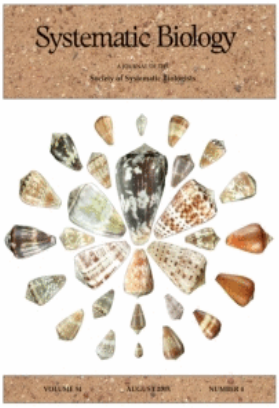

Systematic Biology
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713658732

Morphology and Placental Mammal Phylogeny

Mark S. Springer [a]; Robert W. Meredith [a]; Eduardo Eizirik [b]; Emma Teeling [c]; William J. Murphy [d]
[a] Department of Biology, University of California, Riverside, CA, USA
[b] Faculdade de Biociencias, PUCRS, Porto Alegre, RS, Brazil
[c] School of Biological and Environmental Sciences, University College Dublin Belfield, Dublin, Ireland
[d] Department of Veterinary Integrative Biosciences, Texas A&M University, College Station, TX, USA

First Published on: 01 June 2008

To cite this Article: Springer, Mark S., Meredith, Robert W., Eizirik, Eduardo, Teeling, Emma and Murphy, William J. (2008) 'Morphology and Placental Mammal Phylogeny', Systematic Biology, 57:3, 499 — 503

To link to this article: DOI: 10.1080/10635150802164504
URL: http://dx.doi.org/10.1080/10635150802164504

Points of View

Syst. Biol. 57(3):499–503, 2008

ISSN: 1063-5157 print / 1076-836X online
DOI: 10.1080/10635150802164504

Morphology and Placental Mammal Phylogeny

MARK S. SPRINGER,[1] ROBERT W. MEREDITH,[1] EDUARDO EIZIRIK,[2] EMMA TEELING,[3] AND WILLIAM J. MURPHY[4]

[1]*Department of Biology, University of California, Riverside, CA 92521, USA; E-mail: mark.springer@ucr.edu (M.S.S.)*
[2]*Faculdade de Biociencias, PUCRS, Porto Alegre, RS 90619-900, Brazil*
[3]*School of Biological and Environmental Sciences, University College Dublin Belfield, Dublin 4, Ireland*
[4]*Department of Veterinary Integrative Biosciences, Texas A&M University, College Station, TX 77843-4458, USA*

In a recent article on placental mammal phylogeny (Springer et al., 2007), we discussed evidence for correlated character evolution among morphological characters. We also performed pseudoextinction analyses that assessed whether placental orders remained in the expected superordinal group (Afrotheria, Xenarthra, Euarchontoglires, Laurasiatheria) when molecular and soft-tissue data were coded as missing and only osteological data from Asher et al. (2003) remained for the pseudoextinct taxa. Finally, we examined congruence among 21 molecular data partitions and Asher et al.'s (2003) morphological data. Our results demonstrated that most placental orders moved to a different superordinal group when treated as pseudoextinct and also that Asher et al.'s (2003) morphological data consistently emerged as the most incongruent data partition. Based on these results, we questioned the ability of current morphological data sets and phylogenetic methods to reconstruct higher level relationships among placental mammals. In their response to our paper, Asher et al. (2008) raise several objections including (1) continued debate over our "preferred 4-clade topology" (p. 311) that renders our conclusions on morphology "premature" (p. 311); (2) basing our conclusions "on a single morphological data set" (p. 312); (3) our use of pseudoextinction techniques "to make broad generalizations about the quality of data for mammal phylogeny reconstruction" (p. 313); and (4) our conditional acknowledgement of the primacy "of morphological data to infer phylogeny of fossil taxa" (p. 313). Asher et al. (2008) make a number of useful points, but as discussed below these do not diminish the main conclusions of our earlier paper.

CONGRUENCE AND THE MAJOR CLADES OF PLACENTAL MAMMALS

We agree with Asher et al. (2008) that the placental tree is not fully resolved and that debate continues to surround local polytomies such as the root of the placental tree. Some recent molecular studies (Cannarozzi et al., 2007; Huttley et al., 2007) contradict one or more of the four major clades, but these studies typically have sparse taxon sampling for key taxa such as murid rodents that have accelerated rates of molecular evolution. Phylogenetic analyses that only include one or two murid rodents are especially susceptible to long-branch attraction and typically root the placental tree on the long murid branch. Indeed, this problem has long plagued molecular studies of placental mammal phylogeny and is the basis for the "guinea pig is not a rodent" phenomenon (D'Erchia et al., 1996). Nuclear and mitogenomic studies with adequate taxon sampling recover the four clades listed earlier with robust support. Analyses of nuclear genes that support the four clades include Madsen et al. (2001), Murphy et al. (2001a, 2001b), Scally et al. (2001), Waddell et al. (2001), Delsuc et al. (2002), Amrine-Madsen et al. (2003), Fleming et al. (2003), Waddell and Shelley (2003), Roca et al. (2004), Hallstrom et al. (2007), Nikolaev et al. (2007), and Wildman et al. (2007). Analyses of mitochondrial genes that support the four clades include Hudelot et al. (2003), Waddell and Shelley (2003), Reyes et al. (2004), Kitazoe et al. (2005), and Kjer and Honeycutt (2007). Each of the four clades is also supported by a wealth of indels and retrotransposon insertions (van Dijk et al., 1999; Poux et al., 2002, de Jong et al., 2003; Thomas et al., 2003; Murphy et al., 2004; Springer et al., 2004; Kriegs et al., 2006; Nishihara et al., 2006; Waters et al., 2007). Transposable element insertions are particularly compelling, especially when they corroborate previously well-supported hypotheses based on other types of data. Unlike analyses based on nucleotide sequences, studies of transposable elements are not expected to suffer from long-branch attraction (Lunter, 2007). We maintain that robust, independent support from nuclear, mitogenomic, and genomic analyses make the four clades among the best-supported groups in all of mammalian classification and also that these clades provide an appropriate benchmark for gauging the reliability of morphological characters in

higher level placental phylogenetics. Asher et al. (2008: 312) imply that Springer et al. (2001) may support the association of Chiroptera + Dermoptera as sister taxa and consequently reject the four-clade hypothesis based on Springer et al.'s (2001) analysis of a morphological data set (Simmons and Geisler, 1998) that contained only archontan taxa. Springer et al. (2001) were aware of the potential homoplasy in using flying lemurs as an outgroup to bats; however, they had no choice but to use this data set as it was the most extensive morphological data set for bats at the time. Indeed, Gunnel and Simmons (2005) were also aware of this potential problem and expanded the Simmons and Geisler (1998) data set by scoring more laurasiatherian outgroups for bats.

Asher et al.'s (2008:311) criticism that we used a "single topology to question an entire class of data (e.g., 'morphology') as premature" is a misinterpretation of our results. Our pseudoextinction analyses were not dependent on a single topology. Rather, analyses that treated single orders as pseudoextinct assessed "whether each pseudoextinct order was recovered as monophyletic" and also whether "each pseudoextinct order remained in the same superordinal group (Afrotheria, Xenarthra, Laurasiatheria, Euarchontoglires) as in Figure 1 or moved elsewhere on the tree" (Springer et al. 2007:676). Both intraordinal and interordinal relationships were free to vary in pseudoextinction analyses and were not constrained to follow the topology shown in our Figure 1. Data congruence analyses were based on comparisons of the shortest tree(s) for 22 different data partitions. The finding that morphological data consistently emerged as the most incongruent data partition is not dependent on a single topology.

Asher et al. (2008) argue that some individual genes (e.g., *VWF*) fail to support the four-clade classification of placental mammals and suggest that these genes should not be eschewed just as morphology should not be eschewed. This comparison between individual genes and morphology ignores a fundamental difference between single genes and Asher et al.'s (2003) morphological data set. Namely, that single genes, whether they support the four-clade classification or not, remain more congruent with each other than they do with morphology. This is evident from figure 2 in Springer et al. (2007). Further, robust support for the four-clade classification is recovered when individual genes that fail to support the four-clade classification are analyzed together. This is precisely what we should expect if phylogentic signal is embedded in a background of random homoplastic noise. That is, we expect that random noise will sometimes exceed phylogenetic signal for finite data and that this problem will be more acute as the number of informative characters decreases and also for increasingly shorter internal branches. However, if background noise is random, then phylogenetic signal will overcome this noise as the data set becomes larger (e.g., single genes are concatenated into multigene data sets). Short internal branches will typically require more data for phylogenetic signal to overcome random noise. In contrast to phylogenetic signal against a backdrop of random, homoplastic noise, pervasive correlated character evolution in response to developmental constraints and/or ecological venue may be decidedly nonrandom and generate competing signal that is not overcome by phylogenetic signal simply by adding more and more characters to the data set. There is no doubt that morphology contains phylogenetic signal at the level of interordinal relationships (e.g., Paenungulata, Glires). However, the emerging pattern of incongruence between morphological and molecular data (Springer et al., 2007: fig. 2), in conjunction with numerous

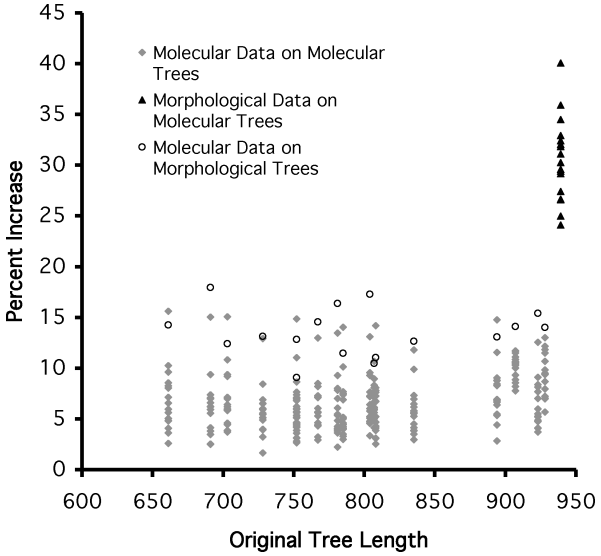

FIGURE 1. Plot of minimum tree length for 18 different data partitions (17 molecular, 1 morphological; *x*-axis) versus the increase in tree length when each data partition is mapped onto the best tree(s) for each of the other data partitions (*y*-axis). The molecular data set included four sequences new to this study (accession numbers EU448989 to EU448992), seven sequences that were obtained from genome assemblies and trace files, and nine additional sequences that were obtained from GenBank (see online Supplemental Material at www.systematicbiology.org). New sequences were manually aligned to the Springer et al. (2007) data set. Each data partition contained 194 informative characters. Molecular data were treated as unordered; morphological data were treated as a combination of ordered and unordered characters following Horovitz's (2004) description of each character. In cases where a partition was mapped onto more than one equally most parsimonious tree for another data partition, we plotted the midpoint value for the percentage increase in tree length. The 17 molecular partitions (P) arbitrarily followed the sequential gene order in the concatenated molecular data set of Springer et al. (2007), irrespective of gene boundaries, and included characters from the following gene segments: P1 (*ADRA2B*); P2 (*ADRA2B, ADORA3*); P3 (*ADORA3, ADRB2*); P4 (*ADRB2, APOB, APP*); P5 (*APP, ATP7A*); P6 (*ATP7A, BDNF*); P7 (*BDNF, BRCA1*); P8 (*BRCA1*); P9 (*BRCA1*); P10 (*BRCA1*); P11 (*BRCA1*); P12 (*BRCA1, CNR1*); P13 (*CNR1, CREM*); P14 (*CREM, GHR*); P15 (*GHR*); P16 (*PLCB4, VWF*); P17 (*VWF*). PAUP 4.0b11 (Swofford, 2003) was used to find the most parsimonious tree(s) for each data partition. We employed heuristic searches with tree-bisection-and-reconnection branch swapping and 1000 randomized taxon input orders. The 30 taxa included in partition congruence analyses were sloth, anteater, armadillo, hedgehog, mole, shrew, tenrec, golden mole, hyrax, elephant, elephant shrew, aardvark, murid, hystricognath, rabbit, pika, flying lemur, tree shrew, strepsirrhine, human, tarsier, microchiropteran, megachiropteran, llama, pig, horse, ceratomorph, cat, caniform, and pangolin.

instances of ecological vicars that group together on published morphological trees (e.g., Shoshani and McKenna, 1998: fig. 1; Asher et al., 2003: fig. 4; Horovitz et al., 2004: fig. 1), suggest that morphological data contain a mixture of homologous characters and correlated homoplastic characters that are not effectively separated from each other by current phylogenetic methods.

Use of a Single Morphological Data Set

We chose to analyze the Asher et al. (2003) data set because it was the largest available morphological data set to include representatives of all extant orders of placental mammals. Asher et al.'s (2003) data set was ideally suited for making comparisons with our molecular data because Asher et al. (2003:132) selected extant taxa "based on the degree to which each was represented by the 22 genes used in Murphy et al. (2001b), most of which were in turn first published by Madsen et al. (2001) and Murphy et al. (2001a)." Asher et al. (2008) object that our conclusions are based on a single data set and suggest several other recent data sets (i.e., Horovitz, 2004; Luo and Wible, 2005; Wible et al., 2007) that may be appropriate for making comparisons. None of these data sets was collected with the explicit intention to maximize taxonomic overlap with the molecular data of Murphy et al. (2001b). All of these data sets also are missing representation from one or more placental orders. Most perplexing is the taxon sampling in Wible et al. (2007), which notably excludes four placental orders (Chiroptera, Dermoptera, Perissodactyla, and Pholidota) that were present in Asher et al. (2003). These same orders were identified by our pseudoextinction analyses to be among the most problematic, and all four moved to different superordinal groups in the analyses. Wible et al. (2007) provide no explanation for these missing representatives, but results of our pseudoextinction analyses suggest that inclusion of these taxa would likely have rendered all four, rather than just two, of the molecular-defined clades poly- or paraphyletic in their study.

Horovitz (2004) is the most complete of the data sets cited by Asher et al. (2008) and shares 30 placental taxa with Springer et al. (2007). We performed congruence analyses for 13 of the 20 nuclear genes in Springer et al. (2007) and the postcranial data set of Horovitz (2004) using our previously described methodology. Morphological data once again emerged as the most incongruent data partition after standardizing partitions so that they contained the same number of informative characters (Fig. 1). We also performed congruence tests that compared molecular sequences for five nuclear genes (*APOB*, *RAG*1, *BRCA*1, *VWF*, *IRBP*) to three different morphological data sets for marsupials (Horovitz and Sánchez-Villagra, 2003; Luo et al., 2003; Sánchez-Villagra et al., 2007) that include representation for the seven extant marsupial orders and obtained results that are similar to those for our comparisons involving placental mammals (Meredith, 2007). Thus, the pattern of incongruence between molecular and morphological data extends beyond Asher et al.'s (2003) morphological data set to other higher level morphological data sets for both placentals and marsupials. The morphological data sets that we have examined for congruence with molecular data are limited to osteological data and it remains possible that soft-tissue morphological characters will prove more compatible with molecular data. Nevertheless, osteological characters are the primary type of data that are available for fossil mammals.

Pseudoexintction Techniques

Asher et al. (2008:313) argue that pseudoextinction "can be used to roughly gauge the performance of a subset of data in placing extinct taxa in a phylogeny, but is insufficient to make broad generalizations about the quality of data for mammal phylogeny reconstruction." In support of their conclusion, Asher et al. point out that "morphology is reliable in recognizing orders" (p. 313) and also that "some superordinal groups were first recognized at least in part (Archonta, Cetartiodactyla) or entirely (Glires, Paenungulata) on morphological evidence" (p. 313). Asher et al. are correct that most placental orders and some superordinal groups (e.g., Glires and Paenungulata) were first recognized based on morphology. Morphology clearly provides useful information for some mammalian superorders. However, our fundamental thesis is that morphological cladistic studies addressing higher level placental relationships conflate homology and homoplasy at the level of superordinal placental groups. This problem is not trivial based on our pseudoextinction analyses with Asher et al.'s (2003) morphological data set. Specifically, the majority of placental orders moved to a different superordinal group when molecular data were coded as missing and the four superordinal groups were never recovered as monophyletic. Minimally, this result demands careful consideration prior to combining molecular and morphological data to examine higher level placental relationships.

Primacy of Morphology for Extinct Taxa

Morphological data are of fundamental importance for constructing hypotheses that place extinct mammals in a phylogenetic context. Asher et al. (2008) correctly note that morphology has not been fully exploited and also that morphological data "remain among the best (and often only) means by which we can incorporate the real data afforded to us via the fossil record into reconstructing the mammalian component of the Tree of Life" (p. 315). However, the primacy of morphological data for reconstructing relationships of extinct forms does not excuse morphology from critical assessments of its performance in analyses of higher level placental mammal phylogenetics. A robust solution for the mammalian component of the Tree of Life that includes living and extinct taxa must address problems of incongruence between molecular and morphological data and potential problems of correlated homoplasy among morphological characters. The problem of teasing homplastic

morphological markers from homologous morphological markers will not simply disappear if it is ignored.

ACKNOWLEDGMENTS

M.S.S. and W.J.M. acknowledge support from NSF AToL program. We thank Dr. Ines Horovitz for providing her 2004 data set in nexus format. Mark Hafner and Jack Sullivan provided helpful comments on this manuscript.

REFERENCES

Amrine-Madsen, H., K.-P. Koepfli, R. K. Wayne, and M. S. Springer. 2003. A new phylogenetic marker, apolipoprotein B, provides compelling evidence for eutherian relationships. Mol. Phylogenet. Evol. 28:225–240.

Asher, R. J., J. H. Geisler, and M. R. Sánchez-Villagra. 2008. Morphology, paleontology, and placental mammal phylogeny. Syst. Biol. 57:311–317.

Asher, R. J., M. J. Novacek, and J. G. Geisler. 2003. Relationships of endemic African mammals and their fossil relatives based on morphological and molecular evidence. J. Mamm. Evol. 10:131–194.

Cannarozzi, G., A. Schneider, and G. Gonnet. 2007. A phylogenomic study of human, dog, and mouse. PLoS Comput. Biol. 3:9–14.

D'Erchia, A. M., C. Gissi, G. Pesole, C. Saccone, and U. Arnason. 1996. The guinea-pig is not a rodent. Nature 381:597–600.

de Jong, W. W., M. A. M. van Dijk, C. Poux, G. Kappe, T. van Rheede, and O. Madsen. 2003. Indels in protein-coding sequences of Euarchontoglires constrain the rooting of the eutherian tree. Mol. Phylogenet. Evol. 28:328–340.

Delsuc, F., M. Scally, O. Madsen, M. J. Stanhope, W. W. de Jong, F. M. Catzeflis, M. S. Springer, and E. J. P. Douzery. 2002. Molecular phylogeny of living xenarthrans and the impact of character and taxon sampling on the placental tree rooting. Mol. Biol. Evol. 19:1656–1671.

Fleming, M. A., J. D. Potter, C. J. Ramirez, G. K. Ostrander, and E. A. Ostrander. 2003. Understanding missense mutations in the BRCA1 gene: An evolutionary approach. Proc. Natl. Acad. Sci. USA 100:1151–1156.

Gunnel, G. F., and N. B. Simmons. 2005. Fossil evidence and the origin of bats. J. Mamm. Evol. 12:209–246.

Hallstrom, B. M., M. Kullberg, M. A. Nilsson, and A. Janke. 2007. Phylogenomic data analyses provide evidence that Xenarthra and Afrotheria are sister groups. Mol. Biol. Evol. 24:2059–2068.

Horovitz, I. 2004. Eutherian mammal systematics and the origins of South American ungulates as based on postcranial osteology. Bull. Carneg. Mus. Nat. Hist. 36:63–79.

Horovitz, I., and M. R. Sánchez-Villagra. 2003. A morphological analysis of marsupial mammal higher-level phylogenetic relationships. Cladistics 19:181–212.

Hudelot, C., V. Gowri-Shankar, H. Jow, M. Rattray, and P. G. Higgs. 2003. RNA-based phylogenetic methods: Application to mammalian mitochondrial RNA sequences. Mol. Phylogenet. Evol. 28:241–252.

Huttley, G. A., M. J. Wakefield, and S. Easteal. 2007. Rates of genome evolution and branching order from whole genome analysis. Mol. Biol. Evol. 24:1702–1730.

Kitazoe, Y., H. Kishino, T. Okabayashi, T. Watabe, N. Nakajima, Y. Okuhara, and Y. Kurihara. 2005. Multidimensional vector space representation for convergent evolution and molecular phylogeny. Mol. Biol. Evol. 22:704–715.

Kjer, K. M., and R. L. Honeycutt. 2007. Site specific rates of mitochondrial genomes and the phylogeny of Eutheria. BMC Evol. Biol. 7:8.

Kriegs, J. O., G. Churakov, M. Keifmann, U. Jordan, J. Brosius, and J. Schmitz. 2006. Retroposed elements as archives for the evolutionary history of placental mammals. PLoS Biol. 4:e91.

Lunter, G. 2007. Dog as an outgroup to human and mouse. PLoS Comput. Biol. 3:e74.

Luo, Z.-X., and J. R. Wible. 2005. A Late Jurassic digging mammal and early mammalian diversification. Science 308:103–107.

Luo, Z.-X, Q. Ji, J. R. Wible, and C.-Y. Yuan. 2003. An early Cretaceous tribosphenic mammal and metatherian evolution. Science 203:1934–1940.

Madsen, O., M. Scally, C. J. Douady, D. J. Kao, W. DeBry, R. Adkins, H. Amrine, M. J. Stanhope, W. W. de Jong, and M. S. Springer. 2001. Parallel adaptive radiations in two major clades of placental mammals. Nature 409:610–614.

Meredith, R. W. 2007. Phylogeny and evolution of diprotodontian marsupials. PhD Dissertation, University of California, Riverside.

Murphy, W. J., E. Eizirik, W. E. Johnson, Y. P. Zhang, O. A. Ryder, and S. J. O'Brien. 2001a. Molecular phylogenetics and the origins of placental mammals. Nature 409:614–618.

Murphy, W. J., E. Eizirik, S. J. O'Brien, O. Madsen, M. Scally, C. J. Douady, E. Teeling, O. A. Ryder, M. J. Stanhope, W. W. de Jong, and M. S. Springer. 2001b. Resolution of the early placental mammal radiation using Bayesian phylogenetics. Science 294:2348–2351.

Murphy, W. J., P. A. Pevzner, and S. J. O'Brien. 2004. Mammalian phylogenomics comes of age. Trends Genet. 20:631–639.

Nikolaev, S., J. I. Montoya-Burgos, E. H. Margulies, NISC Comparative Sequencing Program, J. Rougemont, B. Nyffeler, and S. E. Antonarakis. 2007. Early history of mammals is elucidated with the ENCODE multiple species sequencing data. PLoS Genet. 3:e2.

Nishihara, H., M. Hasegawa, and N. Okada. 2006. Pegasoferae, an unexpected mammalian clade revealed by tracking acient retroposon insertions. Proc. Natl. Acad. Sci. USA 103:9929–9934.

Poux, C., van Rheede, T., O. Madsen, and W. W. de Jong. 2002. Sequence gaps join mice and men: Phylogenetic evidence from deletions in two proteins. Mol. Biol. Evol. 19:2035–2037.

Reyes, A., C. Gissi, F. Catzeflis, E. Nevo, G. Pesole, and C. Saccone. 2004. Congruent mammalian trees from mitochondrial and nuclear genes using Bayesian methods. Mol. Biol. Evol. 21:397–403.

Roca, A., G. K. Bar-Gal, E. Eizirik, K. M. Helgen, R. Maria, M. S. Springer, S. J. O'Brien, and W. J. Murphy. 2004. Mesozoic origin for West Indian insectivores. Nature 429:649–651.

Sánchez-Villagra, M., S. Ladeveze, I. Horovitz, C. Argot, J. J. Hooker, T. E. Macrini, T. Martin, S. Moore-Fay, C. de Muizon, T. Schmelzle, and R. J. Asher. 2007. Exceptionally preserved North American Paleogene metatherians: Adaptations and discovery of a major gap in the opossum fossil record. Biol. Lett. 3:318–322.

Scally, M., O. Madsen, C. J. Doouady, W. W. de Jong, M. J. Stanhope, and M. S. Springer. 2001. Molecular evidence for the major clades of placental mammals. J. Mamm. Evol. 8:239–277.

Shoshani, J., and M. C. McKenna. 1998. Higher taxonomic relationships among extant mammals based on morphology, with selected comparisons of results from molecular data. Mol. Phylogenet. Evol. 9:572–584.

Simmons, N. B., and J. H. Geisler. 1998. Phylogenetic relationships of *Icaronycteris*, *Archaeonycteris*, *Hassianycteris*, and *Palaeochiropteryx* to extant bat lineages, with comments on the evolution of echolocation and foraging strategies in Microchiroptera. Bull. Am. Mus. Nat. Hist. 235:1–182.

Springer, M. S., A. Burk-Herrick, R. Meredith, E. Eizirik, E. Teeling, S. J. O'Brien, and W. J. Murphy. 2007. The adequacy of morphology for reconstructing the early history of placental mammals. Syst. Biol. 56:673–684.

Springer, M. S., M. J. Stanhope, O. Madsen, and W. W. de Jong. 2004. Molecules consolidate the placental mammal tree. Trends Ecol. Evol. 19:430–438.

Springer, M. S., E. C. Teeling, O. Madsen, M. J. Stanhope, and W. W. de Jong. 2001. Integrated fossil and molecular data reconstruct bat echolocation. Proc. Natl. Acad. Sci. USA 98:6241–6246.

Swofford, D. L. 2003. PAUP*: Phylogenetic analysis using parsimony (*and other methods), Version 4.0b10. Sinauer Associates, Sunderland, Massachusetts.

Thomas, J. W., et al. (70 other authors). 2003. Comparative analyses of multi-species sequences from targeted genomic regions. Nature 424:788–793.

van Dijk, M. A. M., E. Paradis, F. Catzeflis, and W. W. de Jong. 1999. The virtues of gaps: Xenarthran (edentate) monophyly supported by a unique deletion in αA-crystallin. Syst. Biol. 48:94–106.

Waddell, P. J., H. Kishino, and R. Ota. 2001. A phylogenetic foundation for comparative mammalian genomics. Genome Informatics 12:141–154.

Waddell, P. J., and S. Shelley. 2003. Evaluating placental interordinal phylogenies with novel sequences including RAG1, γ-fibrinogen, ND6, and mt-tRNA, plus MCMC-driven nucleotide, amino acid, and codon models. Mol. Phylogenet. Evol. 28:197–224.

Waters, P. D., G. Dobigny, P. J. Waddell, and T. J. Robinson. 2007. Evolutionary history of LINE-1 in the major clades of placental mammals. PLoS ONE 2:e158.

Wible, J. R., G. W. Rougier, M. J. Novacek, and R. J. Asher. 2007. Cretaceous eutherians and Laurasian origin for placental mammals near the K/T boundary. Nature 447:1003–1006.

Wildman, D. E., M. Uddin, J. C. Opazo, G. Liu, V. Lefort, S. Guindon, O. Gascuel, W. I. Grossman, R. Romero, and M. Goodman. 2007. Genomics, biogeography, and the diversification of placental mammals. Proc. Natl. Acad. Sci. USA 104:14395–14400.

First submitted 23 January 2008; reviews returned 4 February 2008; final acceptance 14 March 2008
Associate Editor: Mark Hafner

Syst. Biol. 57(3):503–507, 2008

ISSN: 1063-5157 print / 1076-836X online
DOI: 10.1080/10635150802164587

Hemiplasy: A New Term in the Lexicon of Phylogenetics

JOHN C. AVISE[1] AND TERENCE J. ROBINSON[2]

[1]*Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92697, USA; E-mail, javise@uci.edu*
[2]*Evolutionary Genomics Group, Department of Botany and Zoology, University of Stellenbosch, Private Bag X1, Matieland 7602, South Africa*

Homoplasy (trait similarity due to evolutionary convergence, parallelism, or character reversals) is a well-appreciated form of phylogenetic noise that systematists strive to identify and avoid when reconstructing species phylogenies. However, another source of phylogenetic "noise" is often neglected: the idiosyncratic sorting of gene-tree lineages into descendant taxa from character-state polymorphisms retained across successive nodes in a species tree. Here we introduce a term (hemiplasy) that formalizes a category of outcomes that can emerge from this evolutionary lineage-sorting phenomenon, and we make a case for why a wider recognition of hemiplasy (and attempts to ameliorate its complications) can play an important role in phylogenetics.

The word *homoplasy*, meaning shaped (-plasy) in the same (homo-) way, refers to any trait correspondence or similarity not due to common ancestry. A central challenge in phylogenetic reconstruction is thus to distinguish the phylogenetic noise of homoplasy from the phylogenetic signal of homology (similarity in biological features due directly to shared ancestry). However, homology itself bears a subtle relationship to phylogeny, as emphasized by Willi Hennig (1950) more than a half-century ago. Hennig introduced the critical distinction between shared ancestral homology (symplesiomorphic similarity) and shared derived homology (synapomorphic similarity), noting that only the latter is indicative of monophyly within an organismal phylogeny. Hennig's cladistic insights fostered a fundamental revolution in phylogenetic principles and methodologies.

The molecular revolution in biology that began at about that same time added further nuances to the homology concept. For example, DNA sequence homology in a multigene family can be due either to paralogy (similarity tracing to a gene duplication event) or to orthology (similarity tracing to an allelic separation within a particular locus). Orthology and paralogy are both genuine forms of genetic homology, but a failure to distinguish them in comparisons of DNA sequences can lead to errors in phylogenetic reconstruction.

Phylogenetic jargon is already extensive but also important because words such as *homoplasy*, *synapomorphy*, and *orthology* capture and convey sophisticated evolutionary concepts that otherwise might remain opaque or underappreciated. In this spirit, here we formally define a new term—*hemiplasy*—for how the well-known phenomenon of idiosyncratic lineage sorting can lead to fundamental discordances between gene trees and organismal (species) trees. As will be described, hemiplasy is a bona fide form of homology (allelic orthology in this case) that nonetheless can give the illusion of homoplasy in an organismal tree. No other word or simple phrase currently exists to encapsulate the phenomenon that we will define under the suggested term.

CONCEPTUAL BACKGROUND

The nature of Mendelian heredity in sexually reproducing taxa ensures that alleles at unlinked loci transmit through an organismal pedigree via noncoincidental genealogical pathways across multiple generations. Thus, both within and among related species, the true topologies of gene trees inevitably differ somewhat from locus to unlinked locus (Ball et al., 1990). Furthermore, gene genealogies can in principle differ in basic topology from the overall population tree or species tree of which they are a part, if for no other reason than stochastic lineage sorting across successive evolutionary nodes in an organismal phylogeny. These concepts and their corollaries have been available for more than two decades (Hudson, 1983; Tajima, 1983; Takahata and Nei, 1985; Neigel and Avise, 1986), and they are encapsulated