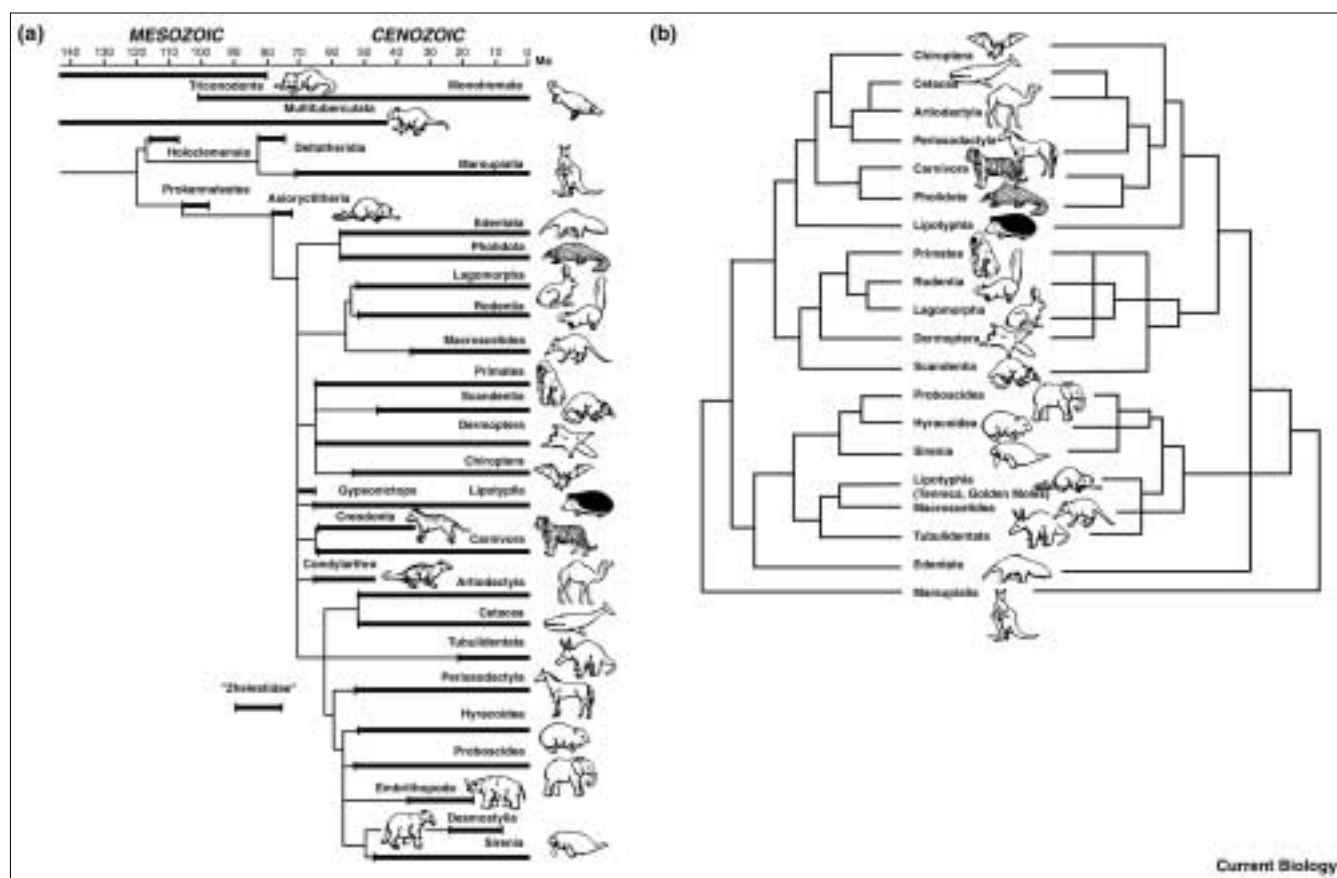
Another aspect of this molecular-based tree [2] that contradicts morphological results is the insertion of the Cetacea (whales and dolphins) within the artiodactyls — hippos, pigs, cows, antelope, camel, deer, sheep and their relatives. This association is consistent with results based on intensive sampling of genes and taxa in artiodactyls and whales [4]. It is noteworthy, however, that this group collapses with a combined analysis of genes and morphological data in both living and fossil taxa [5].

In other respects, the trees produced in the sextet gene analysis of Madsen *et al.* [2] both agree and disagree with morphological results (Figure 1). In addition to Afrotheria, the authors also recognize another superclade, the Laurasiatheria, not supported by morphological data. The Laurasiatheria comprise the hedgehogs, moles and shrews (linked together as a core insectivoran group that excludes tenrecs and golden moles), bats, pangolins, artiodactyls (including whales), perissodactyls and carnivorans. The name Laurasiatheria derives from the former supercontinent Laurasia, the mega-landmass that included North America, Europe and much of Asia, and the alleged place of origin for many of the clades assigned to this group.

A second, even more comprehensive comparison of 15 nuclear genes and three mitochondrial genes for 64 placentals and two marsupials by Murphy *et al.* [6] also produces some results that identify the clade Afrotheria. This study recognizes three other major clades: Xenarthra; a clade consisting of the archontan taxa primates, tree shrews and flying lemurs (but excluding bats) plus Glires (rodents and lagomorphs); and a clade reminiscent of Laurasiatheria that includes bats, artiodactyls (including whales), perissodactyls, carnivorans plus pholidotans, and insectivorans [6]. These results are regarded by the authors as independent corroboration of conclusions drawn by Madsen *et al.* [2].

Given the massive infusion of new DNA data represented by these two studies [2,6], where do we stand in our quest to resolve the major branches of the placental mammal tree? Acceptance of such groups as Afrotheria or Laurasiatheria would suggest a radical shakeout of the placental tree, in ways not anticipated by morphological work. But do we have reason to think the matter is now finally resolved? The question is addressed in a recent

Figure 1



(a) Branching sequence of higher placental mammal clades supported by morphological data [3], showing the geologic ranges of clades (thick horizontal bars). Ages of divergence (thin lines) between any two or more clades are based on the oldest occurrence of one of the sister lineages. Internodal distances are slightly exaggerated to show topology. (b) The branching sequence of higher placental clades reported by Madsen *et al.*

[2] is shown on the left. This solution omits data from one of the five nuclear genes in order to depict relationships of clades based on an expanded set of 52 taxa. The branching sequence of higher placental clades represented by 18 genes and 66 taxa reported by the Murphy *et al.* study [6] is shown on the right. Note discrepancies as well as agreement in branching structure between the two results.

study of a large number of molecular and morphological results by Liu *et al.* [7]. These authors used published trees themselves as data to form a binary matrix from which a resultant 'supertree' is identified. The supertree then is a composite of the agreements and discrepancies between the various contributing trees. The molecular supertree based on 314 contributing molecular trees was found to be largely consistent with such groupings as Afrotheria and a clade that includes whales and other artiodactyls. Nonetheless, the combined morphological and molecular supertree, based on 430 source phylogenies and 315 papers, fails to support these groups and instead is more consistent with a variety of morphological data.

Taken at face value, the new supertree [7] indicates that there is a problem with the major clades, such as Afrotheria and Laurasiatheria, identified in the recent gene studies

[2,6]. It is noteworthy, however, that the supertree analysis [7] did not incorporate these new and more comprehensive gene results. More important is a concern about the methodological justification for supertrees in the first place. The supertree method uses taxonomic congruence based on various published trees — whatever their character range and methodological underpinnings — rather than the congruence among characters in a single character–taxon data matrix. This approach cannot substitute for a much needed combined analysis of the relevant character and taxon evidence [8]. Supertrees, or the less formally constructed summary trees that anecdotally 'paste together' different published results [9], defy any logical basis for a test of the structure of the tree that depends on new, comprehensive data sets. Ironically, Liu *et al.* [7] themselves recognize a better approach would be to combine character data with the taxa under study in a single supermatrix.

Do then the problems inherent in constructing supertrees leave molecular-based theories, such as the existence of the Afrotheria or Laurasiatheria superclade, untainted? Here one should note some limitations of even these more ambitious molecular studies. The taxon and character sampling required to test definitively the monophyly of either group has not been accomplished. Missing, for example, are data for key taxa — such as more generalized relatives of hedgehogs, the echinosoricine ‘moonrats’ — that represent basal lineages important for scrutinizing the support for Afrotheria. Morphology-based analyses are often drawn, either explicitly or implicitly, from much larger sets of relevant taxa. Moreover, the primacy of morphological data for non-living taxa is unchallenged [3,5,10]. In organisms with little or no fossil record, we are of course compelled to rely overwhelmingly on morphological and molecular data on living forms. But fossil evidence has a powerful reality in groups like the mammals, where 75% of mammalian diversity is known only from fossils [11] and is thus inaccessible to molecular analysis. We have yet to muster both the morphological and molecular data in combined form — in super matrices — that are instructive to broad-scale placental mammal phylogeny, yet these approaches have been applied to important sectors of the tree [5].

Another problem with the more recent molecular analyses concerns the connections between data and the choice of method or assumption for the analysis. Madsen *et al.* [2] constrained their data to ensure that many of the end taxa group into traditional mammalian families or higher taxa. However, an unconstrained run of the data should be consistent with the clades proposed, if the signal for such groupings is a strong one. In the much more comprehensive analysis of Murphy *et al.* [6], Afrotheria is supported under maximum likelihood assumptions about the tempo and mode of gene sequence change. This group collapses, however, when either weighted or unweighted maximum parsimony is applied. Finally, despite the congruence claimed in these papers, the results do show marked inconsistencies. The six-gene trees of Madsen *et al.* [2] recognize three main clades of placental mammals, while the 18-gene trees of Murphy *et al.* [6] recognize (in the maximum likelihood and neighbor joining trees) at least four main clades. Given these limitations, it seems that the lack of agreement on certain higher-level placental groups remains a serious matter for both molecular and morphological studies.

The introduction of massive new molecular data sets relevant to this problem is duly recognized, but some of the claims for the thunderous impact of these data seem premature. Proposals based on these studies, such as revised mammalian classifications or biogeographic scenarios, only serve to confuse matters. On the positive side, such recent molecular work sets the stage for a new and exciting phase

of research. The mammalian tree can now be tested with data that encompass a virtually unmatched range of comparative gene information and one of the most enriched fossil records of any organismic group.

References

1. Norell MA, Novacek MJ: **The fossil record and evolution: comparing cladistics and paleontological evidence for vertebrate history.** *Science* 1992, **255**:1609-1693.
2. Madsen O, Scally M, Douady C, Kao DJ, DeBry RW, Adkins R, Amrine H, Stanhope MJ, de Jong WW, Springer MS: **Parallel adaptive radiations in two major clades of placental mammals.** *Nature* 2001, **409**:610-614.
3. Novacek MJ: **Mammalian phylogeny: shaking the tree.** *Nature* 1992, **356**:121-125.
4. Gatesy J, O'Grady P, Baker RH: **Corroboration among data sets in simultaneous analysis: hidden support for phylogenetic relationships among higher level artiodactyl taxa.** *Cladistics* 1999, **15**:271-313.
5. O'Leary MA: **Parsimony analysis of total evidence from extinct and extant taxa and the cetacean-artiodactyl question (Mammalia, Ungulata).** *Cladistics* 1999, **15**:315-330.
6. Murphy WJ, Eizirik E, Johnson WE, Zhang YP, Ryder OA, O'Brien SJ: **Molecular phylogenetics and the origins of placental mammals.** *Nature* 2001, **409**:614-618.
7. Liu F-G R, Miyamoto M M, Freire NP, Ong PQ, Tennat MR, Young TS, Gugel KF: **Molecular and morphological supertrees for eutherian (placental) mammals.** *Science* 2001, **291**:1786-1789.
8. Kluge AG, Wolfe AJ: **Cladistics: what's in a word?** *Syst Biol* 1993, **9**:183-199.
9. Wadell PJ, Okado, N, Hasegawa, M: **Toward resolving the interordinal relationships of placental mammals.** *Syst Biol* 1999, **48**:1-5.
10. Springer MS, De Jong WW: **Which mammalian supertree to bark up?** *Science* 2001, **291**:1709-1711.
11. McKenna MC, Bell SK: *Classification of Mammals above the Species Level.* Columbia University Press: New York; 1997.