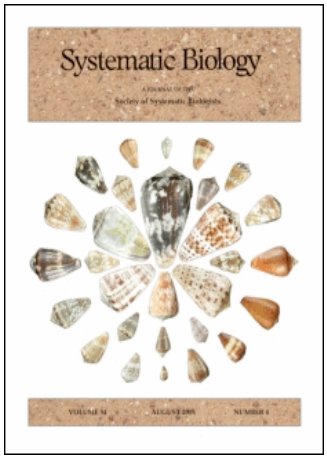

Systematic Biology
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713658732

The Adequacy of Morphology for Reconstructing the Early History of Placental Mammals

Mark S. Springer [a]; Angela Burk-Herrick [a]; Robert Meredith [a]; Eduardo Eizirik [b]; Emma Teeling [c]; Stephen J. O'Brien [d]; William J. Murphy [e]
[a] Department of Biology, University of California, Riverside, Riverside, CA, USA
[b] Faculdade de Biociencias, PUCRS, Porto Allegre, RS, Brazil
[c] School of Biological and Environmental Sciences, University College Dublin Belfield, Dublin, Ireland
[d] Laboratory of Genomic Diversity, National Cancer Institute-Frederick, Frederick, MD, USA
[e] Department of Veterinary Integrative Biosciences, Texas A&M University, College Station, TX, USA

First Published on: 01 August 2007
To cite this Article: Springer, Mark S., Burk-Herrick, Angela, Meredith, Robert, Eizirik, Eduardo, Teeling, Emma, O'Brien, Stephen J. and Murphy, William J. (2007) 'The Adequacy of Morphology for Reconstructing the Early History of Placental Mammals', Systematic Biology, 56:4, 673 - 684
To link to this article: DOI: 10.1080/10635150701491149
URL: http://dx.doi.org/10.1080/10635150701491149

Points of View

Syst. Biol. 56(4):673–684, 2007

ISSN: 1063-5157 print / 1076-836X online
DOI: 10.1080/10635150701491149

The Adequacy of Morphology for Reconstructing the Early History of Placental Mammals

MARK S. SPRINGER,[1] ANGELA BURK-HERRICK,[1] ROBERT MEREDITH,[1] EDUARDO EIZIRIK,[2] EMMA TEELING,[3] STEPHEN J. O'BRIEN,[4] AND WILLIAM J. MURPHY[5]

[1]*Department of Biology, University of California, Riverside, Riverside, CA 92521, USA; E-mail: mark.springer@ucr.edu (M.S.S.)*
[2]*Faculdade de Biociencias, PUCRS, Porto Allegre, RS 90619-900, Brazil*
[3]*School of Biological and Environmental Sciences, University College Dublin Belfield, Dublin 4, Ireland*
[4]*Laboratory of Genomic Diversity, National Cancer Institute–Frederick, Frederick, MD 21702, USA*
[5]*Department of Veterinary Integrative Biosciences, Texas A&M University, College Station, TX 77843-4458, USA*

Mammalian anatomists and paleontologists working primarily with osteological data have long been intrigued with the problem of eutherian diversification, including both interordinal relationships and the timing of the placental mammal radiation (McKenna, 1975; Novacek, 1992). Cladistic analyses of morphological characters for placental mammal orders have suggested a variety of superordinal hypotheses (Table 1). Molecular trees based on single gene segments were often in conflict with each other and with morphology, but larger nuclear gene data sets that include longer and/or multiple gene segments have converged on a well-supported superordinal tree topology that divides placental orders into four major groups: Afrotheria, Xenarthra, Laurasiatheria, and Euarchontoglires (Madsen et al., 2001; Murphy et al., 2001a, 2001b; Scally et al., 2001; Amrine-Madsen et al., 2003). Analyses of independent molecular and genomic data sets, specifically whole mitochondrial genomes and rare genomic changes (RGCs), are congruent with this four-clade classification (Hudelot et al., 2003; Waddell and Shelley, 2003; Murphy et al., 2004; Reyes et al., 2004; Springer et al., 2004, 2005; Gibson et al., 2005; Kriegs et al., 2006; Nishihara et al., 2006; Kjer and Honeycutt, 2007). For example, Nishihara et al. (2006) found six, nine, and nine L1 retroposon insertions supporting the monophyly of Afrotheria, Euarchontoglires, and Laurasiatheria, respectively.

Whereas nuclear, mitochondrial, and RGC data corroborate each other in supporting the four major clades, only Xenarthra had been previously hypothesized based on morphology. In addition, the morphology-based clades Altungulata, Anagalida, Archonta, Edentata, Lipotyphla, Ungulata, and Volitantia (Table 1) are all incompatible with the four-clade classification of mammals derived from molecular and genomic data. The molecular/genomic tree also suggests that morphology-based groups of placental orders (Novacek, 2001) are often homoplastic constellations of taxa that have evolved independently in separate geographic venues rather than monophyletic assemblages that descended from a common ancestor.

Scotland et al. (2003) argued for a more limited role for morphology-based phylogenetic analyses in reconstructing the tree of life. Other authors (Jenner, 2004; Wiens, 2004; Smith and Turner, 2005) have offered detailed critiques of Scotland et al.'s (2003) main thesis and maintain the view that morphological data remain "crucial in reconstructing the phylogeny of the earth's biota" (Smith and Turner, 2005:171). In particular, Wiens (2004), Jenner (2004), and Smith and Turner (2005) all note the importance of morphological data for reconstructing relationships of fossil taxa. We also recognize the primacy of morphological data for reconstructing relationships of extinct taxa. However, the failure of morphological data alone to recover a tree compatible with the four-clade interordinal partitioning of placental mammals raises serious doubts about the ability of current morphological cladistic studies to accurately reconstruct relationships for extinct forms.

Correlated character evolution related to diet and/or locomotion in independent lineages is perhaps the most dangerous pitfall of morphological cladistics, where instead the implicit assumption is that morphological characters evolve independently of each other in separate lineages. This assumption may be warranted for some characters, but there are also studies of the genetics of skeletal variation in mammals that provide evidence for positively and negatively correlated traits in skeletal size and shape. Chase et al. (2002) examined variation in the canid skeleton and found that individuals with relatively small pelvic girdles and lumbar vertebrae also tend to have large attachment sites for jaw and neck muscles, relatively large posterior faces, and small anterior faces. More generally, this finding implies that the

TABLE 1. Superordinal groups based on morphological data.

Superordinal group	Taxa
Altungulata	Paenungulata, Perissodactyla
Anagalida	Glires, Macroscelidea
Archonta	Dermoptera, Chiroptera, Primates, Scandentia
Edentata	Xenarthra, Pholidota
Glires	Rodentia, Lagomorpha
Lipotyphla	Afrosoricida, Eulipotyphla
Paenungulata	Hyracoidea, Proboscidea, Sirenia
Tethytheria	Proboscidea, Sirenia
Ungulata	Altungulata, Cetartiodactyla, Tubulidentata
Volitantia	Dermoptera, Chiroptera

size and strength of the pelvic and head-neck musculoskeletal systems are inversely correlated (Chase et al., 2002). Chase et al. (2002) also found that metrics of skull and limb length are inversely correlated with metrics of skull width and height. Carrier et al. (2005) found a negative correlation between the size of the pelvis and dimensions of distal limb bones. This negative correlation represents a functional trade-off between high-speed, energy-efficient running versus limb strength. Negative correlations between pelvis size and metrics of limb robustness are also seen in the transition from *Australopithecus* (large pelvis and less robust limb bones) to *Homo* (smaller pelvis and more robust limb bones; Wolpoff, 1999).

Kangas et al. (2004) discovered evidence for correlated character evolution in their genetic study of ectodysplasin expression in mouse and found that increased expression of this single protein resulted in an increase in the number of teeth, an increase in the number of cusps on teeth, changes in the shapes and positions of cusps, and the formation of longitudinal crests. Numerous characteristics related to these variables are routinely scored in matrices of dental characters (Luo et al., 2001; Meng et al., 2003). Kangas et al. (2004:211) concluded that "most aspects of tooth shape have the developmental potential for correlated changes during evolution which may, if not taken into account, obscure phylogenetic history." Kangas et al. (2004) did not deny the potential for dental characters to change independently, but their study (p. 214) demands that "developmental nonindependence should not be excluded from the hypotheses considered in evolutionary taxonomy."

Perhaps the most spectacular example of correlated character evolution is Volitantia (Chiroptera + Dermoptera), which is supported by as many as 17 putative synapomorphies (Simmons and Geisler, 1998). However, molecular and genomic studies provide congruent support for the inclusion of chiropterans in the superordinal clade Laurasiatheria, whereas dermopterans are members of the superordinal clade Euarchontoglires (Waddell et al., 1999, 2001; Teeling et al., 2000; Murphy et al., 2001b; Amrine-Madsen et al., 2003; Reyes et al., 2004; Nishihara et al., 2006; Kjer and Honeycutt, 2007). It is now manifest that putative synapomorphies of Volitantia are homoplastic features shared in common by Chiroptera and Dermoptera and that numerous derived characters related to gliding (Dermoptera) and powered flight (Chiroptera) have evolved independently in a highly correlated fashion in these two taxa. Gunnell and Simmons (2005) provide a cogent ecological explanation: that morphological similarities shared by these taxa were derived independently due to the demands of an arboreal environment.

The emerging molecular and genomic consensus of four major groups of placental mammals has implications for early placental biogeography as Afrotheria and Xenarthra are of putative Gondwanan origin based on the fossil record of constituent orders whereas Laurasiatheria and Euarchontoglires are of putative Laurasian origin (Eizirik et al., 2001; Madsen et al., 2001). Molecular dating analyses with relaxed clock methods suggest that the four superordinal groups all diverged from each other in the Cretaceous (Hasegawa et al., 2003, Springer et al., 2003; van Rheede et al., 2006; Murphy et al., 2007). By contrast, Asher et al. (2003, 2005) and Meng et al. (2003) concluded that molecular studies that place interordinal divergences in the Cretaceous (Hasegawa et al., 2003; Springer et al., 2003) are contradicted by phylogenetic analyses that place Cretaceous eutherians outside of crown-group Placentalia. Asher et al. (2003), Zack et al. (2005), and Tabuce et al. (2007) have all suggested that Afrotheria is Holarctic in origin based on the placement of extinct northern hemisphere condylarths within and/or at the base of Afrotheria in phylogenetic analyses with combined molecular-morphological data (molecular data coded as missing for fossil taxa; Asher et al., 2003) or morphological data alone (Zack et al., 2005; Tabuce et al., 2007). For example, Zack et al. (2005, p. 498) stated "identification of macroscelidean relatives in the North American Palaeocene argues against an African origin for Afrotheria, weakening support for linking placental diversification to the break-up of Gondwana." One difficulty with Zack et al.'s (2005) conclusion is that estimating the place of origin based on the oldest afrotherian fossils, which in this case are North American aspheliscines, does not follow logically. Other variables, most importantly the topology itself, are of fundamental importance in reconstructing the place of origin. If we accept the highly nested position that Zack et al. (2005) discovered for North American aspheliscines (see Zack et al., 2005: fig. 3c), it is clear that Afrotheria must be substantially older than the aspheliscine fossils and that the origin of Afrotheria cannot be inferred from the geographic distribution of aspheliscines. A second difficulty with these studies is the assumption that morphological data, and methods for analyzing these data, give reliable phylogenetic solutions for the placement of key fossil taxa, even though morphological cladistics has so far failed to recover three of the fundamental clades of living placentals (Afrotheria, Euarchontoglires, Laurasiatheria) that are supported by multiple lines of independent molecular and genomic evidence. These same studies that question the African origin of Afrotheria based on the placement of extinct condylarths also fail to recover the monophyly of extant afrotherians in analyses based on morphology

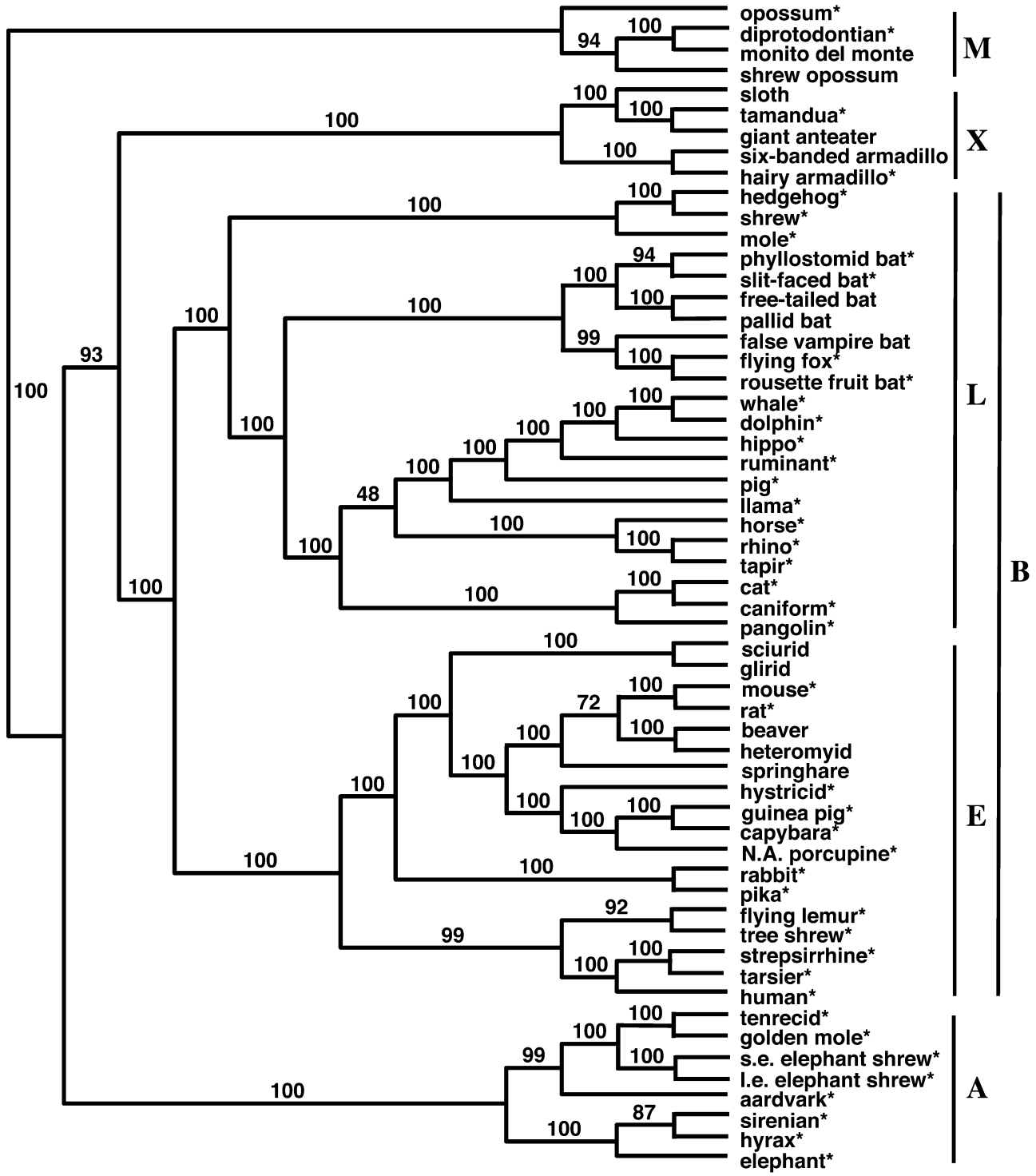

FIGURE 1. Phylogenetic tree with posterior probabilities derived from MrBayes 3.1.1 (Ronquist and Huelsenbeck, 2003) and a 14,326-bp data set (TreeBASE S1792, M3274) that included segments from 20 different nuclear genes (*ADORA3, ADRA2B, ADRB2, APOB, APP, ATP7A, BDNF, BMI, BRCA1, CNR1, CREM, EDG1, GHR, IRBP, PLCB4, RAG1, RAG2, TYR, VWF, ZFX*). Accession numbers for previously published molecular data are given in online Supplemental Material (www.systematicbiology.org). New molecular sequences were obtained as previously described (Madsen et al., 2001; Murphy et al., 2001a, 2001b; Amrine-Madsen et al., 2003). Accession numbers for new sequences are EF104981 to EF104995 (see online Supplemental Information at www.systematicbiology.org). Posterior probabilities are averages based on two independent runs. The average standard deviation of split frequencies for the two runs was 0.004. Asterisks denote 44 taxa that were included in the pseudoextinction analyses. Abbreviations as follows: M = Marsupialia; X = Xenarthra; L = Laurasiatheria; E = Euarchontoglires; A = Afrotheria; B = Boreoeutheria.

alone, and only recover a clade similar to Afrotheria by concatenating morphological characters with molecular data (Asher et al., 2003) or reducing taxon sampling to a small number of placental orders (Zack et al., 2005; Tabuce et al., 2007).

The phylogenetic placement of most mammal fossils cannot be tested directly with molecular data owing to the degradation of DNA. Nevertheless, the emergence of four superordinal clades (Afrotheria, Xenarthra, Laurasiatheria, Euarchontoglires) that are supported by analyses of nuclear gene sequences, mitochondrial genome sequences, and genome-wide screens for rare genomic changes provides an opportunity to conduct tree congruence tests with morphological data. Here, we use a novel method that combines molecular and morphological data and treats living orders as pseudoextinct to assess whether morphological data alone are sufficient to provide congruent phylogenetic results in the absence of molecular data. Unlike Scotland et al. (2003), who criticized morphological cladistics largely on the basis of the presumed behavior of morphological data (see Scotland et al., 2002: fig. 1), our pseudoextinction approach provides a test case of the behavior of real morphological data in phylogenetic analyses. Additional analyses examine whether morphological data are more incongruent with molecular data than are equivalently sized molecular data partitions with each other. Our results call for improved methods for collecting and analyzing morphological data to resolve deep-level phylogenetic relationships among extinct and living mammalian taxa.

Expanded Molecular Sampling

We performed Bayesian analyses with an expanded molecular data set that builds on the nuclear data of Murphy et al. (2001b) and Amrine-Madsen et al. (2003) (Fig. 1) and includes four marsupial outgroups, 53 placental taxa, and 20 different nuclear gene segments. Sequences for each gene segment were aligned using SOAP (Löytynoja and Milinkovitch, 2001) with 25 different combinations of gap opening and gap extension penalties following Westerman et al. (2002). A total of 14,326 aligned sites were conserved across all 25 alignments and were retained for phylogenetic analyses. The concatenated DNA alignment is available in online Supplemental Material (www.systematicbiology.org). ModelTest (Posada and Crandall, 1998) analyses were performed with each gene segment to determine the most appropriate model of sequence evolution as indicated by the Akaike Information Criterion. The 57-taxon molecular data set was analyzed with MrBayes 3.1.1 (Ronquist and Huelsenbeck, 2003; Ronquist et al., 2005), which simultaneously performs two Metropolis-coupled MCMC runs and checks for convergence on the fly after discarding the first 25% of tree samples as burn-in. Analyses were run with three heated chains and one cold chain for 5,000,000 generations. Chains were sampled every 1000 generations. Chains started with random trees and the following priors: all trees equally probable; Dirichlet probability density for base frequencies (1,1,1,1); Dirichlet probability density for rate matrix (all substitution types set at 1.0); uniform distribution for the shape (α) of the gamma distribution (0.1, 50.0); uniform distribution for proportion of invariant sites (0.0, 1.0); exponential (10.0) distribution for branch lengths. Bayesian analyses provided strong support for the four-clade classification (Afrotheria, Xenarthra, Laurasiatheria, Euarchontoglires) as well as for an association of Laurasiatheria and Euarchontoglires in the superordinal group Boreoeutheria (Fig. 1).

Pseudoextinction Analyses

Forty-four of the 57 taxa included in our molecular data set are also represented in the morphological data set of Asher et al. (2003; data set available at http://people.pwf.cam.ac.uk/rja58/; also see online Supplemental Material at www.systematicbiology.org), which is the largest currently available morphological data set to include representatives of all extant orders of placental mammals. The 44 taxa represented in our molecular data set that overlap with Asher et al.'s (2003) morphological data set are denoted in Figure 1 with asterisks. Asher et al.'s (2003) data set includes 185 osteological characters and 11 soft-tissue characters. Molecular and morphological data from Asher et al. (2003) were concatenated into a mixed data set for these 44 taxa. Our combined data set included two marsupial outgroups and representatives of all 18 placental orders.

DNA molecules and soft-tissue morphological characters are not routinely preserved for fossil mammals. In phylogenetic analyses with the 44-taxon data set, taxa representing each placental order (one order at a time) were treated as if the order was extinct by coding both molecular and 11 soft-tissue morphological characters as missing. The remaining 185 morphological characters are osteological and were retained for phylogenetic analyses with the pseudoextinct order. We assessed whether each pseudoextinct order was recovered as monophyletic on the most parsimonious tree(s) and in Bayesian analyses. We also assessed whether each pseudoextinct order remained in the same superordinal group (Afrotheria, Xenarthra, Laurasiatheria, Euarchontoglires) as in Figure 1 or moved elsewhere on the tree. In cases where individual pseudoextinct orders moved elsewhere on the tree we performed additional pseudoextinction analyses using taxonomic subsets of these orders. The superordinal groups Afrotheria, Euarchontoglires, and Laurasiatheria were also treated as pseudoextinct. Bayesian analyses with the 44-taxon mixed data set were performed with MrBayes 3.1.1 (Ronquist and Huelsenbeck, 2003; Ronquist et al., 2005). Analyses were run for a minimum of three million generations with four chains. Longer runs with up to eight chains were performed if necessary to achieve convergence. We used a threshold of 0.05 for the average standard deviation of split frequencies as a diagnostic for convergence. A lower threshold value (i.e., 0.01) is preferable, but achieving this level of convergence proved impractical for some of our analyses with pseudoextinct taxa. We note, however, that more than 75% of our analyses exceeded the desired 0.01

TABLE 2. Results of parsimony and Bayesian analyses when extant placental superordinal groups are treated as pseudoextinct. Parsimony results are based on the most parsimonious tree(s) for each pseudoextinction analysis (1000 randomized taxon input orders; TBR branch-swapping); parsimony bootstrap percentages are based on 500 pseudoreplications with 10 randomized taxon input orders per pseudoreplicate. Posterior probabilities for Bayesian analyses are average values based on two independent runs. Bayesian analyses were run until the average standard deviation of split frequencies was ≤ 0.05. Molecular and soft morphological characters were both scored as missing for pseudoextinct taxa.

	Parsimony results		Bayesian results	
Superordinal group	Monophyletic?	Location on shortest tree(s) when pseudoextinct + bootstrap percentages	Monophyletic?	Location on tree when pseudoextinct
Afrotheria	No	Tubulidentata joins Xenarthra (71%); paenungulates and Afrosoricida join Laurasiatheria (35%) where paenungulates are sister to perissodactyls (25%), tenrecid is sister to soricid (29%), and chrysochlorid is sister to talpid (64%); macroscelideans move to Euarchontoglires and are sister to Glires[a] (31%)	No	Tubulidentata moves to Xenarthra (0.67) and is the sister to Dasypodidae (0.53); all other afrotherians join Laurasiatheria (0.80) where macroscelideans are sister to Erinaceidae (0.87), Afrosoricida joins Talpidae (0.73), and paenungulates form a clade with perissodactyls (0.88)
Euarchontoglires	No	Glires moves to Afrotheria (16%) and is sister to Macroscelidea (2 trees; 24%) or Afrosoricida + Macroscelidea (1 tree; 12%); Dermoptera, Scandentia, and Primates join Laurasiatheria (14%) with Dermoptera sister to Chiroptera (43%) and reconstitution of Archonta (7%)	No	Glires moves to Afrotheria (0.83) and joins Macroscelidea (0.82); Euarchonta remains monophyletic (0.84) and moves to the base of Placentalia where there is a split between Euarchonta and all other placentals (0.60)
Laurasiatheria	No	Pholidota joins Xenarthra (48%) as sister to Myrmecophagidae (74%); all other taxa move to Afrotheria (10%) where perissodactyls + artiodactyls join paenungulates (75%); Erinaceidae sister to Macroscelidea (63%); talpid joins chrysochlorid (59%); soricid sister to talpid + chrysochlorid + tenrecid (40%); chiropterans + cetaceans + carnivores join eulipotyphlans + afrosoricidans + macroscelideans (5%)[b]	No	Pholidota joins Xenarthra (0.95) as sister to Myrmecophagidae (0.97); Chiroptera joins Euarchontoglires (0.75) as sister to Dermoptera (0.65); eulipotyphlans, artiodactyls, and perissodactyls join Afrotheria where Erinaceidae is sister to Macroscelidea (0.93), soricid joins tenrecid (0.50), talpid joins chrysochlorid + soricid + tenrecid (0.93), and artiodactyls + perissodactyls form a monophyletic clade (0.95) that joins Hyracoidea (0.99) inside of a paraphyletic Paenungulata; carnivores and cetaceans are excluded from a clade that contains all other placental mammals (0.78)

[a] On bootstrap trees the most common position for macroscelideans is sister to lipotyphlans (45%).
[b] On bootstrap trees the most common position for chiropterans is sister to Dermoptera (42%).

threshold. Burnin was set to include the first 25% of tree samples. We used the Lewis (2001) model for morphological characters. Settings for molecular partitions were as described in Figure 1 legend. Parsimony analyses with the mixed data set were performed with PAUP 4.0b10 (Swofford, 2003).

Analyses with the 44-taxon data set that treated entire molecular or morphological superordinal groups as pseudoextinct and examined the resulting deployment of taxa over the remaining tree are summarized in Table 2. When all afrotherian orders were pseudoextinct, they dispersed to Euarchontoglires (Macroscelidea),

TABLE 3. Results of parsimony and Bayesian analyses when placental orders are pseudoextinct. Parsimony results are based on the most parsimonious tree(s) for each pseudoextinction analysis (1000 randomized taxon input orders; TBR branch-swapping); parsimony bootstrap percentages are based on 500 pseudoreplications with 10 randomized taxon input orders per pseudoreplicate. Posterior probabilities for Bayesian analyses are average values based on two independent runs. Bayesian analyses were run until the average standard deviation of split frequencies was ≤ 0.05. Molecular and soft morphological characters were both scored as missing for pseudoextinct taxa.

	Parsimony results		Bayesian results	
Order	Monophyletic?	Location on shortest tree(s) when pseudoextinct + bootstrap percentages	Monophyletic?	Location on tree when pseudoextinct
Afrosoricida	No	Moves to Laurasiatheria (96%) and joins Eulipotyphla (92%); chrysochlorid sister to talpid (69%); tenrec sister to soricid (1 tree; 28%) or talpid + chrysochlorid (1 tree; 32%)	No	Moves to Laurasiatheria (1.00); chrysochlorid joins talpid (0.95) and tenrec joins chrysochlorid + talpid (0.80)
Carnivora	Yes (86%)	Moves to Afrotheria (42%) where carnivores are sister to paenungulates (31%)	Yes (1.00)	Moves to base of Placentalia [basal split between carnivores (1.00) and all other placentals (0.99)]
Cetartiodactyla	No	Cetaceans move to Xenarthra (16%); artiodactyls remain in Laurasiatheria (39%) and form clade with perissodactyls (72%)	No	Remains in Laurasiatheria (0.67); artiodactyls are monophyletic (0.96) and join perissodactyls (0.93); cetaceans are a monophyletic (1.00) sister to chiropterans + artiodactyls + perissodactyls + pholidotans + carnivores (0.52 for cetaceans + sister taxon)
Chiroptera	Yes (100%)	Moves to Euarchontoglires (2 trees; 39%) where chiropterans join Dermoptera (33%) or remain in Laurasiatheria (1 tree; 37%) where they are sister to Erinaceidae (17%)	Yes (1.00)	Chiropterans join Euarchontoglires (0.47) as a monophyletic sister taxon to this superordinal group
Dermoptera	NA	Moves to Laurasiatheria (51%) where Dermoptera is sister to Pholidota[a] (16%)	NA	Remains in Euarchontoglires (0.83) and is the sister to Scandentia (0.82)
Eulipotyphla	No	Moves to Afrotheria (78%) where Erinaceidae joins Macroscelidea (68%), talpid joins chrysochlorid (61%), and soricid is sister to talpid + chrysochlorid + tenrecid (47%)	No	Moves to Afrotheria (1.00) with Erinaceidae as sister to Macroscelidea (1.00), talpid as sister to chrysochlorid (0.63), and soricid as sister to tenrecid (0.78)
Hyracoidea	NA	Moves to Laurasiatheria (95%) and is the sister to Perissodactyla (62%)	NA	Moves to Laurasiatheria (1.00) and is the sister to Perissodactyla (1.00)
Lagomorpha	Yes (92%)	Remains in Euarchontoglires (68%) and Glires (73%)	Yes (1.00)	Remains in Euarchontoglires (1.00) but joins murid rodents (0.61)
Macroscelidea	Yes (100%)	Moves to Laurasiatheria (60%) and is sister to Erinaceidae (51%)	Yes (1.00)	Moves to Laurasiatheria (0.99) and is sister to Erinaceidae (0.88)
Perissodactyla	Yes (60%)	Moves to Afrotheria (50%) where perissodactyls are sister to Proboscidea + Sirenia (22%)	Yes (1.00)	Moves to Afrotheria (0.99) as sister to Hyracoidea (0.99)
Pholidota	NA	Moves to Xenarthra (79%) and is sister to Myrmecophagidae (78%)	NA	Moves to Xenarthra (1.00) and is sister to Myrmecophagidae (1.00)
Primates	Yes (42%)	Moves to Afrotheria (2 trees) or remains in Euarchontoglires (5 trees)	Yes (1.00)	Remains in Euarchontoglires (1.00) as sister to Scandentia (1.00)
Proboscidea	NA	Remains in Afrotheria (71%) but is sister to Sirenia (79%)	NA	Remains in Afrotheria (0.98) but is sister to Sirenia (0.98)
Rodentia	Yes (24%) on 2 of 4 minimum length trees	Remains in Euarchontoglires (80%) where Glires is monophyletic 92%)	Yes (1.00)	Remains in Euarchontoglires (1.00) as sister to Lagomorpha (1.00)
Scandentia	NA	Remains in Euarchontoglires (64%) and is sister to primates (2 trees; 31%), Glires (1 tree; 12%), or Glires + Dermoptera (1 tree; 15%)	NA	Remains in Euarchontoglires (0.97) as sister to Dermoptera (0.59)

(Continued on next page)

TABLE 3. Results of parsimony and Bayesian analyses when placental orders are pseudoextinct. Parsimony results are based on the most parsimonious tree(s) for each pseudoextinction analysis (1000 randomized taxon input orders; TBR branch-swapping); parsimony bootstrap percentages are based on 500 pseudoreplications with 10 randomized taxon input orders per pseudoreplicate. Posterior probabilities for Bayesian analyses are average values based on two independent runs. Bayesian analyses were run until the average standard deviation of split frequencies was ≤ 0.05. Molecular and soft morphological characters were both scored as missing for pseudoextinct taxa. *(Continued)*

	Parsimony results		Bayesian results	
Order	Monophyletic?	Location on shortest tree(s) when pseudoextinct + bootstrap percentages	Monophyletic?	Location on tree when pseudoextinct
Sirenia	NA	Remains in Afrotheria (58%) but is sister to Proboscidea (65%)	NA	Remains in Afrotheria (1.00) but is sister to Proboscidea (1.00)
Tubulidentata	NA	Moves to Xenarthra (60%) where Tubulidentata is sister to Dasypodidae (1 tree; 31%) or is sister to Xenarthra (1 tree; 29%)	NA	Moves to Xenarthra (0.81) and joins Dasypodidae (0.53)
Xenarthra	No	Dasypodid moves to Afrotheria (64%) and is sister to Tubulidentata (43%); myrmecophagid moves to Laurasiatheria (84%) and is sister to Pholidota (91%)	No	Dasypodid moves to Afrotheria (1.00) and is sister to Tubulidentata (1.00); myrmecophagid moves to Laurasiatheria (1.00) and is sister to Pholitoda (1.00)

[a] Dermoptera is sister to Laurasiatheria rather than Pholitoda on the bootstrap tree.

Laurasiatheria (Afrosoricida, Proboscidea, Sirenia, Hyracoidea), and Xenarthra (Tubulidentata). With the exception of pangolins, which joined Xenarthra, all of the pseudoextinct laurasiatherian orders moved to Afrotheria. Among the five orders in Euarchontoglires, rodents and lagomorphs joined elephant shrews in Afrotheria, whereas tree shrews, primates, and flying lemurs joined bats to reconstitute the morphological clade Archonta (parsimony) or moved to the base of the tree where they were a monophyletic sister-group to all other placentals (Bayesian).

The results of analyses that treated individual orders as pseudoextinct are summarized in Table 3. In both parsimony and Bayesian analyses, 4 of 18 orders (Afrosoricida, Eulipotyphla, Cetartiodactyla, Xenarthra) were not recovered as monophyletic when they were coded as pseudoextinct. Rodents were paraphyletic or monophyletic in parsimony analyses and monophyletic in Bayesian analyses (Table 3). However, only 5 of 18 orders (rodents, lagomorphs, scandentians, sirenians, proboscideans) consistently remained in the molecularly defined superordinal group, whereas the other 13 orders were deployed to one or more of the other superordinal groups on some or all of the most parsimonious trees and/or in the Bayesian analyses (Table 3). Among afrotherian orders, Afrosoricida, Hyracoidea, and Macroscelidea moved to Laurasiatheria and aardvark nested within Xenarthra as the sister-taxon to the armadillo. Among the orders that moved to Laurasiatheria, afrosoricidans joined eulipotyphlan insectivores (shrews, moles, hedgehogs), macroscelideans joined Erinaceidae (hedgehogs), and hyracoids became the sister taxon to Perissodactyla. When laurasiatherian orders were treated as pseudoextinct, pholidotans joined Xenarthra; carnivores, perissodactyls, and eulipotyphlans moved to Afrotheria; chiropterans either stayed in Laurasiatheria or moved to Euarchontoglires; and cetartiodactyls split into a cetacean group that either joined Xenarthra (parsimony) or remained in Laurasiatheria (Bayesian), and an artiodactyl assemblage (paraphyletic in parsimony analyses and monophyletic in Bayesian analyses) that remained within Laurasiatheria. When xenarthrans were pseudoextinct, the myrmecophagid (anteater) moved to Laurasiatheria (sister-taxon to Pholidota), whereas the dasypodid (armadillo) moved to Afrotheria (sister-taxon to Tubulidentata). Placental orders in Euarchontoglires were the most stable when treated as pseudoextinct and the only movements to other superordinal groups were Dermoptera joining Pholidota in Laurasiatheria and primates moving to Afrotheria (Table 3).

Table 4 shows the results of pseudoextinction analyses for taxonomic subsets of topologically unstable placental orders (i.e., orders that moved to different superordinal groups in Table 3). Most taxa remained in the same superordinal group that they reside in on the molecular tree shown in Figure 1. Macroscelideans, which moved to Laurasiatheria as the sister to Erinaceidae when both species were coded as pseudoextinct, remained in Afrotheria when only one species was pseudoextinct. Similarly, carnivores remained in Laurasiatheria when only one of two carnivore taxa was pseudoextinct. However, there were six taxa (Myrmecophagidae, Dasypodidae, Chrysochloridae, Tenrecidae, Talpidae, Erinaceidae) that moved to a different superordinal group when they alone were pseudoextinct in both parsimony and Bayesian analyses (Table 4).

DATA CONGRUENCE

Current morphological data sets are small relative to molecular data sets and it remains possible that larger morphological data sets will overcome problems that beset current data sets if these problems are statistical in nature and are tied to small sample size. Alternatively, there may be fundamental differences between

TABLE 4. Results of parsimony and Bayesian analyses when subsets of placental orders are pseudoextinct. Parsimony results are based on the most parsimonious tree(s) for each pseudoextinction analysis (1000 randomized taxon input orders; TBR branch-swapping); parsimony bootstrap percentages are based on 500 pseudoreplications with 10 randomized taxon input orders per pseudoreplicate. Posterior probabilities for Bayesian analyses are average values based on two independent runs. Bayesian analyses were run until the average standard deviation of split frequencies was ≤ 0.05. Molecular and soft morphological characters were both scored as missing for pseudoextinct taxa.

	Parsimony results		Bayesian results	
Taxon	Monophyletic?	Location on shortest tree(s) when pseudoextinct + bootstrap percentages	Monophyletic?	Location on tree when pseudoextinct
Myrmecophagid	N/A	Moves to Laurasiatheria (92%) and is sister to Pholidota (90%)	N/A	Moves to Laurasiatheria (1.00) and is sister to Pholidota (1.00)
Dasypodid	N/A	Moves to Afrotheria (75%) and is sister to Tubulidentata (46%)	N/A	Moves to Afrotheria (0.78) and is sister to Tubulidentata (0.77)
Chrysochlorid	N/A	Moves to Laurasiatheria (84%) and is sister to talpid (62%)	N/A	Moves to Laurasiatheria (84%) and is sister to talpid (62%)
Tenrecid	N/A	Moves to Laurasiatheria (96%) and is sister to talpid (2 trees; 33%), erinaceid (1 tree; 20%), or soricid (1 tree; 20%)	N/A	Moves to Laurasiatheria (1.00) and is sister to soricid (0.83)
Macroscelides	Yes (100%)	Remains in Afrotheria (86%) and is sister to *Elephantulus* (100%)	Yes (1.00)	Remains in Afrotheria (1.00) and is sister to *Elephantulus* (1.00)
Elephantulus	Yes (100%)	Remains in Afrotheria (72%) and is sister to *Macroscelides* (100%)	Yes (1.00)	Remains in Afrotheria (1.00) and is sister to *Macroscelides* (1.00)
Soricid[a]	No	Remains in Laurasiatheria (54%) and is sister to talpid (40%)	N/A	Moves to Afrotheria (0.71) and is sister to tenrecid (0.70)*
Talpidae	No	Moves to Afrotheria (91%) and is sister to Afrosoricida (76%)	N/A	Moves to Afrotheria (0.98) and is sister to chrysochlorid (0.84)
Erinaceidae	No	Moves to Afrotheria (75%) and is sister to Macroscelidea (57%)	N/A	Moves to Afrotheria (0.85) and is sister to Macroscelidea (0.85)
Yangochiroptera	Yes (83%)	Remains in Laurasiatheria (100%) and is sister to Pteropodidae (100%)	Yes (1.00)	Remains in Laurasiatheria (1.00) and is sister to *Rousettus* (0.80)
Pteropodidae	Yes (87%)	Remains in Laurasiatheria (100%) and is sister to Yangochiroptera (100%)	Yes (1.00)	Remains in Laurasiatheria (100%) and is sister to Phyllostomidae (100%)
Feliformia	N/A	Remains in Laurasiatheria (99%) and is sister to Caniformia (81%)	N/A	Remains in Laurasiatheria (1.00) and is sister to Caniformia (1.00)
Caniformia	N/A	Remains in Laurasiatheria (99%) and is sister to Feliformia (78%)	N/A	Remains in Laurasiatheria (1.00) and is sister to Feliformia (1.00)
Ceratomorpha	No	Remains in Laurasiatheria (84%); Ceratomorpha is paraphyletic to Hippomorpha with tapirid as sister to Hippomorpha (2 trees; 44%) or rhinocerotid as sister to Hippomorpha (1 tree; 33%)	Yes (0.96)	Remains in Laurasiatheria (0.96) and is sister to Hippomorpha (0.96)
Hippomorpha	N/A	Remains in Laurasiatheria (99%) and is sister to tapirid (2 trees; 45%) or rhinocerotid (1 tree; 32%)	N/A	Remains in Laurasiatheria (0.99) and is sister to tapirid (0.93)
Cetacea	Yes (100%)	Moves to Xenarthra (22%) and is sister to myrmecophagid (2 trees; 22%) or remains in Laurasiatheria (62%) and is sister to Chiroptera (1 tree; 20%)	Yes (1.00)	Remains in Laurasiatheria (0.87) in a clade with carnivores and Pholidota (0.63)
Lama	N/A	Remains in Laurasiatheria (100%) and is sister to Ruminantia (66%)	N/A	Remains in Laurasiatheria (1.00) and is sister to Ruminantia (1.00)
Sus	N/A	Remains in Laurasiatheria (95%) and is sister to a clade that contains other cetartiodactyls + perissodactyls (40%)	N/A	Remains in Laurasiatheria (0.93) and is sister to *Lama* (0.80)
Ruminantia	N/A	Remains in Laurasiatheria (99%) and is sister to *Lama* (73%)	N/A	Remains in Laurasiatheria (1.00) and is sister to *Lama* (1.00)
Hippopotamus	N/A	Remains in Laurasiatheria (92%) and is sister to Ruminantia (24%)	N/A	Remains in Laurasiatheria (0.92) and is sister to *Lama* (0.69)
Homo	N/A	Remains in Euarchontoglires (73%) and is sister to *Tarsius* (65%)	N/A	Remains in Euarchontoglires (1.00) and is sister to *Tarsius* (0.96)
Strepsirrhine	N/A	Remains in Euarchontoglires (75%) and is sister to *Tarsius* + *Homo* (67%)	N/A	Remains in Euarchontoglires (1.00) and is sister to *Tarsius* (0.91)
Tarsius	N/A	Remains in Euarchontoglires (78%) and is sister to *Homo* (68%)	N/A	Remains in Euarchontoglires (1.00) and is sister to *Homo* (0.59)

[a] The average standard deviation of split frequencies remained at 0.10 after Bayesian analyses were run for 80 million generations with eight chains (one cold, seven hot) for each analysis.

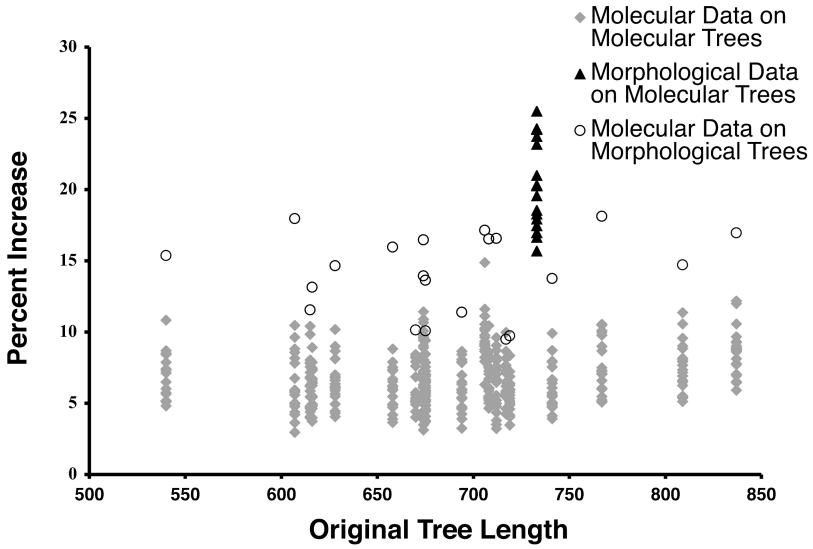

FIGURE 2. Plot of minimum tree length for 22 different data partitions (21 molecular, one morphological; *x*-axis) versus the increase in tree length when each data partition is mapped on to the best tree(s) for each of the other data partitions (*y*-axis). In cases where a partition was mapped on to more than one equally most parsimonious tree for another data partition, we plotted the midpoint value for the percentage increase in tree length. The 21 molecular partitions (P) arbitrarily followed the sequential gene order in our concatenated molecular data set, irrespective of gene boundaries, and included characters from the following gene segments: P1 (*ADRA2B*); P2 (*ADRA2B*); P3 (*ADRA2B, ADORA3, ADRB2*); P4 (*ADRB2*); P5 (*ADRB2, APOB*); P6 (*APOB, APP*); P7 (*APP, ATP7A*); P8 (*ATP7A*); P9 (*BDNF, BRCA1*); P10 (*BRCA1*); P11 (*BRCA1*); P12 (*BRCA1*); P13 (*BRCA1*); P14 (*BRCA1*); P15 (*BRCA1*); P16 (*BRCA1*); P17 (*BRCA1, CREM, EDG1*); P18 (*EDG1*); P19 (*EDG, PLCB4, VWF*); P20 (*VWF)*; P21 (*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 32 taxa included in partition congruence analyses were tamandua, armadillo, hedgehog, mole, shrew, tenrec, golden mole, sirenian, hyrax, elephant, elephant shrew, aardvark, mouse, hystricid, North American porcupine, pika, flying lemur, tree shrew, strepsirrhine, human, flying fox, rousette fruit bat, whale, dolphin, hippo, ruminant, pig, horse, ceratomorph, cat, caniform, pangolin.

molecular and morphological data. We examined congruence within the molecular data set and between the molecular and morphological data after partitioning the molecular data into 21 partitions that contained the same number of phylogenetically informative characters as the morphological data set. First, we identified 12 gene segments (*ADORA3, ADRA2B, ADRB2, APOB, APP, ATP7A, BDNF, BRCA1, CREM, EDG1, PLCB4, VWF*) with taxonomic overlap that comprised all placental orders as well as five lineages of afrosoricidan and eulipotyphlan insectivores. Taxonomic overlap for these 12 segments included 32 placental taxa. This resulted in an alignment that included 9179 nucleotide positions, of which 3818 were phylogenetically informative. Next, the 3818 phylogenetically informative characters were used to generate 21 different molecular partitions, each of which contained 175 informative characters and matched the size of the morphological data set (175 informative characters) of Asher et al. (2003) for the same 32 placental taxa. Data for each partition were then mapped on to the most parsimonious tree(s) for each of the other data partitions and the percentage increase in the number of steps was calculated. Incongruence was indexed as the percentage increase in the number of steps for each partition relative to the number of steps on the most parsimonious tree(s) for each data partition.

The morphological data consistently emerged as the most incongruent data partition, even though molecular data were segregated into partitions that contained the same number of informative characters (175) as the morphological data set (Fig. 2). The increase in tree length associated with mapping molecular data for one partition on to the best tree(s) for another molecular partition ranged from 2.9% to 14.9% (mean = 6.7%, gray diamonds in Fig. 2) for 420 comparisons; mapping molecular data partitions on to the morphological tree (21 comparisons) resulted in increases in tree length that ranged from 9.5% to 18.1% (mean = 14.2%, open circles in Fig. 2); mapping morphological data on to trees derived from molecular data partitions (21 comparisons) resulted in increases in tree length that ranged from 15.7% to 25.5% (mean = 19.9%, black triangles in Fig. 2).

With one exception, these results show that for every molecular data partition, the alternate tree from morphology was worse (longer) than any tree from another molecular data partition (containing the same number of parsimony informative characters). In addition, all of the mappings of morphological data on to molecular trees were worse than any molecular tree with different molecular data. Thus, there is more conflict between the morphological data and any molecular data partition than between all possible pairs of molecular data partitions. These findings also suggest that current molecular and morphological data for placental mammal orders are not readily miscible and that the disagreement between morphological and molecular data

is not simply a sample size (i.e., number of characters) problem.

Osteology remains the only source of data for most fossils, but methods for coding and analyzing such data are ineffective at discriminating between homology and homoplasy at the level of placental interordinal relationships. The previously mentioned problem of correlated character evolution (Chase et al., 2002; Kangas et al., 2004; Carrier et al., 2005) may contribute to incongruence between molecular and morphological data. Given that character matrices used by Asher et al. (2003), Zack et al. (2005), and Tabuce et al. (2007) include precisely the kinds of characters that might be expected to be convergent in mammals with similar diet and/or locomotion, the potential problem of correlated character evolution must be addressed.

CONCLUSIONS

We tested the ability of the most comprehensive morphological data set available to recapitulate the accepted phylogeny for living placental mammals, in the absence of molecular data, as a way to evaluate the accuracy of cladistic studies of fossil specimens that rely exclusively on morphology. We show that as many as 72% of the living placental orders move to a different superordinal group when molecular data are missing and that four of the 18 orders are rendered poly- or paraphyletic. Superordinal groups are never recovered as monophyletic when treated as pseudoextinct. We also show that there is fundamental incongruence between molecular and morphological data at the level of placental interordinal relationships. These results suggest that morphological studies of eutherian interordinal relationships have failed to separate homology and homoplasy and have consistently been misled by the latter. Cases in point include the recovery of Insectivora when either Eulipotyphla or Afrosoricida are pseudoextinct, Volitantia when chiropterans are pseudoextinct, Altungulata when perissodactyls are pseudoextinct, and Edentata when pholidotans are pseudoextinct. Further, these homoplastic groups that are recovered based on morphology often mix placental orders that originated on Gondwana versus Laurasia. The inadequacies of even the most extensive published morphological data for reconstructing higher level placental mammal phylogeny become even more apparent when entire superordinal groups are treated as extinct. Each of the four major clades disappears when it is pseudoextinct and is transmogrified into a polyphyletic assemblage that is integrated elsewhere across the tree. Notably, the apposition of southern versus northern hemisphere clades is all but lost when molecular data are missing for Afrotheria, Euarchontoglires, or Laurasiatheria.

Can we trust morphological cladistic analyses that place extinct aspheliscines within or at the base of Afrotheria? This question is comparable to asking whether or not pseudoextinct taxa remain in their original superordinal group or move to a different superordinal group. Our results show numerous instances of clades composed of one or two species (e.g., Afrosoricida, Perissodactyla, Macroscelidea, Carnivora, Pholidota, Xenarthra, Dasypodidae, Myrmecophagidae, Chrysochloridae, Tenrecidae, Talpidae, Erinaceidae) that move to a different superordinal group when they are pseudoextinct. These results suggest that Zack et al.'s (2005) and Tabuce et al.'s (2007) placement of aspheliscines could be a similar artifact and underscore the difficulty of reconstructing mammalian phylogenetic history for ancient, extinct lineages that are only known for osteological characters. Another problem associated with real fossils, and one that we did not address in our pseudoextinction analyses, is that fossil skeletons are rarely complete and some or many osteological characters must be coded as missing.

We agree with Jenner (2004), Wiens (2004), and Smith and Turner (2005) that systematists should continue to collect data for morphology-based phylogenetic analyses. However, our findings suggest that new methods for coding and analyzing morphological characters should be explored (also see Wiens, 2004), at least for analyzing difficult phylogenetic problems such as placental interordinal relationships. Lewis (2001) suggested scoring all variable morphological characters (i.e., informative and autapomorphic) rather than just informative characters and developed a model for analyzing appropriately scored data that will diminish certain long-branch attraction problems. Another strategy is to score morphological characters for stem or early crown representatives of placental orders. The rationale is that these taxa will more closely approximate ancestral states for living orders and reveal instances of homoplasy in more derived taxa. In one of their morphological analyses, Zack et al. (2005) scored an early Eocene perissodactyl (*Hyracotherium*) rather than a living representative of this order. However, perissodactyls still clustered as the sister-taxon to Hyracoidea. Tabuce et al. (2007) included stem representatives for four afrotherian orders (Macroscelidea, Proboscidea, Sirenia, Hyracoidea) and recovered a clade that grouped these taxa together to the exclusion of an early perissodactyl (i.e., *Hyracotherium*) and an early artiodactyl (i.e., *Diacodexis*). However, *Hyracotherium* grouped closer to paenungulates than to *Diacodexis* in some analyses (fig. 3b) and only joined *Diacodexis* when the taxonomic matrix was manipulated to exclude the Eocene fossil genus *Anthracobune*. These results highlight the challenges that face morphologists studying early placental diversification. Finally, studies that integrate genetics, developmental biology, and morphology and that unravel the causal links between genetic variation and morphological features (Kangas et al., 2004; Kassai et al., 2005) remain critical for developing a proper foundation for the phylogenetic analysis of morphological characters.

Problems associated with reconstructing relationships among extant placental orders with morphology alone are largely overcome by combining morphological data with molecular data because the phylogenetic signal associated with the relatively large molecular data set overwhelms the conflicting signal from the smaller morphological data set. However, combined data only mask

the problem for extinct taxa where molecular data are still missing. Recovering a morphological tree that approximates the molecular/genomic tree of living placental mammals with its four superordinal clades is an important stepping-stone on the way to inferring phylogenetic relationships of extinct eutherian orders. Robust solutions for the phylogenetic placement of fossil taxa are a prerequisite for testing hypotheses concerning the biogeography and timing of the early diversification of mammals. In this context, one of the greatest challenges ahead for mammalian systematists is to tease apart homology and homoplasy in morphological characters.

ACKNOWLEDGMENTS

We thank Ron DeBry, Michael Woodburne, Michael Westerman, Carey Krajewski, John Gatesy, and two anonymous reviewers for reviewing earlier versions of the manuscript. This work was partially supported by the National Science Foundation (M.S.S. and W.J.M.).

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. Meng, J. R. Wible, M. C. McKenna, G. W. Rougier, D. Dashzeveg, and M. J. Novacek. 2005. Stem Lagomorpha and the antiquity of Glires. Science 307:1091–1094.

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

Carrier, D. R., K. Chase, and K. G. Lark. 2005. Genetics of canid skeletal variation: Size and shape of the pelvis. Genome Res. 15:1825–1830.

Chase, K., D. R. Carrier, F. R. Adler, T. Jarvik, E. A. Ostrander, T. D. Lorentzen, and K. G. Lark. 2002. Genetic basis for systems of skeletal quantitative traits: Principal component analysis of the canid skeleton. Proc. Natl. Acad. Sci. USA 99:9930–9935.

Eizirik, E., W. J. Murphy, and S. J. O'Brien. 2001. Molecular dating and biogeography of the early placental mammals. J. Hered. 92:212–219.

Gibson, A., V. Gowri-Shankar, P. G. Higgs, and M. Rattray. 2005. A comprehensive analysis of mammalian mitochondrial genome base composition and improved phylogenetic methods. Mol. Biol. Evol. 22:251–264.

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

Hasegawa, M., J. L. Thorne, and H. Kishino. 2003. Time scale of eutherian evolution estimated without assuming a constant rate of molecular evolution. Genes Genet. Syst. 78:267–283.

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.

Jenner, R. A. 2004. Accepting partnership by submission? Morphological phylogenetics in a molecular millennium. Syst. Biol. 53:333–342.

Kangas, A. T., A. R. Evans, I. Thesleff, and J. Jernvall. 2004. Nonindependence of mammalian dental characters. Nature 432:211–214.

Kassai, Y, P. Munne, Y. Hotta, E. Penttila, K. Kavanagh, N. Ohbayashi, S. Takada, I. Thesleff, J. Jernvall, and N. Itoh. 2005. Regulation of mammalian tooth cusp patterning by ectodin. Science 309:2067–2070.

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.

Lewis, P. O. 2001. A likelihood approach to estimating phylogeny from discrete morphological character data. Syst. Biol. 50:913–925.

Löytynoja, A., and M. C. Milinkovitch. 2001. SOAP, cleaning multiple alignments from unstable blocks. Bioinformatics 17:573–574.

Luo, Z.-X., R. L. Cifelli, and S. Kielan-Jaworoska. 2001. Dual origin of tribosphenic mammals. Nature 409:53–57.

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.

McKenna, M. C. 1975. Toward a phylogenetic classification of the Mammalia. Pages 21–46 *in* Phylogeny of the primates: A multidisciplinary approach (W. P. Luckett and F. S. Szalay, eds.). Plenum Press, New York and London.

Meng, J., Y. Hu, and C. Li. 2003. The osteology of Rhombomylus (Mammalia, Glires): Implications for phylogeny and evolution of Glires. Bull. Am. Mus. Nat. Hist. 275:1–247.

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.

Murphy, W. J., T. H. Pringle, T. A. Crider, M. S. Springer, and W. Miller. 2007. Using genomic data to unravel the root of the placental mammal phylogeny. Genome Res. 17:413–421.

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

Novacek, M. J. 1992. Mammalian phylogeny: Shaking the tree. Nature 356:121–125.

Novacek, M. 2001. Mammalian phylogeny: Genes and supertrees. Curr. Biol. 11:R573–R575.

Posada, D., and K. Crandall. 1998. ModelTest: Testing the model of DNA substitution. Bioinformatics 14:817–818.

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.

Ronquist, F., and J. P. Huelsenbeck. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572–1574.

Ronquist, F., J. P. Huelsenbeck, and P. van der Mark. 2005. MrBayes 3.1 manual.

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

Scotland, R. W., R. G. Olmstead, and J. R. Bennett. 2003. Phylogeny reconstruction: The role of morphology. Syst. Biol. 52:539–548.

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.

Smith, N. D., and A. H. Turner. 2005. Morphology's role in phylogeny reconstruction: Perspectives from paleontology. Syst. Biol. 54:166–173.

Springer, M. S., W. J. Murphy, E. Eizirik, and S. J. O'Brien. 2003. Placental mammal diversification and the Cretaceous-Tertiary boundary. Proc. Natl. Acad. Sci. USA 100:1056–1061.

Springer, M. S., W. J. Murphy, E. Eizirik, and S. J. O'Brien. 2005. Molecular evidence for major placental clades. Pages 37–49 *in* The rise of placental mammals (K. D. Rose and J. D. Archibald, eds.). The Johns Hopkins University Press, Baltimore.

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.

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

Tabuce, R., L. Marivaux, M. Adaci, M. Bensalah, J.-L. Hartenberger, M. Mahboudi, F. Mebrouk, P. Tafforeau, and J.-J. Jaeger. 2007. Early Tertiary mammals from North Africa reinforce the molecular Afrotheria clade. Proc. Roy. Soc. B 274:1159–1166.

Teeling, E. C., M. Scally, D. J. Kao, M. L. Romagnoli, M. S. Springer, and M. J. Stanhope. 2000. Molecular evidence regarding

the origin of echolocation and flight in bats. Nature 403:188–192.
van Rheede, T., T. Bastiaans, D. N. Boone, S. B. Hedges, W. W. de Jong, and O. Madsen. 2006. The platypus in its place: nuclear genes and indels confirm the sister group relation of monotremes and therians. Mol. Biol. Evol. 23:587–597.
Waddell, P. J., H. Kishino, and R. Ota. 2001. A phylogenetic foundation for comparative mammalian genomics. Genome Informatics 12:141–154.
Waddell, P. J., N. Okada, and M. Hasegawa. 1999. Towards resolving the interordinal relationships of placental mammals. Syst. Biol. 48:1–5.
Waddell, P. J., and S. Shelley. 2003. Evaluating placental inter-ordinal 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.
Westerman, M., A. Burk, H. M. Amrine-Madsen, G. J. Prideaux, J. A. Case, and M. S. Springer. 2002. Molecular evidence for the last survivor of an ancient kangaroo lineage. J. Mammal. Evol. 9:209–223.
Wiens, J. J. 2004. The role of morphological data in phylogeny reconstruction. Syst. Biol. 53:653–661.
Wolpoff, M. H. 1999. Paleoanthropology. McGraw-Hill, Boston, Massachusetts.
Zack, S.P., T. A. Penkrot, J. I. Bloch, and K. D. Rose. 2005. Affinities of "hyopsodontids" to elephant shrews and a Holarctic origin of Afrotheria. Nature 434:497–501.

First submitted 11 January 2006; reviews returned 6 June 2006; final acceptance 2 May 2007
Associate Editor: Ron DeBry

Syst. Biol. 56(4):684–689, 2007

ISSN: 1063-5157 print / 1076-836X online
DOI: 10.1080/10635150701494127

Diversity, Nomenclature, and Taxonomy of Protists

SINA M. ADL,[1] BRIAN S. LEANDER,[2] ALASTAIR G. B. SIMPSON,[1] JOHN M. ARCHIBALD,[3] O. ROGER. ANDERSON,[4] DAVID BASS,[5] SAMUEL. S. BOWSER,[6] GUY BRUGEROLLE,[7] MARK. A. FARMER,[8] SERGEY KARPOV,[9] MARTIN KOLISKO,[1] CHRISTOPHER E. LANE,[3] DEBORAH J. LODGE,[10] DAVID G. MANN,[11] RALF MEISTERFELD,[12] LEONEL MENDOZA,[13] ØJVIND MOESTRUP,[14] SHARON E. MOZLEY-STANDRIDGE,[15] ALEXEY V. SMIRNOV,[16] AND FREDERICK SPIEGEL[17]

[1]*Department of Biology, Dalhousie University, Halifax, NS B3H 4J1, Canada; E-mail: sadl@dal.ca (S.M.D.)*
[2]*Departments of Botany and Zoology, University of British Columbia, Vancouver, BC V6T 1Z4, Canada*
[3]*Department of Biochemistry, Dalhousie University, Halifax, NS B3H 4J1, Canada*
[4]*Lamont-Doherty Earth Observatory, Palisades, NY 10964, USA*
[5]*Department of Zoology, University of Oxford, Oxford, OX1 3PS, UK*
[6]*Wadsworth Center, New York State Department of Health, Albany, NY 12201, USA*
[7]*Biologie des Protistes, Université Blaise Pascal de Clermont-Ferrand, F63 177 Aubierre cedex, France*
[8]*Center for Ultrastructural Research, Department of Cellular Biology, University of Georgia, Athens, GA 30602, USA*
[9]*Biological Faculty, Herzen State Pedagogical University of Russia, St. Petersburg 191186, Russia*
[10]*Center for Forest Mycology Research, USDA Forest Service, Forest Products Laboratory, Luquillo, Puerto Rico*
[11]*Royal Botanic Garden, Edinburgh, EH3 5LR, UK*
[12]*Institute for Biology II, Unit of Cellular Neurobionics, RWTH Aachen University, D-52056 Aachen, Germany*
[13]*Medical Technology Program, Department of Microbiology and Molecular Genetics, Michigan State University, East Lansing, MI 48824-1030, USA*
[14]*Department of Phycology, Kobenhavns Universitet, Copenhagen, DK-1353, Denmark*
[15]*Biology, Middle Georgia College, Cochran, GA 31014, USA*
[16]*Department of Invertebrate Zoology, Saint Petersburg State University, 199034 St. Petersburg, Russia*
[17]*Department of Biological Sciences, University of Arkansas, Fayetteville, AR 72701, USA*

The new classification of protists from the International Society of Protistologists (Adl et al., 2005) could not apply both the International Code of Botanical Nomenclature and the International Code of Zoological Nomenclature because the two are incompatible. The classification designated one name for each clade where multiple names from different codes had previously existed, traced authorities, and provided a classification based on nameless ranks. Here we review important issues that remain to be resolved. Current rules governing validation of new species, from various codes of nomenclature, have become an impediment to naming of new protists. Standard requirements for protist species descriptions and type specimens need to be modernized to accommodate the rapid discovery of new species made possible by modern microscopic and molecular techniques. Although we agree with the criticisms of the botanical and zoological codes made by proponents of the Phylocode, we did not all agree that the current Phylocode is the solution, nor does it currently address species typification. Accordingly, new guidelines are needed to govern standards in protist species descriptions and classification.

Over the past 25 years, molecular phylogenetic studies have led to extensive modification of traditional classification schemes for eukaryotes. The most dramatic changes have occurred within protists, from which multicellular organisms evolved. The names of many protist groups and the genera they include have been changed so many times that the classification scheme is unclear, and it is difficult to determine which names apply. Two recent reviews have provided a modern phylogenetic perspective on the overall organization of eukaryote clades (Keeling et al., 2005; Simpson and Roger, 2004).

A necessary extension of this phylogenetic research was to establish a new classification that reflected the general consensus on the taxonomic names and their authorities (Adl et al., 2005). This classification scheme for protists breaks with tradition by not using either the International Code of Botanical Nomenclature (ICBN) (Greuter et al., 2000) or the International Code of Zoological Nomenclature (ICZN; International Commission on Zoological Nomenclature, 1999) regarding ranks, because neither of these codes are presently adequate for protists. The decision to do so was primarily practical. Where possible, well-known names referring to recognized monophyletic groups were retained. Although it did not try to follow the Phylocode, groups of named lineages were defined by apomorphies (derived characters) as much as possible, but node-based and stem-based definitions were used as necessary, even though they were not identified as such in the final presentation. In this classification, name endings that conveyed hierarchical information in a traditional code (e.g., -idae, -inae, -ales, -aceae) were retained to avoid unnecessary name changes but are not intended to convey hierarchical information. We believe this scheme to be more utilitarian as it recognizes one name for each clade where multiple names for the same clade were used previously. Furthermore, the classification is intended to facilitate future modification in light of improved phylogenetic information, without requiring a cascade of name changes. Further changes to the classification will no doubt be necessary given that our knowledge of some groups and our geographical sampling are still far from complete. Several critical issues remain to be resolved and we must continue to work towards a practical consensus.

Diversity of Protists

Adl et al. (2005) defined protists as eukaryotic organisms with unicellular, colonial, filamentous, or parenchymatous organization that lack vegetative tissue differentiation, except for reproduction. Metazoa Haeckel 1874, Plantae Haeckel 1866, and some Phaeophyceae Hansgirg 1886 are recognized as being truly multicellular. The current number of described protist species, including fungi, is widely acknowledged to be a fraction of the total diversity in nature (Table 1; May, 1988; Corliss, 2002). Many geographic regions have not been sampled at all and most regions and habitats are insufficiently sampled. The rate of discovery of new species from environmental samples remains high. Indeed, most soil, freshwater, or marine samples collected contain a multitude of undescribed species (Foissner, 1999, 2006; Slapeta et al., 2005) that are found through microscopy or environmental DNA samples. Owing to insufficient environmental sampling and reisolation, the geographical distribution of most species remains unknown. A metadata statistical analysis of species richness indicated that unicellular organisms showed high relative local species richness, which is consistent with most species being locally rare (Hillebrand et al., 2001). Species composition for protists was statistically less similar between samples with geographical distance, suggesting a regionally restricted distribution for some or many species, likely due to limitations to protist dispersal over long distances. Interpretation of these results is complicated, however, because species identification is typically based

TABLE 1. Approximate number of described species and estimated total number of species in each group.*

Group name	Number of known extant species	Potential number of species
Amoebozoa		
Lobose, naked	180	600
Arcellinida	1100	10^3 to 10^4
Myxogastria	>900	1200 to 1500
Dictyostelia	>100	300
Protostelia	36	150
Eumycetozoa	655	10,000
Other Ameobozoa	35	50
Opisthokonta		
Fungi (excluding Zygomycota)	335,000	$n \times 10^6$
Zygomycota	70,000	1.5×10^6
Chytridiomycetes	1000	<2000
Microsporidia	1200	10,000
Mesomycetozoa	47	$n \times 10^3$
Choanomonada	120	300
Rhizaria		
Cercozoa	<500	$n \times 10^3$
Haplosporidia	31	$n \times 10^2$
Foraminifera	>10,000	15,000
Acantharia	160	<200
Polycystinea	700 to 1000	1,500
Nucleohelea	160 to 180	200
Archaeplastida		
Glaucophyta and Rhodophyceae	4000 to 6000	20,000
Chloroplastida, excluding Charophyta	8000 to 10,000	$1–2 \times 10^5$
Charophyta, excluding Plantae	4300	5000
Chromalveolata		
Cryptophyceae	70	200
Haptophyta	350	<400
Phaeophyceae	1500 to 2000	2000
Actinophryidae	5	<10
Opalinata	400	500
Bicosoecida	72	100
Labyrinthulomycetes	40	<100
Hyphochytriales	25	$n \times 100$
Peronosporomycetes	676	10^3 to 10^4
Chrysophyceae	1000	2000
Dictyophyceae	15	30
Eustigmatophyceae	15	30
Pelagophyceae	12	20
Phaeothamniophyceae	25	40
Pinguiophyceae	5	20
Raphidophyceae	20	40
Synurophyceae	200	350
Xanthophyceae	600	800
Bacillariophyta	$1\text{-}2 \times 10^4$	2×10^5
Apicomplexa	6,000	$1.2–10 \times 10^6$
Dinozoa	2000	<3000
Ciliophora	3500	30,000
Excavata		
Fornicata	146	<200
Parabasalia	466	500
Preaxostyla	96	<120
Jakobida	10	200
Heterolobosea	80	200
Euglenozoa	1520	2000
Incertae sedis		
Eukaryota		
Apusomonadida	12	20

*Potential number of species were estimated by authors for each group based on number of unknown DNA sequences found in environmental samples.

on morphology, which often may not distinguish between species with similar or identical morphologies (Hillebrand et al., 2001; Adl and Gupta, 2006; Foissner, 2006).

Which Classification?

For historical reasons, protists traditionally fell under the jurisdiction of the ICBN if they were "algae" or "fungi" and under the jurisdiction of the ICZN if they were "protozoa." This system has been unraveling for several decades, as a number of groups were described in parallel by zoologists (ICZN) and by botanists (ICBN) each with distinct names (Corliss, 1995). To give just one example, Diatomea Dumortier 1821 and Bacillariophyta Haeckel 1878 both describe the same clade: the diatoms. The ranks within this group received a parallel series of names independently by zoologists and botanists to accommodate rank endings appropriate for each code (the so-called ambireginal classification). These unnecessary duplications introduced a double language throughout protist classification schemes that resulted in confusion.

The situation was exacerbated from the 1960s onward, as many genera were reclassified to accommodate new research and discoveries of new taxa. The traditional classification of protozoa and algae collapsed during the 1970s and 1980s as many groups were subsequently shuffled. Many ranks contained genera that were described under one code and other genera under the other code. More dramatically, it became evident through molecular phylogenies that fungi (governed by the ICBN) are a sister lineage of animals (governed by the ICZN), and novel protists discovered at the base of both of these clades were described following ICZN rules (Mendoza et al., 2002; James et al., 2006).

Lastly, the recognition of monophyletic groups based on modern phylogenetic concepts forces us to do things that are awkward with the traditional codes. For example, we would be forced to place classes within classes, and kingdoms within kingdoms, or invent many new ranks. These issues were elaborated fully elsewhere and will not be repeated here (Cantino, 2004; Pleijel and Rouse, 2003). Previous attempts at synthesis of a classification for eukaryotes, based on identifying successive evolutionary steps and providing a Linnaean name for each rank in the hierarchy, required numerous novel rank names (Cavalier-Smith, 1993) and never became widely used by protistologists. In part, this valiant effort was premature because most of the molecular phylogenetic information necessary became available subsequently. Several alternative classifications were proposed in this new light, with new competing names for the same groups of organisms (Cavalier-Smith, 1998; Patterson, 1999, 2002), with accompanying changes in ranks and authority as required by the ICBN or the ICZN. As a result, authors resorted to selecting one of several possible names for each group or, more commonly, used informal names without specifying an authority or a definition. This further added to the confusion. Without a memory of the history of changes associated with a taxon name, rank, and clade, identifying a group and its composition became very difficult for professionals, and almost impossible for those entering the field of protistology. There was simply no common rationale for deciding which name and which classification to use.

How Did We Get into This Mess?

The purpose of classification is to arrange biological diversity in such a way as to facilitate communication and accurate information retrieval. This system must operate within a phylogenetic context and must be able to accommodate modification while retaining name stability. This is a particularly onerous task as there are millions of phylogenetic entities at different hierarchical levels, with thousands more being discovered annually (May and Nee, 1995). The mess that arose in the classification of protists attests to the failure of the ICBN and ICZN to arrive at a mutually satisfactory accommodation, at accommodating changes in the classification, and providing unambiguous name stability in a modern evolutionary context.

The ICBN and ICZN were created based on preevolutionary principles laid out by Aristotle and Linnaeus, using a species binomial nomenclature of *Genus epithet*. Binomial nomenclature is responsible for much of the instability in the classification, as each time a taxon is moved, its generic name is changed (Cantino, 1998). This is not problematic for a small number of taxa, but the extent of change required to the classification was unforeseen. The fundamental division of life into plants versus animals appeared distinct and stable enough at the time, but protists blurred that distinction. The flexibility that would later be required of the traditional schemes, with the rapid expansion of protist taxa and extensive reclassification, simply could not be accommodated while retaining name stability.

Other problems with the Linnaean rank-based nomenclature have been the subject of many papers over the past 15 years (de Queiroz and Gauthier 1992, 1994; Cantino et al., 1997; de Queiroz, 1997; Kron, 1997; Hibbett and Donoghue, 1998; Pleijel and Rouse, 2003; Cantino, 2004). Some of the more problematic issues raised are that (1) rank dictates priority and synonymy under separate codes, instead of clades; (2) rank changes cause a cascade of name changes following even minor changes in phylogenetic hypotheses (shifting to a new rank changes both the name, and the authority of a group, even though the organisms it describes and the clade remains the same); (3) the codes are essentially silent on what is considered today to be the overriding concern in classification—the principle of common descent. It is permissible for the members of well supported clades to be separated into paraphyletic categories, even if doing so introduces misleading information about evolutionary relatedness; and (4) more emphasis is placed on who named or moved a group than the group and its name. Several other issues concern outdated approaches to describing species. For example, the requirement for Latin descriptions in the ICBN and what is acceptable as a type specimen

and holotype under both codes are impractical for protists and need modernizing, as discussed below. Unfortunately, the Phylocode is not much help on this point. Although it has attempted to introduce "tree-thinking" from molecular phylogenies into its rules of nomenclature, it deals only with rules governing clades and not with naming species.

TOWARD A SOLUTION

It has been argued that the traditional codes can be revised to accommodate some of the problems mentioned above, and that many of the identified problems are not serious (Barkley et al., 2004). That may be true for extant Animalia and Plantae, although some disagree (Cantino, 2004), but for protists that is simply not the case. An example of the many profound difficulties that can be encountered was recently provided for *Pneumocystis*, a pathogen that was traditionally treated as a protozoan under the ICZN but is now known to be a fungus and must be treated by the ICBN (Redhead et al., 2006). These difficulties are encountered with well-known isolates that exist in many laboratories. The problem is insurmountable with isolates that can be fully described but cannot be cultured or cryopreserved. To place the issue in perspective, imagine a situation where plant species descriptions would be acceptable only if the new specimen was domesticated enough to be cultivated! For example, a protist specimen that is digitally photographed and then used to obtain DNA for phylogenetic information will no longer physically exist to be deposited as a holotype. However, the resulting digital images, sequence data, and DNA sample—which are all necessary, sufficient, and more useful than a microscope slide for subsequent identification—continue to exist. At some point, so much modification is needed that the original code is no longer the same code but becomes something new (Cantino, 2004). It is impossible to be familiar with the diversity and classification of protists on the one hand and to claim that the ICBN and the ICZN have been stabilizing and accommodating on the other.

In our view, the following issues need to be addressed in the formulation of articles for a code that would be useful for the classification of protists and all eukaryotes. Below we highlight several approaches that would help in species typification by working towards standardized rules. We emphasize parameters that are useful to protist species delineation using a variety of biological and molecular approaches. Next, we discuss nomenclatural issues that would provide name stability.

Standardized Data Acquisition

Fundamentally, there is widespread acceptance that identification of protist species using light microscopy alone is no longer sufficient or adequate. Many well-studied morphotypes, including those with sufficient biogeographical sampling, are known to represent a variety of morphologically indistinguishable species (see Adl and Gupta, 2006; Foissner, 2006). These cryptic species can be distinguished by mating types if sexual, by feeding preferences, from DNA sequences, from excystment requirements, and from temperature or habitat optima. Descriptions based on microscopy and holotypes deposited in designated institutes as fixed slide preparations or photographs can help to describe a morphotype but fail to identify species. Accurate identification of morphotypes by microscopy depends on the array of morphotypes known to the microscopist. Identification errors are common because type specimens generally have limited accessibility, being kept inside "designated institutes" far away, or are not useful, and older published drawings and photographs are often of poor quality and insufficient on their own. We therefore recommend adoption of some combination of standardized requirements for microscopy that include using digital still-images of live specimens or digital video showing patterns of motility (in motile specimens), scanning or transmission electron micrographs, DNA sequence information, habitat and feeding preferences, and, where possible, a description of life cycle stages. For reference material, both the images and the sequence information must be freely available in electronic public databases.

Use of Molecular Data

DNA sequence information is commonly used both for understanding relatedness between clades as well as for identifying species. The most commonly used DNA sequences for phylogenetic reconstruction of eukaryotic groups, such as small subunit ribosomal DNA (18S rDNA), may underestimate intrageneric diversity in some clades but may be less conserved in others (Keeling et al., 2005; Simpson and Roger, 2004). Careful consideration is required to supplement the 18S rDNA data with sequence information from other genes, such as the mitochondrial *cox1* or from the ribosomal ITS regions. Choices about the number and identity of genes necessary for sufficient resolution may be different for different clades, and this needs to be established by experts. The sequence information should be compared with similar isolates from across the geographical range of morphotypes to obtain a sense of how much variation or diversity is represented by each morphotype. It is only with repeated isolation and comparison at a variety of locations that intrageneric protist diversity can be described adequately.

When comparing sequence differences, a recurring issue has been to ask, how much sequence divergence warrants a new species or genus? If a fixed amount of sequence difference is preset and applied uniformly, would many of the large primates, antelopes, or the brown kelps in the *Laminariales* converge to a small number of species and genera? Solving this question by setting a number, even with a more complicated formula, would be an arbitrary delineation of isolates into categories. We do not advocate a standardized and uniformly applied fixed amount of sequence difference to delineate species. Rather, the emphasis for species delineation would be placed on the combination of phylogenetic analysis of sequence data plus physiological

adaptations to a multidimensional niche space (i.e., an ecologically relevant parameter; Whittaker, 1972). These would include parameters described above in obtaining a standard set of information in species delineation. This shift in emphasis would require more species characterization than is currently done in describing isolates, a problem that will no doubt intensify as molecular data become cheaper and easier to acquire. A practical solution will have to be accommodated.

Name Stability

In seeking unambiguous and stable names in biological classifications, the PhyloCode (Cantino and de Queiroz, 2006) proposed a shift from Linnaean rank based nomenclatures toward naming nested clades with stem-, node-, or apomorphy-based definitions (de Queiroz and Gauthier, 1994; Cantino, 2004). This approach, as the authors argued, separates naming clades from assembling nested hierarchies, in contrast to rank-based nomenclature, which treats these steps as part of the same process. Therefore, clade composition is determined by the interaction of a clade definition with a phylogenetic hypothesis. The names do not necessarily change when the phylogenetic hypothesis (the classification) changes. This goes a long way toward providing name stability while accommodating changes in the classification. We do not suggest that the current PhyloCode provides solutions to all of the problems of protist classification, but it is clearly a step in the right direction. As we work toward adopting rules that will work for protists, the criticisms made of Linnaean rank-based nomenclatures, and its benefits, ought to be considered seriously by the biological community. The danger with the current situation is that out of necessity, protist species descriptions will occur outside of the guidelines established by the existing codes, and thus without standards, as researchers continue to ignore them as unworkable. The ambiguous situation has now expanded beyond protists, as descriptions of animal species are also occurring according to PhyloCode (even though it has not been formally implemented) or the ICZN, creating an ambireginal situation in Animalia (Hillis, 2007; Dubois, 2007), where clades are named and described according to two different and parallel set of rules.

Concluding Remarks

The new classification of eukaryotes reflects our current knowledge of protist evolution, has reintroduced some formality with group names and their authority, and provides a point of reference for protist systematics (Adl et al., 2005). Unresolved cases remain where relationships between clades are unclear (Adl et al., 2005; Keeling et al., 2005; Patterson, 1999). Some of the most undersampled groups include the most diverse eukaryotes, such as red algae, fungi, and apicomplexan parasites (Table 1). The rate of new species description should be limited only by how fast individual cells can be collected, photographed, and their genes sequenced, not by antiquated codes of nomenclature. Using high-throughput methods developed for genomic studies, potentially hundreds of new species could be discovered weekly from environmental samples, with accompanying phenotypic information from microscopy. Images and sequence information need to be publicly available electronically in searchable databases, such as in Discover Life (www.discoverlife.org) or Microscope (http://starcentral.mbl.edu/microscope/). Ultimately, all species need to be transferred into digital searchable catalogues that contain both DNA sequence information and images, as well as additional biological information. (A task of this magnitude can be accomplished with sufficient resources. For example, it was accomplished for much of the published scientific literature of the last century in two decades, despite many pessimists claiming it would take an unreasonably long time to do so.) There has to be a shift away from the emphasis on authorities and ranks toward clade name stability. Possibilities for dealing with the *genus epithet* binomial ambiguity can be handled simply, by combining the two (*genus.epithet* or *genus-epithet* or *genusepithet*) into a single unambiguous name (Cantino et al., 1999). The rules to standardize the process must be simple, few, and practical.

Perhaps the most serious consequence of not having had a classification with name stability for protists over the past decades has been the gradual omission of protists from biology textbooks (Adl, 2005). Without a classification with stable names to teach students, or to search the literature, the significance of the diversity of protists to the biology community has been diminishing. This has dire consequences to research funding in protistology, as long as these organisms are considered few and unimportant despite their key role in ecosystems and the evolution of life. This is unfortunate because protists cause many of the world's deadliest human diseases and include the most damaging crop pathogens such as *Phytophthora* that caused the Irish famine. The World Health Report (2004) ranked respiratory tract infections, diarrheal diseases, and malaria respectively as first, fourth, and sixth in number of deaths caused by communicable diseases, maternal and perinatal conditions, and nutritional deficiencies. Each of these categories contains a variety of pathogenic protists (Corliss, 2002). Yet, most people remain unaware of the diversity and complexity of protist cell biology, which is necessary to prevent crop damage, maintain livestock health, and to save human lives.

Acknowledgments

Several colleagues have contributed through responding to emails or in conversations at conferences and elsewhere. These include Robert A. Andersen, Louise Lewis, Richard M. McCourt, Robert Fensome, Timothy Y. James, Laura Katz, Georges Merinfeld, and Gary Saunders.

References

Adl, S. M. 2005. Microbiology, microbial diversity and microbial life. J. Euk. Microbiol. 52:546–548.

Adl, S. M., and V. V. S. R. Gupta. 2006. Protists in soil ecology and forest nutrient cycling. Can. J. Forest Res. 36:1805–1817.

Adl, M. S., A. G. B. Simpson, M. A. Farmer, R. A. Andersen, O. R. Anderson, J. Barta, S. S. Bowser, G. Brugerolle, R. A. Fensome, S. Fredericq, T. Y. James, S. Karpov, P. Kugrens, J. Krug, C. Lane, L. A. Lewis, J. Lodge, D. H. Lynn, D. G. Mann, R. M. McCourt, L. Mendoza, Ø. Moestrup, S. E. Mozley-Standridge, T. A. Nerad, C. A. Shearer, A. V. Smirnov, F. Spiegel, and F. J. R. Taylor. 2005. The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. J. Euk. Microbiol. 52:399–451.

Barkley, T. M., P. DePriest, V. Funk, R. W. Kiger, W. J. Kress, and G. Moore. 2004. Linnaean nomenclature in the 21st century: A report from a workshop on integrating traditional nomenclature and phylogenetic classification. Taxon 53:153–158.

Cantino, P. D. 1998. Binomials, hyphenated uninomials, and phylogenetic nomenclature. Taxon 47:425–429.

Cantino, P. D. 2004. Classifying species versus naming clades. Taxon 53:795–798.

Cantino, P. D., H. D. Bryant, K. de Queiroz, M. J. Donoghue, T. Eriksson, D. M. Hillis, M. S. Y. Lee. 1999. Species names in phylogenetic nomenclature. Syst. Biol. 48:790–807.

Cantino, P. D., and K. de Queiroz. 2006. PhyloCode: A phylogenetic code of biological nomenclature, version 3 [http://www.ohiou.edu/phylocode/].

Cantino, P. D., R. G. Olmstead, and S. J. Wagstaff. 1997. A comparison of phylogenetic nomenclature with the current system: A botanical case study. Syst. Biol. 46:313–331.

Cavalier-Smith, T. 1993. Kingdom Protozoa and its 18 phyla. Microbiol. Rev. 57:953–994.

Cavalier-Smith, T. 1998. A revised six-kingdom system of life. Biol. Rev. 73:203–266.

Corliss, J. O. 1995. The ambireginal protists and the codes of nomenclature: A brief review of the problem and of proposed solutions. Bull. Zool. Nomencl. 52:11–17.

Corliss, J. O. 2002. Biodiversity and biocomplexity of the protists and an overview of their significant roles in maintenance of our biosphere. Acta Protozool. 41:199–219.

de Queiroz, K. 1997. The Linnaean hierarchy and the evolutionization of taxonomy, with emphasis on the problem of nomenclature. Aliso 15:125–144.

de Queiroz, K., and J. Gauthier. 1992. Phylogenetic taxonomy. Annu. Rev. Ecol. Syst. 23: 449–480.

de Queiroz, K., and J. Gauthier. 1994. Toward a phylogenetic system of biological nomenclature. Trends Ecol. Evol. 9:27–31.

Discover Life. http://www.discoverlife.org/ (accessed November 28, 2006).

Dubois, A. 2007. Naming taxa from cladograms: A cautionary tale. Molec. Phylogen. Evol. 42:317–330.

Foissner, W. 1999. Protist diversity: Estimates of the near imponderable. Protist 150:363–368.

Foissner, W. 2006. Biogeography and dispersal of micro-organisms: A review emphasizing protists. Acta Protozool. 45:111–136.

Greuter, W., J. McNeill, F. R. Barrie, H. M. Burdet, V. Demoulin, T. S. Filgueiras, D. H. Nicholson, P. C. Silva, J. E. Skog, P. Trehane, N. J. Turland, and D. L. Hawksorth. 2004. International code of botanical nomenclature (St Louis Code). Königstein, Koeltz Scientific Books.

Hibbett, D. S., and M. J. Donoghue. 1998. Integrating phylogenetic analysis and classification in fungi. Mycologia 90:347–356.

Hillebrand, H., F. Watermann, R. Karez, U.G. Berninger. 2001. Difference in species richness patterns between unicellular and multicellular organisms. Oecologia 126:114–124.

Hillis, D. M. 2007. Constraints in naming parts of the tree of life. Mol. Phylogenet. Evol. 42:331–338.

International Commission on Zoological Nomenclature. 1999. International Code of Zoological Nomenclature, 4th edition. International Trust for Zoological Nomenclature, London.

James, T. Y., F. Kauff, C. L. Schoch, P. B. Matheny, V. Hofstetter, C. J. Cox, G. Celio, C. Gueidan, E. Fraker, J. Miadlikowska, H. T. Lumbsch, A. Rauhut, V. Reeb, A. E. Arnold, A. Amtoft, J. E. Stajich, K. Hosaka, G.-H. Sung, D. Johnson, B. O'Rourke, M. Crockett, M. Binder, J. M. Curtis, J. C. Slot, Z. Wang, A. W. Wilson, A. Schüszligler, J. E. Longcore, K. O'Donnell, S. Mozley-Standridge, D. Porter, P. M. Letcher, M. J. Powell , J. W. Taylor, M. M. White, G. W. Griffith, D. R. Davies, R. A. Humber, J. B. Morton, J. Sugiyama, A. Y. Rossman, J. D. Rogers, D. H. Pfister, D. Hewitt, K. Hansen, S. Hambleton, R. A. Shoemaker, J. Kohlmeyer, B. Volkmann-Kohlmeyer, R. A. Spotts, M. Serdani, P. W. Crous, K. W. Hughes, K. Matsuura, E. Langer, G. Langer, W. A. Untereiner, R. Lücking, B. Büdel, D. M. Geiser, A. Aptroot, P. Diederich, I. Schmitt, M. Schultz, R. Yahr, D. S. Hibbett, F. Lutzoni, D. J. McLaughlin, J. W. Spatafora, and R. Vilgalys. 2006. Reconstructing the early evolution of fungi using a six gene phylogeny. Nature 443:818–822.

Keeling, P. J., G. Burger, D. G. Durnford, B. F. Lang, R. W. Lee, R. E. Pearlman, A. J. Roger, and M. W. Gray. 2005. The tree of eukaryotes. Trends Ecol. Evol. 20:670–676.

Kron, K. A. 1997. Exploring alternative systems of classification. Aliso 15:105–112.

May, R. M. 1988. How many species are there on earth? Science 241:1441–1449.

May, R. M., and S. Nee. 1995. The species alias problem. Nature 378:447–448.

Mendoza, L., J. W. Taylor, and L. Ajello. 2002. The class Mesomycetozoa: a heterogeneous group of microorganisms at the animal-fungal boundary. Ann. Rev. Microbiol. 56:315–344.

Microscope. http://starcentral.mbl.edu/microscope/portal.php (accessed December 4, 2006).

Patterson, D. J. 1999. The diversity of eukaryotes. Am. Nat. 154(Suppl.):S96–S124.

Patterson, D. J. 2002. Changing views of protistan systematics. Pages 2–9 *in* An illustrated guide to the protozoa (J. J. Lee, G. F. Leedale and P. C. Bradbury, eds.). Society of Protozoologists, Lawrence, Kansas.

Pleijel, F., and G. W. Rouse. 2003. Ceci n'est pas une pipe: Clades and phylogenetic nomenclature. J. Zool. Syst. Evol. Res. 41:162–174.

Redhead, S. A., M. T. Cushion, J. K. Frenkel, and J. R. Stringer. 2006. *Pneumocystis* and *Trypanosoma cruzi*: Nomenclature and typifications. J. Euk. Microbiol. 53:2–11.

Simpson, A. G. B., and A., Roger. 2004. The real kingdoms of eukaryotes. Curr. Biol. 14:R693–R696.

Slapeta, J., D. Moreira, and P. Lopez-Garcia. 2005. The extent of protist diversity: insights from molecular ecology of freshwater eukaryotes. Proc. Royal Soc. B Biol. Sci. 272:2073–2081.

Whittaker, R. H. 1972. Evolution and measurement of species diversity. Taxon 21:213-251.

The World Health Report. http://www.who.int/whr/2004/en/ (accessed July 15, 2006).

First submitted 22 January 2007; reviews returned 31 March 2007; final acceptance 1 May 2007

Associate Editor: Tim Collins

Editor: Jack Sullivan