Syst. Biol. 61(1):150–164, 2012

DOI:10.1093/sysbio/syr089
Advance Access publication on September 7, 2011

Phylogenomic Analysis Resolves the Interordinal Relationships and Rapid Diversification of the Laurasiatherian Mammals

XUMING ZHOU, SHIXIA XU, JUNXIAO XU, BINGYAO CHEN, KAIYA ZHOU, AND GUANG YANG*

*Jiangsu Key Laboratory for Biodiversity and Biotechnology, College of Life Sciences, Nanjing Normal University, Nanjing 210046, China; *Correspondence to be sent to: Jiangsu Key Laboratory for Biodiversity and Biotechnology, College of Life Sciences, Nanjing Normal University, Nanjing 210046, China; E-mail: gyang@njnu.edu.cn.*

Received 28 October 2010; reviews returned 6 April 2011; accepted 23 June 2011
Associate Editor: Michael Charleston

Abstract.—Although great progress has been made in resolving the relationships of placental mammals, the position of several clades in Laurasiatheria remain controversial. In this study, we performed a phylogenetic analysis of 97 orthologs (46,152 bp) for 15 taxa, representing all laurasiatherian orders. Additionally, phylogenetic trees of laurasiatherian mammals with draft genome sequences were reconstructed based on 1608 exons (2,175,102 bp). Our reconstructions resolve the interordinal relationships within Laurasiatheria and corroborate the clades Scrotifera, Fereuungulata, and Cetartiodactyla. Furthermore, we tested alternative topologies within Laurasiatheria, and among alternatives for the phylogenetic position of Perissodactyla, a sister-group relationship with Cetartiodactyla receives the highest support. Thus, Pegasoferae (Perissodactyla + Carnivora + Pholidota + Chiroptera) does not appear to be a natural group. Divergence time estimates from these genes were compared with published estimates for splits within Laurasiatheria. Our estimates were similar to those of several studies and suggest that the divergences among these orders occurred within just a few million years. [Laurasiatheria; phylogenomics; rapid divergence; placental mammals.]

Mammalian phylogeny has recently received extensive attention, leading to great progress in this field. One widely recognized superclade of mammals is Laurasiatheria, one of several clades first proposed by Waddell, Okada, and Hasegawa (1999). Monophyly of Laurasiatheria has been corroborated by analyses of long concatenations of genetic sequences (e.g., Madsen et al. 2001; Murphy, Eizirik, Johnson, et al. 2001; Murphy, Eizirik, O'Brien, et al. 2001; Scally et al. 2001; Waddell et al. 2001; Waddell and Shelley 2003; Hallström and Janke 2010) and recently by inserted transposable elements that have very low levels of homoplasy (Kriegs et al. 2006). High support has also been found for many subclades of Laurasiatheria, many of which have traditionally been recognized as mammalian orders: Cetartiodactyla (artiodactyls and cetaceans), Perissodactyla (rhinoceroses, equids, and tapirs), Carnivora (carnivores), Chiroptera (megabats + microbats), Pholidota (pangolins), and Eulipotyphla (hedgehogs, soricid shrews, and moles). Interordinal relationships within this superorder, however, remain unclear even after extensive molecular analyses. Many laurasiatherian mammals are endowed with a highly specialized anatomy and behavior (e.g., flying, swimming, insectivorous, carnivorous, fast-running, etc.), and it has been hypothesized that many of these biological innovations may have evolved in Laurasia (modern day North America, Europe, and Asia) together with those observed in Supraprimates (Waddell et al. 2001) (or Euarchontoglires [Murphy, Eizirik, O'Brien, et al. 2001], e.g., primates and rodents). Knowledge of the phylogenetic relationships within Laurasiatheria and the elucidation of the position of its root are important for interpreting biogeographic patterns and evolutionary processes involved in the early diversification of placental mammals.

Among the orders of Laurasiatheria, hypotheses for the phylogenetic positions of Eulipotyphla and Chiroptera have played a central role in discussions of the root and center of origin for Laurasiatheria. Morphologists (e.g., Butler 1988; MacPhee and Novacek 1993) originally placed some eulipotyphlans in the order Lipotyphyla; however, Stanhope et al. (1998) presented genetic sequence evidence for a polyphyletic origin for lipotyphlans and they found one lipotyphlan clade positioned inside Afrotheria, which they named Afrosoricida. The second clade of "lipotyphylans," which positioned deep inside Laurasiatheria, was named Eulipotyphyla by Waddell, Okada, and Hasegawa (1999). However, based primarily on the incorporation of the complete mitochondrial genome of a mole (*Talpa europaea*), Eulipotyphla was subsequently suggested to be a paraphyletic group as well, with the mole more closely related to Fereuungulata (Waddell, Cao, Hauf, and Hasegawa 1999; Cetartiodactyla + Perissodactyla + Carnivora + Pholidota) and hedgehogs placed at a basal position in Eutheria (Mouchaty et al. 2000). However, paraphyly of Eulipotyphla is inconsistent with several subsequent molecular analyses of concatenated sequences of nuclear exons and mitochondrial rRNA genes that placed a monophyletic Eulipotyphla as the most basal lineage in the Laurasiatheria (Madsen et al. 2001; Matthee et al. 2001; Waddell et al. 2001; Kriegs et al. 2006; Nishihara et al. 2006). Additionally, some molecular studies (Onuma et al. 1998; Mouchaty et al. 2000; Narita et al. 2001; Nikaido et al. 2001) strongly supported a sister-group relationship between Eulipotyphla and Chiroptera named Insectiphillia (Waddell et al. 2001). Thus, the position of Eulipotyphla, either as a basal clade of Laurasiatheria or within Insectiphillia, has been a matter of great controversy. Many morphological studies have instead placed Chiroptera within the Superorder Archonta (Gregory 1910; Primates + Dermoptera + Scandentia + Chiroptera) (Novacek 1992; Shoshani and McKenna 1998); however, many molecular analyses have strongly contradicted

this clade and placed Chiroptera within Laurasiatheria (Waddell, Okada, and Hasegawa 1999; Nikaido et al. 2000; Madsen et al. 2001; Murphy, Eizirik, O'Brien, et al. 2001; Scally et al. 2001; Waddell et al. 2001). The perceived phylogenetic position of Chiroptera within Laurasiatheria varies depending on the combination of genes selected for analysis. For example, based on mitochondrial genomes, Pumo et al. (1998) presented strong evidence that Chiroptera is closely related to Fereuungulata. By contrast, several other studies that combined analyses of mitochondrial genomes and functional genes (Onuma et al. 1998; Mouchaty et al. 2000; Narita et al. 2001; Nikaido et al. 2001) support monophyly of Insectiphillia.

The phylogenetic position of Perissodactyla within Laurasiatheria has been another controversial issue. Based on morphological data (e.g., Shoshani and McKenna 1998), Perissodactyla was originally grouped with other hoofed animals in the group Ungulata (McKenna 1975), which also included Cetartiodactyla, Tubulidentata (aardvarks), Proboscidea (elephants), Sirenia (dugongs and manatees), and Hyracoidea (hyraxes). Subsequently, many molecular studies have placed Perissodactyla and Cetartiodactyla together within Laurasiatheria as members of the clade Fereuungulata (Pumo et al. 1998; Waddell, Okada, and Hasegawa 1999; Madsen et al. 2001; Murphy, Eizirik, Johnson, et al. 2001; Murphy, Eizirik, O'Brien, et al. 2001; Waddell et al. 2001). However, relationships within Fereuungulata are still unclear, and a cascade of studies over the last decade has resulted in many alternative topologies (Fig. 1). Most authors agree with one aspect of the phylogeny of Shoshani and McKenna (1998), which placed Carnivora and Pholidota in the clade Ferae (e.g., Waddell, Okada, and Hasegawa 1999; Murphy, Eizirik, Johnson, et al. 2001; Murphy, Eizirik, O'Brien, et al. 2001; Waddell et al. 2001), but all possible hypotheses for the relationships of the remaining clades of Fereuungulata have received at least some support: Perissodactyla + Cetartiodactyla (Irwin and Wilson 1993, a close second

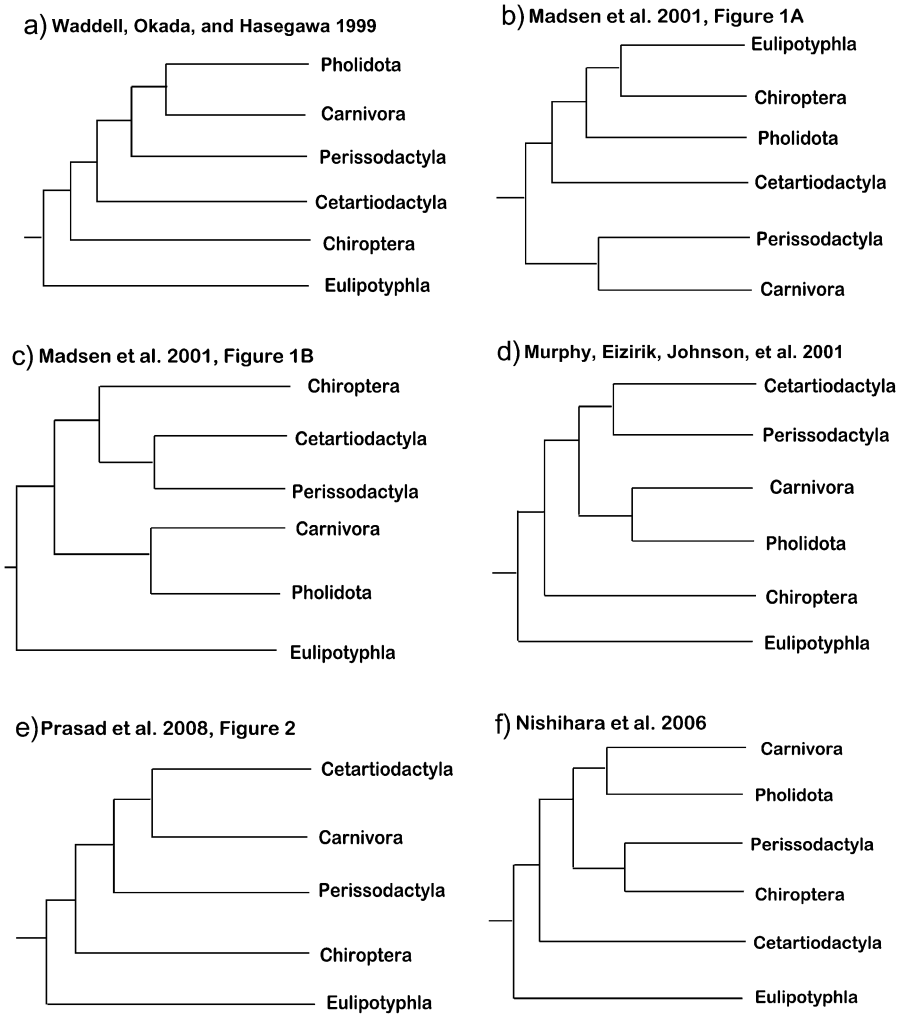

FIGURE 1. Alternative hypotheses for the interorder relationships of major Laurasiatherian lineages based on molecular sequences data (a–e) and retroposon analysis (f). Clade support values and the size of the data sets used are available in Table 1.

in Waddell, Okada, and Hasegawa [1999] and favored in Murphy, Eizirik, Johnson, et al. [2001] and Waddell et al. [2001]), Perisodactyla and Ferae (Waddell, Okada, and Hasegawa 1999; Murphy, Eizirik, O'Brien, et al. 2001; Kullberg et al. 2006; Waters et al. 2007), and even Carnivora + Cetartiodactyla (Prasad et al. 2008, although the exact position of Pholidota was not determined). The name Cetungulata was attached to the clade (crown or stem group unspecified) of Cetartiodactyla plus Perissodactlya by Irwin and Wilson (1993), but it was unclear if they also intended Cetungulata to include Hyracoidea, which they did not sample in their analyses because it was widely thought at the time to be the closest mammalian order to Perissodactyla. Independently, the exclusive crown group of Cetartiodactyla + Perissodactyla was named Euungulata in Waddell et al. (2001). Here we use the name Euungulata with the meaning of Cetungulata. In addition, one supermatrix (5708 bp; mitochondrial RNA gene and 3 nuclear genes) compiled by Madsen et al. (2001) even supported a basal position for Zooamata (Waddell, Okada, and Hasegawa 1999; Perissodactyla + Ferae); however, their hypothesis for the phylogenetic positions of fereuungulates as well as the others described above generally did not receive strong Bayesian posterior probabilities (PPs) or bootstrap (BP) supports. The relative support values of alternative hypotheses through time is summarized in Table 1.

Independent assessments for the phylogenetic positions of Perissodactyla and Chiroptera have been developed using retroposons. Four retroposon loci have been found to contradict Fereuungulata by placing Chiroptera within a clade with Carnivora, Pholidota, and Perissodactyla. This grouping was named Pegasoferae (Nishihara et al. 2006). Thus, instead of favoring one of the phylogenetic hypotheses described in the preceding paragraph, retroposon data support yet another hypothesis. Clearly, the controversy over the phylogenetic positions of Perissodactyla and Chiroptera is an important issue that needs to be settled.

Here, we present a phylogenetic analysis using both concatenated and partitioned analyses of a data set that includes nearly 100 nuclear protein-coding genes sampled from representatives all major lineages of laurasiatherian mammals (including sequences for taxa, e.g., pangolin, that are often missing from other analyses). We used a large amount of newly sequenced data to produce a data set that is largely independent of data upon which existing hypotheses of relationships among Laurasiatherian mammals have been based. This allows us to proceed in a hypothesis-testing framework. We used our data to perform topology tests between previously proposed, competing hypotheses. We also applied a recently developed method for estimating divergence times to these new data and then compared our estimates with those published previously (Waddell, Okada, and Hasegawa 1999; Waddell et al. 2001; Springer et al. 2003; Kitazoe et al. 2007). Finally, we used completed draft genomes to construct and analyze a data set with fewer taxa, but even more genes.

Materials and Methods

Taxonomic Sampling

Data set A.—The 97 markers that were assembled for data set A were developed from OrthoMaM, a database that consists of orthologous genomic markers for placental mammalian phylogenetics (Ranwez et al. 2007). The details of search criterion, characters, and primers for markers have been introduced in our previous study (Zhou et al. 2011). All these markers were amplified and sequenced from the following laurasiatherian species: common bottlenose dolphin (*Tursiops truncatus*), minke whale (*Balaenoptera acutorostrata*), hippopotamus (*Hippopotamus amphibius*), wild boar (*Sus scrofa*), bactrian camel (*Camelus bactrianus*), white rhinoceros (*Ceratotherium simum*), and Chinese pangolin (*Manis pentadactyla*). All sequences in our data set A are new to our study and were deposited in GenBank (679 sequences including GU301691–GU301692, GU301694–GU301696, GU301699, GU301703, GU301705–GU301709, GU301713–GU301714, GU301716–GU301718, GU301721, GU301723–GU301725, GU301727, GU301729, GU301731, GU301734, GU301736, GU301738, GU301740, GU301742–GU301743, GU301746–GU301747, GU301749, GU301751, GU301754–GU301755, GU301757, GU301759–GU301762, GU301765, GU990619–GU991248, GU991264–GU991270). We also retrieved these regions for the cow (*Bos taurus*), horse (*Equus caballus*), dog (*Canis familiaris*), microbat (*Myotis lucifugus*), shrew (*Sorex araneus*), hedgehog (*Erinaceus europaeus*), human (*Homo sapiens*), and mouse (*Mus musculus*) from OrthoMaM for subsequent analyses. The present OrthoMaM database was missing several markers for *Erinaceus europaeus* and *Sorex araneus*; thus, we amplified "replacement" markers from the hedgehog (*E. europaeus*) and the Asian gray shrew (*Crocidura attenuata*) and incorporated them into our analyses (also deposited in GenBank with accession numbers GU991249–GU991263). We also created three additional smaller data sets and alignments using data set A to incorporate taxa for which less than the 97 markers had been sequenced. These are a 67-gene data set with sequences for 16 taxa (including the cat *Felis catus*), an 82-gene data set that includes 17 taxa (including the elephant *Loxodonta africana* and the opossum *Monodelphis domestica)*, and a 60-gene dataset of 18 taxa (with addition of the guinea pig *Cavia porcellus*, rabbit *Oryctolagus cuniculus*, and squirrel *Spermophilus tridecemlineatus*).

Data set B.—We also did an "exhaustive" search for all the orthologs of >800 bp from draft genome sequences of the dolphin, cow, alpaca (*Vicugna pacos*), horse, dog, cat, megabat (*Pteropus vampyrus*), microbat, shrew, hedgehog, armadillo (*Dasypus novemcinctus*), human, mouse, elephant, and opossum. As a result, 2771 genes that met this criterion were downloaded and then genes with relatively slow and moderate evolutionary sites (relative evolutionary rate < 1.2) were retained (1608 genes) (Criscuolo et al. 2006; Ranwez et al. 2007).

TABLE 1. Some typical clades within Laurasiatheria as suggested in the previous studies and their supports received from Bayesian PPs and BP analyses

	Laurasiatheria (Waddell, Okada, and Hasegawa 1999)	Eulipotyphla (Waddell, Okada, and Hasegawa 1999)	Scrotifera (Waddell, Cao, Hauf, and Hasegawa 1999)	Fereuungulata (Waddell, Cao, Hauf, and Hasegawa 1999)	Ferae (Shoshani and McKenna 1998)	Euungulata (Waddell et al. 2001)	Zooamata (Waddell, Okada, and Hasegawa 1999)	Perissodactyla + Chiroptera	Cetartiodactyla + Carnivora	Pegasoferae (Nishihara et al. 2006)	
Waddell, Okada, and Hasegawa (1999)	~90%	~60%	~90%	~60%	~70%	<50%	~60%				Nu+ Mitogenome
Arnason et al. (2002)	100% (ML) 63% (NJ) 71% (MP)		60% (ML) 63% (NJ)	97% (ML) 89% (NJ) 71% (MP)	82% (ML)		97% (ML) 96% (NJ) 83% (MP)				Mitogenome
Murphy, Eizirik, Johnson, et al. (2001)	74% (NJ) 99% (MP) 99% (ML)	100% (NJ) 97% (MP) 98% (ML)			<50% (NJ) <50% (MP) 63% (ML)	63% (NJ) 72% (MP) 67% (ML)					18 Nu genes (9779 bp)
Waddell et al. (2001)	100% (ML) 100% (ML) 100% (ML)	 100% (ML) 100% (ML)	95% (ML) 69% (ML) 98% (ML)	82% (ML) 56% (ML) 96% (ML)	 89% (ML) 65% (ML)	 95% (ML) 92% (ML)	73% (ML)				Mitogenome Nu Nu+ Mitogenome (23,997 bp)
Scally et al. (2001)	99% (ML)	96% (ML)					<50% (ML)				Nu+ Mitogenome (5708 bp)
	100% (ML) 100% (ML)	100% (ML) 99% (ML)	<50% (ML) <50% (ML)	<50% (ML) <50% (ML)	<50% (ML) <50% (ML)	<50% (ML)	 <50% (ML)				Nu (2947 bp) Nu+ Mitogenome (8655 bp)
Murphy, Eizirik, O'Brien, et al. (2001)	1.00 (BI) 100% (ML)	1.00 (BI) 100% (ML)	1.00 (BI) 94% (ML)	0.98 (BI) 59% (ML)	1.00 (BI) 91% (ML)		0.74 (BI) 42% (ML)				Nu (16,397 bp)
Waddell and Shelley (2003)	0.99 (BI)	1.00 (BI)			0.63 (BI)			0.93 (BI)		1.00 (BI)	6 Nu genes + mt tRNA
Nishihara et al. (2006)	9 loci		3 loci			1 locus	1 locus			4 loci	Retroposons
Springer et al. (2007)	1.00 (BI)	1.00 (BI)	1.00 (BI)	1.00 (BI)	1.00 (BI)	0.48 (BI)					20 Nu genes (14,326 bp)
Prasad et al. (2008)			48% (RY-ML) 1.00 (BI) 98% (RY-ML) 1.00 (BI)	 64% (RY-ML) 1.00 (BI)		18% (RY-ML) 1.00 (BI)			44% (RY-ML) 1.00 (BI)		Nu (60 Mb) coding Nu (60 Mb) coding + noncoding
Hou et al. (2009)						97.5% (ML) 1.00 (BI)					2705 genes (33,991,128 bp)

NJ = neighbor-joining method; Nu = nuclear data.

A series of Perl scripts was written to help in extracting and preparing the relevant alignment sequences for subsequent phylogenetic analyses.

Laboratory Protocols

Methods for isolating genomic DNA and sequencing are the same as those described by Zhou et al. (2011).

Alignment and Combination of Data Sets

Initial sequence assembly was carried out using the ContigExpress software (Invitrogen). Chromatograms for exon sequences were viewed individually to ensure accuracy and were aligned with the consensus sequence from public database (OrthoMaM). Individual alignments for data set A were produced using Clustal W (Thompson et al. 1997) on the translated protein sequences. All alignments were checked by Muscle (Edgar 2004), and poorly aligned regions were improved manually. The final concatenated sequence alignments used for subsequent analyses were chosen using Gblock software (Castresana 2000). Intergenomic phylogenetic congruence was tested using a likelihood ratio test (LRT) implemented in Concaterpillar 1.4 (Leigh et al. 2008). Concaterpillar performed per-genome maximum likelihood (ML) analyses on identical taxon sets (15 taxa) using the general time reversible (GTR) model implemented in RAxML 7.2.3 (Stamatakis 2006; Stamatakis et al. 2008) at $\alpha = 0.05$ level.

Phylogenetic Analyses

The effects of mutational saturation at third codon positions in the 97-gene combined data set were investigated. Uncorrected pairwise distances for transitions and transversions among all taxa in the data set were plotted against the GTR model combined with a gamma distribution or with an inverse Gaussian distribution across invariant sites distances (Waddell and Steel 1997) to detect mutational saturation at combined positions of the first and second codons, as well as at third positions. Third codon position transitions displayed the strongest evidence of mutational saturation (Supplementary Information [SI] Fig. 1, available from http://www.sysbio.oxfordjournals.org); thus, we performed phylogenetic analyses on amino acid sequences, combined first and second codon positions, and the third codon position for the 97-gene data set. In an attempt to reduce noise, phylogenetic analyses were performed by coding the nucleotides as purines and pyrimidines (RY coding) (Phillips et al. 2004) for these positions and the more extreme approach of removing all third codon position sites for data set B. Phylogenetic trees were reconstructed using ML (RAxML software, version 7.2.3) and Bayesian inference (BI) (MrBayes software, version 3.1.2; Huelsenbeck and Ronquist 2001). The best-scoring ML tree was inferred using the novel rapid BP algorithm and ML searches after conducting 1000 RAxML runs. Akaike information criterion (AIC) as calculated by MrModeltest software (Nylander 2004) was used to determine the most appropriate model for each of the 97 genes and for the full data matrix. In MrBayes 3.1.2, two independent sets of Markov chain Monte Carlo (MCMC) chains were run, each with three heated and one unheated chain for 10,000,000 generations. Trees were sampled every 1000 generations, and convergence was confirmed with the help of are we there yet? (AWTY) graphical analysis (Nylander et al. 2008). A consensus tree was obtained from BI trees for the 97-gene data set using the *majrule* command in Clann, version 3.0.0 (Creevey and McInerney 2005). Because standard Bayesian PP values can be very unreliable when an incorrect model is used, and nuclear mammalian genes show clear deviations in base composition and relative rates of substitution types from those of commonly employed models (Waddell et al. 2009), we also implemented gene partitions and specified a model for each gene included in the Bayesian analyses of the 97-gene data set. Additionally, approximately unbiased (AU) tests (Shimodaira 2002) were conducted in Consel version 0.1 (Shimodaira and Hasegawa 2001) between the best ML tree and alternative topologies to evaluate whether their likelihoods were significantly different. Kishino–Hasegawa (KH; Kishino and Hasegawa 1989) and Shimodaira–Hasegawa (SH) tests (Shimodaira and Hasegawa 1999) were performed to measure the congruence between single-gene trees.

Molecular Dating Analyses

Molecular dating analyses were conducted with BEAST v1.5.0 (Drummond and Rambaut 2007) using the 97 and 60 nuclear-gene data matrices, with a relaxed-clock MCMC approach under an uncorrelated lognormal model. In view of modern Bayesian methods that allow for the incorporation of a prior distribution of ages ("age constraints") into these fossil calibrations, we chose prior age distributions so that the youngest age of the distribution corresponded to the youngest possible age of the fossil. We chose a standard deviation (SD) of the distribution so that 95% of the distribution fell within the geological time period of the fossil (i.e., 5% of the tail extended into older ages). Five fossils were used as calibration age constraints in the present study. The first calibration point is the age of the divergence between Hippopotamidae and Cetacea, which is at least Ypresian (Eocene: 55.8–48.6 Ma) in age based on the fossil cetacean *Pakicetus* (Thewissen et al. 2001). We chose a lognormal distribution so that the earliest possible sampled age corresponded to 48.6 Ma and the older 95% credible interval (CI) encompassed the beginning of the Ypresian (55.8 Ma) (SD = 1.2). Previous studies have also used this fossil as a calibration point, with a similar age and confidence interval (52 Ma with a confidence interval of 49–61 Ma) (Waddell et al. 2001; Waddell 2008). The second calibration point is the horse/rhino split, which has been updated to 58 Ma by Waddell (2008). This divergence for fossil

calibration data close to the age of placental intraordinal divergences (Waddell, Okada, and Hasegawa 1999) and SD was 1.5 (normal distribution). The age of the root of crown Laurasiatheria was the third calibration point, and it is based on the earliest record of Eulipotyphla, the late Cretaceous *Otlestes* (Averianov and Archibald 2005; Prasad et al. 2007). We chose a lognormal distribution so that the earliest possible sampled age corresponded to 93.5 Ma and the older 95% CI encompassed the beginning of the late Cretaceous (99.6 Ma) (SD = 1.099). The phylogenetic affiliation of *Otlestes* is controversial (Archibald 2003), so we also performed an analysis in which the third calibration point was replaced by a calibration based on the origin of Rodentia. To do this, we used the 60-gene data set in which additional taxa (guinea pig, rabbit, and squirrel) are included. It appears that Rodentia most likely evolved in Asia and moved into North America, which yields an estimate of 60 Ma (SD = 1.5, normal distribution) (Waddell 2008). Each BEAST analysis consisted of 2×10^7 generations with a random starting tree, birth–death default priors (with the exception that we used a uniform [0, 100] prior distribution for the GTR substitution rates), and sampled every 1000 generations. Cumulative PP plots for each clade were constructed to determine convergence with the help of AWTY (Nylander et al. 2008), and PPs $\geqslant$0.95 were considered to indicate strong support for a clade (Huelsenbeck and Rannala 2004).

Results

Two main data sets were generated and compiled in the present study. Data set A contained a total of 97 aligned genes (46,152 bp of sequence data) for 15 taxa (SI file 1), and its subsets include an alignment of 82 genes (36,441 bp) for 17 taxa (with addition of the elephant and opossum, SI file 2), an alignment of 67 genes (29,928 bp) for 16 taxa (with addition of the cat, SI file 3), and an alignment of 60 genes (27,405bp) for 18 taxa (with addition of the guinea pig, rabbit, and squirrel, SI file 4). These alignments have been submitted to TreeBASE (http://purl.org/phylo/treebase/phylows/study/TB2:S11415). Markers assembled in these alignments were most likely to be orthologous in the above species because they have been annotated by ENSEMBLE and amplified as a single band with identical and ideal size in most polymerase chain reactions (Zhou et al. 2011). Data set B consisted of a total of 1608 named orthologs (2,175,102 bp) for 17 taxa (SI file 5), including all laurasiatherian mammals for which draft genome sequences are available (except for the pig because its whole genomic sequences are in low coverage). Each data set was partitioned by codon or gene, and phylogenetic trees were obtained by BI and ML methods (Fig. 2; SI Fig. 2). The AIC selected GTR+I+G as the optimal model for Bayesian searches for all codons and combined the first and second codon positions of the concatenated 97 and 82 gene sequences. The best-scoring ML searches were performed under the model GTRGAMMA or GAMMAWAG. In addition, a mixed branch length model (Kolaczkowski and Thornton 2008) and an empirical profile mixture model (Quang et al. 2008) were applied in ML analyses to reduce the effect of heterogeneity among genes and sites. To accomplish this we used the software of Phyesta (Guindon and Gascuel 2003; Hanson-Smith et al. 2009) and RAxML version 7.2.3 (Stamatakis 2006; Stamatakis et al. 2008, CAT model).

In all the analyses, (human, mouse), (elephant, armadillo), and opossum were successive sister groups to laurasiatherian mammals with strong support regardless of partitioning strategy, data set, or phylogenetic optimality criterion (Fig. 2; SI Fig. 2). Our results show that phylogenetic analyses of the combined data with one model and analyses with each gene partition given its own model yield the same tree (97-gene alignment). ML method combined with a mixed branch length and an empirical profile mixture model also recovered a phylogeny that is identical to that recovered with GTR series model. We found inconsistency only in the analyses of different codon positions. For example, analyses of all codon positions or only the third codon position of the 97-gene data set both strongly supported the separation of Perissodactyla from Cetartiodactyla, and the former analysis even placed Perissodactyla as basal within Laurasiatheria or Fereuungulata (SI Fig. 2a,b). Combined analyses of the first and second codon positions of the 97-gene data set, however, identified Perissodactyla as the sister group to Cetartiodactyla to form the clade Euungulata (Fig. 2a). A similar topology was also inferred when we analyzed the 97-gene data set based on amino acid sequences level (Fig. 2c). In addition, although the LRT implemented in Concaterpillar detected significant topological incongruence between the nuclear genes with three codon positions included (the largest concatenated data set consisted of 37 genes), it does not reject congruence among the 97 nuclear genes with the first + second codon positions at an αlevel of 0.05 (the final concatenated data set consists of 96 genes, except *zscan29*). So LRT indicates that combined analyses of first + second codon positions for 97 genes, which do not violate model assumptions, are permissible. We further analyzed the data set with the slow–fast method (Kostka et al. 2008) that was designed to reduce the effect of long-branch attraction (LBA) and increase the phylogenetic signal-to-noise ratio (Delsuc et al. 2005). The eight most trimmed alignments (S0–S7) containing increasing fractions of fast-evolving positions were analyzed with the ML method. The results showed that the alignments (S1 and S2) with comparably slow-evolving positions support the Euungulata hypothesis, whereas alignments S3–S7 support the separation of Perissodactyla from Cetartiodactyla and a sister-group relationship between Cetartiodactyla and Ferae (SI Fig. 2d,e). It seems that the sister-group relationship between Cetartiodactyla and Ferae is artificial and caused by LBA among third codon positions.

BI and ML analyses of both the 97- and the 82-gene data sets strongly supported the sister-group

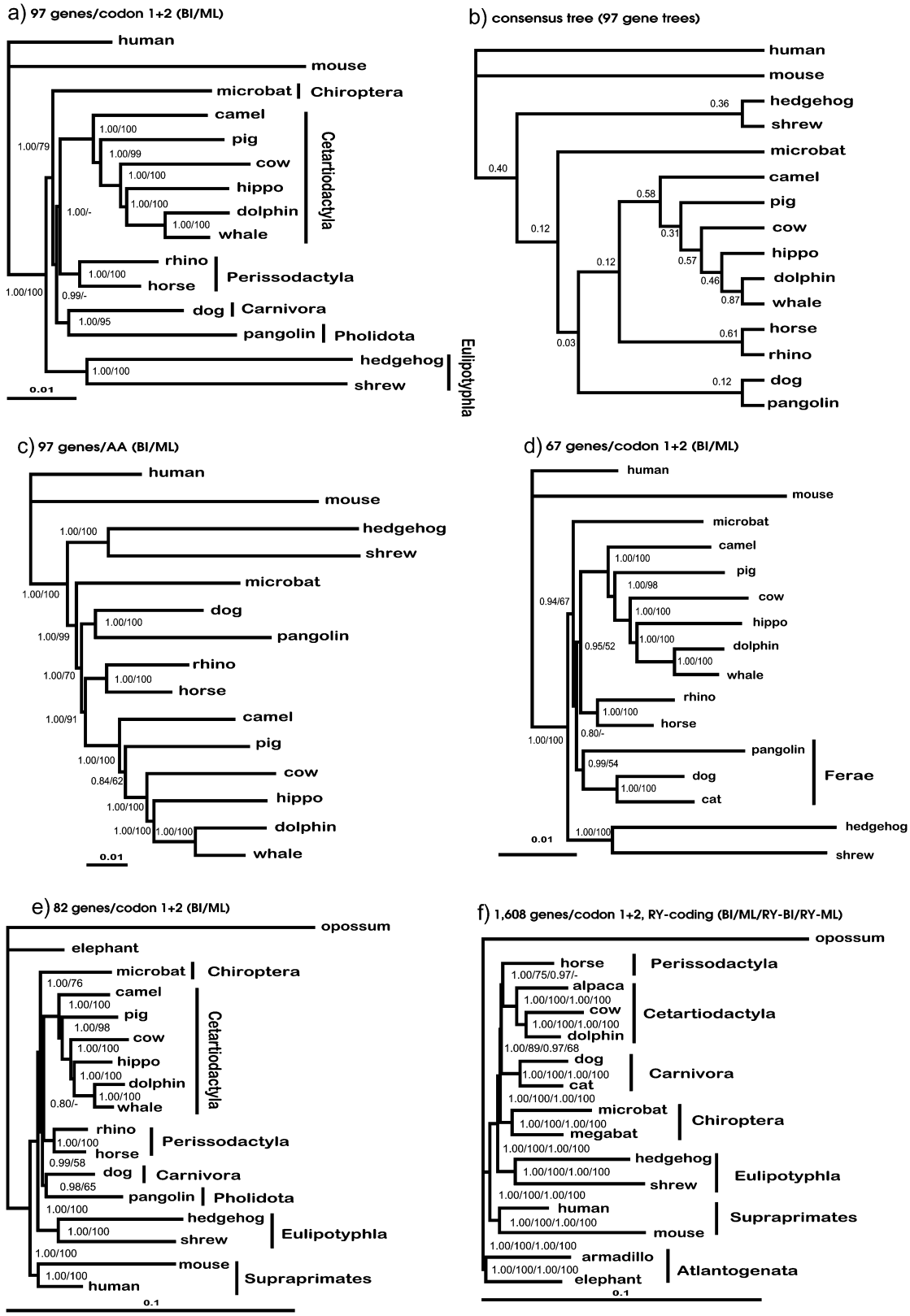

FIGURE 2. Phylogenetic trees reconstructed using BI and ML methods based on two different data sets. Integers associated with branches are BP support values from ML analyses, whereas values of 1 or less are Bayesian PPs. Dash denotes BP support values that are <50%. Codon 1+2 and RY coding refer to the concatenated sequences combined with the first and second codon positions and combined with RY recoding, respectively. A consensus tree based on a majority rule of 97 BI trees from the 97-gene data set is shown in (b). Numbers associated with each branch represent the proportion of universally distributed input trees that contain that particular split.

relationship between the dog (Carnivora) and the pangolin (Pholidota) (Fig. 2a,c). We also introduced a new 67-gene data set incorporating the cat (Carnivora) in order to break the carnivore edge. In the resulting trees, the cat clusters with the dog and together they are sister to the pangolin (BP = 0.99; PP = 54%, Fig. 2d). ML method recovered trees with Chiroptera as sister to other fereuungulates, and it also found moderate BP support (79%) for this clade in the combined analyses of the first and second codon positions for the 97-gene data set (Fig. 2a). This arrangement was further supported by maximal BP support (100%) when the larger data set was analyzed (i.e., data set B) (Fig. 2f). Yet two nodes in the ML reconstruction still did not receive maximal support even with the larger data set B: the node supporting the monophyly of Euungulata (BP = 75%, RY-BP < 50%) and the node supporting the sister-group relationship of Euungulata and Carnivora (BP = 89%, RY-BP = 68%) (Fig. 2f). In addition, the BI and ML methods when applied to amino acid sequences of 1608 genes even supported a different phylogenetic relationship, specifically Perissodactyla as the sister group to Carnivora (SI Fig. 2c), as depicted in Waddell, Okada, and Hasegawa (1999).

Using the 97-gene data set, the AU test successfully rejected four of the six alternative topologies (Fig. 1) involving Perissodactyla, Cetartiodactyla, Carnivora, and Pholidota at the 0.05 significance level (Table 2). The AU test using the 1608-gene data set successfully rejected the clade Insectiphillia and a sister-group relationship between Cetartiodactyla and Carnivora at a 0.001 significance level. The nodes within Cetartiodactyla received strong PP and BP support based on data set A and its subalignments, and, in addition, all these analyses interpreted cetaceans as within artiodactyls as the sister to the hippopotamus (Fig. 2; SI Fig. 2). However, the AU test did not reject the "pig + camel" grouping at a 0.05 significance level (Table 2).

We also investigated the importance of prior hypotheses (Fig. 1) using all codon positions and the combined first and second codon positions of the 97-gene data set. The likelihood scores from each gene were combined into a support value in two different ways, that is, the methods of Adachi and Hasegawa (1996) and Waddell, Cao, Hasegawa, and Mindell (1999), both shown in Waddell and Shelley (2003) as well. The $\sum z^2$ statistic does not reject any of the six prior trees shown in Figure 1. The z statistic gave a similar result and does not reject the trees of Waddell, Okada, and Hasegawa (1999), Nishihara et al. (2006), and Madsen et al. (2001, figure 1B) (SI Table 1 and SI Table 2). The BI tree was reconstructed for each gene and both SH and KH tests of all those 97 trees (SI Tree file) showed that no tree is identical to any of the others or to the resolved tree based on the combined genes alignment. However, a majority-rule consensus tree (Fig. 2b) derived from the 97-gene BI trees provided a tree topology that was identical to the tree based on combined analyses of the first and second codon positions as shown in Figure 2a.

The divergence times inferred by the Bayesian-relaxed clock analyses were based on two different fossil constraints (Table 3). Generally, the divergence estimates in the present study are somewhat closer to those of Waddell, Okada, and Hasegawa (1999) and Waddell et al. (2001) than those of Springer et al. (2003) and Kitazoe et al. (2007), which are somewhat younger. Our estimates based on different calibration points gave similar results within Cetartiodactyla and for the split between Perissodactyla and Cetartiodactyla, but they varied with regard to the origin of Laurasiatheria (Table 3). For example, calibration with the root of crown Laurasiatheria suggests that extant Boreotheria began to diversify in the early Cretaceous ≈107 Ma (95% HPD, 94–124 Ma), whereas calibration with Rodentia proposed a period in the late Cretaceous ≈90 Ma. Despite these differences, both sets of divergence dates indicate that the six major laurasiatherian lineages diversified relatively rapidly in the late Cretaceous, with the earliest divergence dated to 95 Ma (95% HPD, 94–98 Ma) and the latest divergence (i.e., Carnivora and Pholidota) dated to 66 Ma (95% HPD, 52–82 Ma) (Table 3). The origin of the extant crown group of fereungulates was dated somewhat later in the Cretaceous, ≈86 Ma/78 Ma (Table 3).

Discussion

Eulipotyphla: The Basal Branch of Laurasiatheria?

Resolution of the phylogenetic relationship among basal clades of laurasiatherian mammals is essential for addressing many important questions concerning the diversification and evolution of placental mammals. Despite the numerous, intensive phylogenetic studies (Onuma et al. 1998; Waddell, Okada, and Hasegawa 1999; Mouchaty et al. 2000; Madsen et al. 2001; Murphy, Eizirik, Johnson, et al. 2001; Murphy, Eizirik, O'Brien, et al. 2001; Narita et al. 2001; Waddell et al. 2001; Kriegs et al. 2006; Nishihara et al. 2006), considerable disagreement still persists for basal laurasiatherian relationships. The present phylogenetic analyses provide strong support for the position of Eulipotyphla as the most basal branch of Laurasiatheria and thus the sister group to the remaining extant orders (Waddell et al. 2001) (Fig. 2; SI Fig. 2c). Resolution of this issue is clearly the result of the addition of more genes as compared with previous studies. For example, Narita et al. (2001) conducted maximum parsimony MP) and ML analyses with two cDNA sequences, which supported the Eulipotyphla + Chiroptera basal hypothesis. In the present study, trees derived from the BI and ML analyses of first and second codon positions from 97 genes (data set A) placed Eulipotyphla alone as the sister group to the remaining laurasiatherian mammals, whereas analyses that used RY coding supported the hypothesis of a basal position for Eulipotyphla + Chiroptera (data not shown). However, trees reconstructed by BI and ML methods for the 1608-gene data set fully support

TABLE 2. LRTs of alternative topologies as implemented in CONSEL package

Tree	ln L	Δ ln L	KH probability	SH probability	AU probability
97 genes/codon 1+2 data set					
This study (Murphy, Eizirik, Johnson, et al. 2001)	−99171.8	Best	0.800	0.990	0.875
Prasad et al. (2008, figure 2)	−99183.4	11.6	0.200	0.627	0.243
Madsen et al. (2001, figure 1B)	−99198.0	26.2	0.034	0.365	0.037
Waddell, Okada, and Hasegawa (1999)	−99198.5	26.7	0.014	0.334	<0.001
Nishihara et al. (2006)	−99220.3	48.5	0.013	0.115	0.013
Madsen et al. (2001, figure 1A)	−99330.2	158.4	0.000	0.000	<0.001
97 genes/codon 1+2 data set					
((Fereuungulata, Chiroptera), Eulipotyphla)	−99171.8	Best	0.812	0.812	0.815
(Fereuungulata, (Chiroptera, Eulipotyphla))	−99188.6	16.8	0.188	0.188	0.185
97 genes/codon 1+2 data set					
(Euungulata, (Carnivora, Pholidota))	−99171.8	Best	1.000	1.000	1.000
((Euungulata, Carnivora), Pholidota)	−99265.6	93.8	<0.001	<0.001	<0.001
97 genes/codon 1+2 data set					
((Cetruminantia, Suina), Camelidae)	−99171.8	Best	0.938	0.972	0.963
(Cetruminantia, (Suina, Camelidae))	−99203.8	32.0	0.062	0.080	0.067
((Cetruminantia, Camelidae), Suina)	−99210.7	38.9	0.031	0.037	0.034
1608 genes/codon 1+2 data set					
((Fereuungulata, Chiroptera), Eulipotyphla)	−5181495.1	Best	1.000	1.000	1.000
(Fereuungulata, (Chiroptera, Eulipotyphla))	−5182386.3	891.2	<0.001	<0.001	<0.001
1608 genes/ codon 1+2 data set					
((Cetartiodactyla, Perissodactyla), Carnivora)	−5181495.1	Best	0.649	0.779	0.652
(Cetartiodactyla, (Carnivora, Perissodactyla))	−5181537	41.9	0.351	0.500	0.350
((Cetartiodactyla, Carnivora), Perissodactyla)	−5181860.2	365.1	<0.001	0.001	0.001

(100%/1.00/100%) a basal position for Eulipotyphla and a more apical position for Chiroptera. This result was obtained when first and the second codon positions are analyzed, as well as when RY coding analyses were employed. (Fig. 2f). In addition, an AU test rejected the Eulipotyphla + Chiroptera topology with statistical significance (Table 2), which further supports our preferred phylogenetic hypothesis (Fig. 2a–f).

Phylogenetic Relationships among Other Laurasiatherian Orders

Despite extensive phylogenetic analyses of placental, phylogenetic relationships, especially within Scrotifera (Waddell, Cao, Hauf, and Hasegawa 1999; Chiroptera + Pholidota + Carnivora + Perissodactyla + Cetartiodactyla), remain controversial. Morphological and molecular evidence supports several alternative positions for Chiroptera (Pettigrew 1986; Novacek 1992; Pumo et al. 1998; Shoshani and McKenna 1998; Nikaido et al. 2000), and as a consequence, a systematic assignment of this order was lacking in the most recently published review of Laurasiatheria (Nishihara et al. 2006). In the present study, Chiroptera was the sister group to Fereuungulata in almost all the ML and BI trees (except for the combined analyses of the 97-gene data set and RY coding, data not shown), which provides the strongest support so far for the placement of Chiroptera (Fig. 2; SI Fig. 2). Furthermore, the monophyletic status of Chiroptera (megabats + microbats) also received high BP and PP support in analyses of data set B both at the nucleotide and at the amino acid levels (Fig. 2f; SI Fig. 2c), which is consistent with previous studies on a large and diverse set of morphological features (Gunnell and Simmons 2005) as well as mitochondrial and nuclear nucleotide sequence data (Miyamoto 1996; Van den Bussche and Hoofer 2001). Interestingly, the nesting of bats within Fereuungulata was also recovered in some analyses of the present study. The 97- and 82-gene data sets, when analyzed with RY coding, suggest that Chiroptera + Eulipotyphla is the sister clade of Cetartiodactyla + Ferae, whereas MP analysis of first and second codon positions from data set B support Chiroptera as the sister group of Cetartiodactyla (data not shown). These analytical results disagree with some previous studies, which identified a "([(Perissodactyla, Carnivora), Chiroptera], Cetartiodactyla)" or "([(Perissodactyla, Artiodactyla), Chiroptera], Carnivora)" topology (Nishihara et al. 2006; Prasad et al. 2008) but agree to some extent with the finding of Hou et al. that a cow + bat clade was the sister group of a horse + dog clade (Hou et al. 2009; also found in Hallström and Janke 2010). Considering the weak BP support and very narrow divergence times among Scrotifera lineages, we attribute little significance to these new arrangements and suspect that these patterns are more reflective of incomplete lineage sorting.

As described in the introduction, previous morphological and molecular studies have come to very different placements for the laurasiatherian order Perissodactyla. Waddell, Okada, and Hasegawa (1999) suggested a sister-group relationship between

TABLE 3. Divergence times of lineages estimated from Bayesian phylogenetic analyses of 97- (calibrated with (1), (2), and (3), see Materials and methods section) and 60-gene data sets (calibrated with (1), (2), and (4)) using a lognormal relaxed molecular clock

Clade	Waddell, Okada, and Hasegawa (1999)	Waddell et al. (2001)		Springer et al. (2003)	Kitazoe et al. (2007)		Age	
		HR	WH	1st + 2nd	MVS-F_{IR}	ML- F_{IR}	97 genes/codon 1+2	60 genes/codon 1+2
Cetacea	—	20	25	29	24	21	27 (15–39)	28 (15–42)
Whippomorpha	—	42	52	53	49	49	**50** (49–53)	**50** (49–52)
Cetruminantia	—	47	59	54	53	57	54 (50–58)	53 (50–58)
Artiofabula	—	55	68	60	—	—	64 (56–71)	61 (54–69)
Cetartiodactyla	~ 66	60	75	63	55	64	68 (60–77)	65 (57–73)
Perissodactyla	—	55	69	57	52	52	**58** (55–61)	**58** (55–61)
Euungulata	—	77	96	—	63	73	83 (76–91)	76 (66–89)
Ferae	~ 73	76	95	75	—	—	76 (58–87)	66 (52–82)
Fereuungulata	~ 81	81	101	79	66	75	86 (79–93)	78 (68–91)
Scrotifera	~ 85	86	107	80	71	81	90 (83–95)	81 (70–95)
Eulipotyphla	—	79	99	71	69	80	72 (55–87)	66 (45–86)
Laurasiatheria	~ 95	91	114	83	72	86	**95** (94–98)	87 (74–103)
Supraprimates	~ 96	94	117	85	70	82	93 (67–119)	67 (60–77)
Boreotheria	~ 98	101	126	91	74	96	107 (94–124)	90 (76–107)

Notes: Units are in Ma. HPD along with the age for each node is listed within parentheses. HR means horse/rhino split at 55 Ma, and WH means whale/hippo split at 52 Ma as used in Waddell et al. (2001). F_{IR} means local rate variability, and MVS (multidimensional vector space) means the model of multidimensional vector space used in Kitazoe et al. (2007). Numbers in bold font indicate nodes where fossil calibrations were integrated into the molecular clock analysis.

Perissodactyla and Carnivora, which together form the clade Zooamata. This hypothesis received support from subsequent analyses of 19 nuclear and 3 mitochondrial genes (Murphy, Eizirik, O'Brien, et al. 2001). In contrast, the present phylogenetic analyses provides strong support for the placement of Perissodactyla as the sister group to Cetartiodactyla (Fig. 2), and the AU test fully rejected alternative topologies (Table 2). Specifically, both BI and ML analyses of first and second codon positions from data sets A and B as well as RY coding of data set B all support a close phylogenetic relationship between Perissodactyla and Cetartiodactyla (Fig. 2). The AU test combined with the first and second codon positions of the data set B rejected a sister-group relationship between Cetartiodactyla and Carnivora (Prasad et al. 2008, Figure 1; Fig. 2) at a significance level of 0.1%, whereas no significant difference was found between two other alternative placements of Perissodactyla, as shown in Fig. 1 (Waddell, Okada, and Hasegawa 1999; Waddell et al. 2001). In addition, when applied to data set A, the AU test rejected a sister-group relationship between Perissodactyla and Carnivora + Pholidota (an arrangement similar to that of Waddell, Okada, and Hasegawa 1999) at 0.05 significance level (Table 2). A recent study (Hou et al. 2009) based on 2705 protein-coding genes (~40 Mb) from the dog, cow, and horse also favored a close affinity between the cow and horse, although they were not able to reject a horse + dog sister-group relationship. Thus, this recent study provides further evidence to support Euungulata (or Cetungulata), which is also our preferred hypothesis.

Although retroposon analyses (Schwartz et al. 2003; Nishihara et al. 2006) discovered many L1 loci that occur in Carnivora and Perissodactyla but not in Cetartiodactyla (cow or/and pig), we consider it premature to conclude that Zooamata and Pegasoferae are monophyletic because it remains unclear whether the L1 loci are present or absent in the orthologs of Pholidota (pangolins) or other important cetartiodactyl lineages, such as whales and dolphins. It is noteworthy that one locus (INT 283) isolated by Nishihara et al. (2006) supports the monophyly of Carnivora, Perissodactyla, and Cetartiodactyla, which is inconsistent with their preferred hypothesis, which includes monophyly of Pegasoferae. As suggested by Shedlock et al. (2004), such inconsistency can result when species diverged over a short evolutionary time span and there is incomplete lineage sorting of ancestral polymorphisms. Considering the very short duration over which the divergence of main laurasiatherian lineages occurred (Table 3), extensive inconsistencies among retroposon insertions are a distinct possibility. If so, then it maybe difficult to use retroposon insertions to infer the phylogeny within Laurasiatheria as more lineages are sampled, similar to the problems encountered in a study on the phylogeny of baleen whales (Nikaido et al. 2006). Many efforts have been made to resolve conflict among retropsoson insertions, presumably as a result of incomplete lineage sorting of polymorphisms. For example, Waddell et al. (2001) have introduced a method that uses likelihood statistics

to test the validity of a superordinal clade whose monophyly is based on retroposon data. However, whether this kind of method is suitable for rapid divergences has not been established.

Based on morphological data, some systematists (e.g., Novacek 1992) have allied Pholidota with Xenarthra (e.g., armadillo and sloth) at the base of the eutherian tree. However, many other systematists have challenged this proposal and instead have advocated a sister-group relationship between Pholidota and Carnivora (Shoshani and McKenna 1998) based on morphological data, with Perissodactyla as the sister group of this clade (Waddell, Okada, and Hasegawa 1999; Murphy, Eizirik, O'Brien, et al. 2001) based on molecular data. Arnason et al. (2002) also suggested that the definition of Cetferungulata (which is in fact a synonym of Fereuungulata) be modified to include Pholidota. However, a topology with Pholidota as basal to other cetferungulates could not be statistically rejected by Arnason et al. (2002). In the present study, we demonstrated strong support for a sister-group relationship between Pholidota and Carnivora, regardless of partitioning strategy or phylogenetic optimality criterion for the 97-, 67-, and 82-gene data sets. This was corroborated with an AU test, which rejected Pholidota as being basal to the remaining fereuungulates at a significance level of 0.1% (Table 2).

Relationships between Major Clades of Cetartiodactyla

Relationships among cetartiodactyl clades have been the focus of numerous phylogenetic studies during the past decade (reviewed in Gatesy et al. 2002). A consensus has been reached on the placement and taxonomic content of some clades, and our results are stronglycongruent with this consensus (Fig. 2; SI Fig. 1a,b). In addition, our analyses provide resolution to some controversial issues in cetartiodactyl relationships.

The paraphyly of conventional Artiodactyla (i.e., terrestrial artiodactyls) is widely supported by recent phylogenetic analyses of molecular data, regardless if supermatrix or supertree methods are employed (Shimamura et al. 1997; Nikaido et al. 1999; Madsen et al. 2001; Murphy, Eizirik, Johnson, et al. 2001; Beck et al. 2006). Although morphological data now weakly support paraphyly of Artiodactyla as well, they continue to support a close relationship between Ruminantia and Camelidae (e.g., Geisler and Uhen 2003). By contrast, molecular data support a sister relationship between Ruminantia and Whippomorpha (Waddell, Okada, and Hasegawa 1999; Cetacea + Hippopotamidae) (Nikaido et al. 1999; Matthee et al. 2001). The present phylogenomic analysis strongly supports cetaceans nested within artiodactyls as follows: (Camelidae, [Suina, (Ruminantia, Whippomorpha)]). This arrangement has also been supported by recent studies that combine morphological and molecular data (O'Leary and Gatesy 2008; Geisler and Theodor 2009; Spaulding et al. 2009). To corroborate the monophyly of Cetartiodactyla, we sequenced two cetartiodactylan exons (NFE2L2 exon 11 and PTPN22 exon 13) and then aligned them with the other laurasiatherian species in our data sets. The alignments indicated that there were several single amino acid insertion/deletion events specific to Cetartiodactyla (SI Fig. 2). In addition, we found single and triple amino acid insertion and deletion events in NPAS4 exon 7 that suggest a close relationship between Whippomorpha and Ruminantia (Ursing and Arnason 1998) (SI Fig. 2).

Phylogenetic placement of Suina and Camelidae is the sole unresolved problem regarding the basal relationships within Cetartiodactyla. Their positions have varied among different molecular studies of mitochondrial and/or nuclear genes (Ursing and Arnason 1998; Gatesy et al. 1999; Nikaido et al. 1999; Arnason et al. 2000; Murphy, Eizirik, Johnson, et al. 2001). Here, as mentioned above, the Camelidae was placed as the basal member of Cetartiodactyla, and an AU test rejected a basal position of Suina at a 5% level of significance. However, our AU test did not reject the basal placement of Suina + Camelidae that was supported by Arnason et al. (2000).

Rapid Diversification of Laurasiatherian Lineages

Three models, that is, explosive, long fuse, and short fuse, have been suggested for placental mammal diversification, with a focus on the timing of the placental mammal radiation relative to the Cretaceous–Paleogene (Tertiary) boundary (Springer et al. 2003). Our study generally supports the long fuse model, placing most interordinal divergences of laurasiatherian mammals in the Cretaceous and intraordinal divergences in the Cenozoic. In particular, the divergence times we estimated using the first and second codon positions of two of our data sets (97- and 60-gene) suggest that interordinal splits of laurasiatherian mammals were concentrated in the Cretaceous, whereas basal cladogenesis within most of the orders, such as Perissodactyla and Cetartiodactyla, was in the Cenozoic (Table 3). Similar conclusions have been reached by several previous studies (Waddell, Okada, and Hasegawa 1999; Cao et al. 2000; Waddell et al. 2001; Douady and Douzery 2003; Springer et al. 2003; Kitazoe et al. 2007; Arnason et al. 2008). Moreover, late Cretaceous zhelestids, which may or may not be near the root of fereuungulates (Archibald 1996; Waddell, Cao, Hauf, and Hasegawa 1999; Wible et al. 2007), combined with the discovery of a possible late Cretaceous ungulate mammal (*Kharmerungulatum vanvaleni*) (Prasad et al. 2007), coupled with endemic perissodactyls (cambaytherids) and artiodactyls (raoellids) in the Early or Middle Eocene (55–50 Ma) rocks of India (Bajpai et al. 2005; Thewissen et al. 2007; Prasad 2009), suggest that ungulate-like characters evolved on multiple occasions and that "ungulates" diversified in the Cretaceous and early Palaeocene. However, as noted elsewhere (Benton and Donoghue 2007), most orders within Laurasiatheria do not have fossil records earlier than Eocene or Paleocene. Although a number of

hypotheses have been advanced to explain this discrepancy, final resolution will only come when phylogenetic methods for converting sequence data to relative time on a clock-like tree get improved as increased accuracy of fossil calibration points (Waddell et al. 2001). And to improve the latter, if the molecular divergence estimates for laurasiatherian mammals are compatible with the long fuse model, intensive paleontological exploration should be conducted on Gondwanan continents, particularly in Africa and India.

Regardless of the absolute divergence time of each clade, our Bayesian relaxed clock analyses suggest a spread of 20 My from the emergence of Laurasiatheria to the emergence of Ferae, which is inconsistent with the short branches in phylogenetic trees (Fig. 2 and Table 3). This discrepancy may be caused by the model used in the BEAST software, which cannot fully adjust for a sharp rate slowdown right after a calibration point, as discussed by Kitazoe et al. (2007). By contrast, branch lengths and our molecular clock analyses do agree that Fereuungulata diversified rapidly, probably over just a few million years (Fig. 2). In addition, most first fossil occurrences of laurasiatherian orders occur in the early Paleocene and early Eocene of Asia and/or North America, which is further evidence for the rapid diversification of fereuungulates. For this reason, the relationships among these clades were difficult to resolve with high BP support in the previous studies, even with the introduction of phylogenomic data. This persistent difficulty in resolving the phylogeny of Fereuungulata could be caused by their genomes being mosaics of conflicting genealogies that resulted from rapid speciation and incomplete lineage sorting (Pollard et al. 2006; Hallström and Janke 2010).

Conclusions

Based on two data sets that sample all orders of Laurasiatheria, lack missing data, and share no genes with previous studies, we were able to resolve most interordinal relationships within Laurasiatheria and determined that the divergences within this clade occurred in the late Cretaceous. Based on our analyses, the first laurasiatherian lineage to diverge was Eulipotyphla, which was then followed by Chiroptera. This sequence thus strongly supports monophyly of Scrotifera and Fereuungulata. Within Fereuungulata, a sister-group relationship between Perissodactyla and Cetartiodactyla was strongly supported by high BP and, PP values, and an AU test rejected alternative topologies. A sister-group relationship between Carnivora and Pholidota was also strongly supported by our data sets; thus, we conclude that Fereuungulata consists of two main subgroups: Euungulata and Ferae. Accordingly, the Pegasoferae hypothesis of Nishihara et al. (2006) was not supported by our analyses. In addition, we found an amino acid insertion or deletion event indicating monophyly of Cetartiodactyla and support for a basal position of Camelidae within this clade, although topology tests could not rule out a basal position of a combined Camelidae + Suina clade. The interordinal laurasiatherian phylogeny presented here should contribute to further development of the higher-level classification of Laurasiatheria and may help elucidate the pace and geographic pattern of its diversification. Future analyses that combine our data directly with morphological data or that use our results as a molecular scaffold may also help place extinct species. However, our molecular dating analyses suggest that the Fereuungulata diversified rapidly, which implies that the relationships within this clade will be difficult to resolve. This potential problem was first noted by Waddell, Okada, and Hasegawa (1999) when he named the clade Laurasiatheria. Although we view our study as a significant advance, additional studies are needed before a consensus on all aspects of laurasiatherian phylogeny emerges.

Supplementary Material

Supplementary material can be found at http://www.sysbio.oxfordjournals.org/.

Funding

This research was financially supported by the National Natural Science Foundation of China (No. 30830016), the Program for New Century Excellent Talents in University, the Ministry of Education of China (No. NCET-07-0445), the major project of the Natural Science Foundation of the Jiangsu Higher Education Institutions of Jiangsu Province, China (07KJA18016), and the Priority Academic Program Development of Jiangsu Higher Education Institutions to G.Y.

Acknowledgments

We thank Professors Ron DeBry, Michael Charleston, Hirohisa Kishino, and an anonymous reviewer for their valuable comments and revision suggestions on our manuscript. Dr. Jonathan Geisler edited versions of this manuscript and provided important scientific comments. We are grateful to Dr. Deborah Ciszek for carefully going through the whole manuscript and providing some valuable comments, particularly on the supplemental materials. Many thanks also to Professors Sibao Wu, Xiaobin Wu, Yi Wu, Peicheng Xiong, Caiquan Zhou, Qisen Yang, Jianping Jiang, Zhengkun Wang, Dr. Yuqing Chen, and Mr. Xinrong Xu for their assistance in sample collection.

References

Adachi J., Hasegawa M. 1996. MOLPHY: version 2.3: programs for molecular phylogenetics based on maximum likelihood. Computer Science Monographs, No. 28. Tokyo (Japan): Institute of Statistical Mathematics.

Archibald J.D. 1996. Fossil evidence for a late Cretaceous origin of "hoofed" mammals. Science. 272:1150–1153.

Archibald J.D. 2003. Timing and biogeography of the eutherian radiation: fossils and molecules compared. Mol. Phylogenet. Evol. 28:350–359.
Arnason U., Adegoke J.A., Bodin K., Born E.W., Esa Y.B., Gullberg A., Nilsson M., Short R.V., Xu X., Janke A. 2002. Mammalian mitogenomic relationships and the root of the eutherian tree. Proc. Natl. Acad. Sci. U.S.A. 99:8151–8156.
Arnason U., Adegoke J.A., Gullberg A., Harley E. H., Janke A., Kullberg M. 2008. Mitogenomic relationships of placental mammals and molecular estimates of their divergences. Gene. 421:37–51.
Arnason U., Gullberg A., Gretarsdottir S., Ursing B., Janke A. 2000. The mitochondrial genome of the sperm whale and a new molecular reference for estimating eutherian divergence dates. J. Mol. Evol. 50:569–578.
Averianov A., Archibald J.D. 2005. Mammals from the mid-Cretaceous Khodzhakul Formation, Kyzylkum Desert, Uzbekistan. Cretaceous. Res. 26:593–608.
Bajpai S., Das D.P., Tiwari B.N. 2005. Early Eocene land mammals from Vastan lignite mine, District Surat (Gujarat), Western India. J. Paleontol. Soc. India. 50:101–113.
Beck R.M., Bininda-Emonds O.R., Cardillo M., Liu F.G., Purvis A. 2006. A higher-level MRP supertree of placental mammals. BMC Evol. Biol. 6:93.
Benton M.J., Donoghue P.C. 2007. Paleontological evidence to date the tree of life. Mol. Biol. Evol. 24:26–53.
Butler P.M. 1988. Phylogeny of the insectivores. In: Benton M.J., editor. The phylogeny and classification of the tetrapods. Oxford: Clarendon. p. 117–141 (Mammals; vol. 2).
Cao Y., Fujiwara M., Nikaido M., Okada N., Hasegawa M. 2000. Interordinal relationships and timescale of eutherian evolution as inferred from mitochondrial genome data. Gene. 259: 149–158.
Castresana J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. 17: 540–552.
Creevey C.J., McInerney J.O. 2005. Clann: investigating phylogenetic information through supertree analyses. Bioinformatics. 21: 390–392.
Criscuolo A., Berry V., Douzery E.J.P., Gascuel O. 2006. SDM: a fast distance-based approach for (super)tree building in phylogenomics. Syst. Biol. 55:740–755.
Delsuc F., Brinkmann H., Philippe H. 2005. Phylogenomics and the reconstruction of the tree of life. Nat. Rev. Genet. 6:361–375.
Douady C.J., Douzery E.J. 2003. Molecular estimation of eulipotyphlan divergence times and the evolution of "Insectivora". Mol. Phylogenet. Evol. 28:285–296.
Drummond A.J., Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7:214.
Edgar R.C. 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics. 5:113.
Gatesy J., Matthee C., DeSalle R., Hayashi C. 2002. Resolution of a supertree/supermatrix paradox. Syst. Biol. 51:652–664.
Gatesy J., Milinkovitch M., Waddell V., Stanhope M. 1999. Stability of cladistic relationships between Cetacea and higher-level artiodactyl taxa. Syst. Biol. 48:6–20.
Geisler J.H., Theodor J.M. 2009. Hippopotamus and whale phylogeny. Nature. 458:E1–E4.
Geisler J.H., Uhen M.D. 2003. Morphological support for a close relationship between hippos and whales. J. Vertebr. Paleontol. 23:991–996.
Gregory W.K. 1910. The orders of mammals. Bull. Am. Mus. Nat. Hist. 27:1–524.
Guindon S., Gascuel O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. 52:696–704.
Gunnell G.F., Simmons N.B. 2005. Fossil evidence and the origin of bats. J. Mol. Evol. 12:209–246.
Hallström B.M., Janke A. 2010. Mammalian evolution may not be strictly bifurcating. Mol. Biol. Evol. 27:2804–2816.
Hanson-Smith V., Kolaczkowski B., St. John J., Thornton J. W. 2009. Phyesta. Program distributed by the author. Oregon (OR): The Institute of Ecology and Evolution, University of Oregon.
Hou Z.C., Romero R., Wildman D.E. 2009. Phylogeny of the Fereuungulata (Mammalia: Laurasiatheria) as determined from phylogenomic data. Mol. Phylogenet. Evol. 52:660–664.
Huelsenbeck J., Rannala B. 2004. Frequentist properties of Bayesian posterior probabilities of phylogenetic trees under simple and complex substitution models. Syst. Biol. 53:904–913.
Huelsenbeck J.P., Ronquist F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics. 17:754–755.
Irwin D.M., Wilson A.C. 1993. Limitations of molecular methods for establishing the phylogeny of mammals, with special reference to the position of elephants. In: Szalay F.S., Novacek M.J., McKenna M.C., editors. Mammal phylogeny: placentals. New York: Springer-Verlag. p. 257–267.
Kishino H., Hasegawa M. 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. J. Mol. Evol. 29: 170–179.
Kitazoe Y., Kishino H., Waddell P.J., Nakajima N., Okabayashi T., Watabe T., Okuhara Y. 2007. Robust time estimation reconciles views of the antiquity of placental mammals. PLoS One. 2:e384.
Kolaczkowski B., Thornton J.W. 2008. A mixed branch length model of heterotachy improves phylogenetic accuracy. Mol. Biol. Evol. 25:1054–1066.
Kostka M., Uzlikova M., Cepicka I., Flegr J. 2008. SlowFaster, a user-friendly program for slow-fast analysis and its application on phylogeny of Blastocystis. BMC Bioinformatics. 9:341.
Kriegs J.O., Churakov G., Kiefmann M., Jordan U., Brosius J., Schmitz J. 2006. Retroposed elements as archives for the evolutionary history of placental mammals. PLoS Biol. 4:e91.
Kullberg M., Nilsson M.A., Arnason U., Harley E.H., Janke A. 2006. Housekeeping genes for phylogenetic analysis of eutherian relationships. Mol. Biol. Evol. 23:1493–1503.
Leigh J.W., Susko E., Baumgartner M., Roger A.J. 2008. Testing congruence in phylogenomic analysis. Syst. Biol. 57:104–115.
MacPhee R., Novacek M. 1993. Definition and relationships of Lipotyphla. In: Szalay F.S., Novacek M.J., McKenna M.C., editors. New York: Springer. p. 13–31.
Madsen O., Scally M., Douady C.J., Kao D.J., DeBry R.W., Adkins R., Amrine H.M., Stanhope M.J., de Jong W.W., Springer M.S. 2001. Parallel adaptive radiations in two major clades of placental mammals. Nature. 409:610–614.
Matthee C.A., Burzlaff J.D., Taylor J.F., Davis S.K. 2001. Mining the mammalian genome for artiodactyl systematics. Syst. Biol. 50: 367–390.
McKenna M.C. 1975. Toward a phylogenetic classification of the Mammalia. In: Luckett W.P., Szalay F.S., editors. Phylogeny of the primates: an interdisciplinary approach. New York: Plenum. p. 21–46.
Miyamoto M.M. 1996. A congruence study of molecular and morphological data for eutherian mammals. Mol. Phylogenet. Evol. 6: 373–390.
Montgelard C., Catzeflis F.M., Douzery E. 1997. Phylogenetic relationships of artiodactyls and cetaceans as deduced from the comparison of cytochrome b and 12S rRNA mitochondrial sequences. Mol. Biol. Evol. 14:550–559.
Mouchaty S.K., Gullberg A., Janke A., Arnason U. 2000. The phylogenetic position of the Talpidae within eutheria based on analysis of complete mitochondrial sequences. Mol. Biol. Evol. 17:60–67.
Murphy W.J., Eizirik E., Johnson W.E., Zhang Y.P., Ryder O.A., O'Brien S.J. 2001. Molecular phylogenetics and the origins of placental mammals. Nature. 409:614–618.
Murphy W.J., Eizirik E., O'Brien S.J., Madsen O., Scally M., Douady C.J., Teeling E., Ryder O.A., Stanhope M.J., de Jong W.W., Springer M.S. 2001. Resolution of the early placental mammal radiation using Bayesian phylogenetics. Science. 294:2348–2351.
Narita Y., Oda S., Takenaka O., Kageyama T. 2001. Phylogenetic position of Eulipotyphla inferred from the cDNA sequences of pepsinogens A and C. Mol. Phylogenet. Evol. 21:32–42.
Nikaido M., Hamilton H., Makino H., Sasaki T., Takahashi K., Goto M., Kanda N., Pastene L.A., Okada N. 2006. Proceedings of the SMBE Tri-National Young Investigators' Workshop. 2005. Baleen whale phylogeny and a past extensive radiation event revealed by SINE insertion analysis. Mol. Biol. Evol. 23:866–873.

Nikaido M., Harada M., Cao Y., Hasegawa M., Okada N. 2000. Monophyletic origin of the order chiroptera and its phylogenetic position among mammalia, as inferred from the complete sequence of the mitochondrial DNA of a Japanese megabat, the Ryukyu flying fox (Pteropus dasymallus). J. Mol. Evol. 51:318–328.
Nikaido M., Kawai K., Cao Y., Harada M., Tomita S., Okada N., Hasegawa M. 2001. Maximum likelihood analysis of the complete mitochondrial genomes of eutherians and a reevaluation of the phylogeny of bats and insectivores. J. Mol. Evol. 53: 508–516.
Nikaido M., Rooney A.P., Okada N. 1999. Phylogenetic relationships among cetartiodactyls based on insertions of short and long interpersed elements: hippopotamuses are the closest extant relatives of whales. Proc. Natl. Acad. Sci. U.S.A. 96:10261–10266.
Nishihara H., Hasegawa M., Okada N. 2006. Pegasoferae, an unexpected mammalian clade revealed by tracking ancient retroposon insertions. Proc. Natl. Acad. Sci. U.S.A. 103:9929–9934.
Novacek M.J. 1992. Mammalian phylogeny: shaking the tree. Nature. 356:121–125.
Nylander J.A. 2004. MrModeltest v2. Program distributed by the author. Uppsala (Sweden): Evolutionary Biology Centre, Uppsala University.
Nylander J.A., Wilgenbusch J.C., Warren D.L., Swofford D.L. 2008. AWTY (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics. Bioinformatics. 24:581–583.
O'Leary M.A., Gatesy J. 2008. Impact of increased character sampling on the phylogeny of Cetartiodactyla (mammalia): combined analysis including fossils. Cladistics. 24:397–442.
Onuma M., Kusakabe T., Kusakabe S. 1998. Phylogenetic positions of insectivora in eutheria inferred from mitochondrial cytochrome c oxidase subunit II gene. Zool. Sci. 15:139–145.
Pettigrew J. D. 1986. Flying primates? Megabats have the advanced pathway from eye to midbrain. Science. 231:1304–1306.
Phillips M.J., Delsuc F., Penny D. 2004. Genome-scale phylogeny and the detection of systematic biases. Mol. Biol. Evol. 21: 1455–1458.
Pollard D.A., Iyer V.N., Moses A.M., Eisen M.B. 2006. Widespread discordance of gene trees with species tree in Drosophila: evidence for incomplete lineage sorting. PLoS Genet. 2:e173.
Prasad A.B., Allard M.W., Green E.D. 2008. Confirming the phylogeny of mammals by use of large comparative sequence data sets. Mol. Biol. Evol. 25:1795–1808.
Prasad G.V. 2009. Divergence time estimates of mammals from molecular clocks and fossils: relevance of new fossil finds from India. J. Biosci. 34:649–659.
Prasad G.V., Verma O., Sahni A., Parmar V., Khosla A. 2007. A Cretaceous hoofed mammal from India. Science. 318:937.
Pumo D.E., Finamore P.S., Franek W.R., Phillips C.J., Tarzami S., Balzarano D. 1998. Complete mitochondrial genome of a neotropical fruit bat, Artibeus jamaicensis, and a new hypothesis of the relationships of bats to other eutherian mammals. J. Mol. Evol. 47:709–717.
Quang L.S., Gascuel O., Lartillot N. 2008. Empirical profile mixture models for phylogenetic reconstruction. Bioinformatics. 24: 2317–2323.
Ranwez V., Delsuc F., Ranwez S., Belkhir K., Tilak M.K., Douzery E.J. 2007. OrthoMaM: a database of orthologous genomic markers for placental mammal phylogenetics. BMC Evol. Biol. 7:241.
Scally M., Madsen O., Douady C.J., de Jong W.W., Stanhope M.J., Springer M.S. 2001. Molecular evidence for the major clades of placental mammals. J. Mamm. Evol. 8:239–277.
Schwartz S., Elnitski L., Li M., Weirauch M., Riemer C., Smit A., Green E.D., Hardison R.C., Miller W. 2003. MultiPipMaker and supporting tools: alignments and analysis of multiple genomic DNA sequences. Nucleic Acids Res. 31:3518–3524.
Shedlock A.M., Takahashi K., Okada N. 2004. SINEs of speciation: tracking lineages with retroposons. Trends Ecol. Evol. 19: 545–553.
Shimamura A.M., Yasue H., Ohshima K., Abe H., Kato H., Kishiro T., Goto M., Munechika I., Okada N. 1997. Molecular evidence from retroposons that whales form a clade within even-toed ungulates. Nature. 388:666–670.
Shimodaira H. 2002. An approximately unbiased test of phylogenetic tree selection. Syst. Biol. 51:492–508.
Shimodaira H., Hasegawa M. 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol. Biol. Evol. 16:1114–1116.
Shimodaira H., Hasegawa M. 2001. CONSEL: for assessing the confidence of phylogenetic tree selection. Bioinformatics. 17: 1246–1247.
Shoshani J., McKenna M.C. 1998. Higher taxonomic relationships among extant mammals based on morphology, with selected comparisons of results from molecular data. Mol. Phylogenet. Evol. 9:572–584.
Spaulding M., O'Leary M.A., Gatesy J. 2009. Relationships of Cetacea (Artiodactyla) among mammals: increased taxon sampling alters interpretations of key fossils and character evolution. PLoS One. 4:e7062.
Springer M.S., Burk-Herrick A., Meredith R., Eizirik E., Teeling E., O'Brien S.J., Murphy W.J. 2007. The adequacy of morphology for reconstructing the early history of placental mammals. Syst. Biol. 56:673–684.
Springer M.S., Murphy W.J., Eizirik E., O'Brien S.J. 2003. Placental mammal diversification and the Cretaceous-Tertiary boundary. Proc. Natl. Acad. Sci. U.S.A. 100:1056–1061.
Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics. 22:2688–2690.
Stamatakis A., Hoover P., Rougemont J. 2008. A rapid bootstrap algorithm for the RAxML web servers. Syst. Biol. 57:758–771.
Stanhope M.J., Waddell V.G., Madsen O., de Jong W., Hedges C.B., Cleven G.C., Kao D., Springer S.M. 1998. Molecular evidence for multiple origins of Insectivora and for a new order of endemic African insectivore mammals. Proc. Natl. Acad. Sci. U.S.A. 95:9967–9972.
Thewissen J.G., Cooper L.N., Clementz M.T., Bajpai S., Tiwari B.N. 2007. Whales originated from aquatic artiodactyls in the Eocene epoch of India. Nature. 450:1190–1194.
Thewissen J.G., Williams E.M., Roe L.J., Hussain S.T. 2001. Skeletons of terrestrial cetaceans and the relationship of whales to artiodactyls. Nature. 413:277–281.
Thompson J.D., Gibson T.J., Plewniak F., Jeanmougin F., Higgins D.G. 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25:4876–4882.
Ursing B.M., Arnason U. 1998. Analyses of mitochondrial genomes strongly support a hippopotamus-whale clade. Proc. Biol. Sci. 265:2251–2255.
Van den Bussche R.A., Hoofer S.R. 2001. Evaluating monophyly of Nataloida (Chiroptera) with mitochondrial DNA sequences. J. Mammal. 82:320–327.
Waddell P.J. 2008. Fit of fossils and mammalian molecular trees: dating inconsistencies revisited. arXiv 0812.5114.
Waddell P.J., Cao Y., Hasegawa M., Mindell D.P. 1999. Assessing the Cretaceous superordinal divergence times within birds and placental mammals by using whole mitochondrial protein sequences and an extended statistical framework. Syst. Biol. 48:119–137.
Waddell P.J., Cao Y., Hauf J., Hasegawa M. 1999. Using novel phylogenetic methods to evaluate mammalian mtDNA, including AA invariant sites-LogDet plus site stripping, to detect internal conflicts in the data, with special reference to the position of hedgehog, armadillo, and elephant. Syst. Biol. 48:31–53.
Waddell P.J., Kishino H., Ota R. 2001. A phylogenetic foundation for comparative mammalian genomics. Genome Inform. 12: 141–154.
Waddell P.J., Okada N., Hasegawa M. 1999. Towards resolving the interordinal relationships of placental mammals. Syst. Biol. 48:1–5.
Waddell P.J., Ota R., Penny D. 2009. Measuring fit of sequence data to phylogenetic model: gain of power using marginal tests. J. Mol. Evol. 69:289–299.
Waddell P.J., Shelley S. 2003. Evaluating placental inter-ordinal phylogenies with novel sequences including RAG1, gamma-fibrinogen, ND6, and mt-tRNA, plus MCMC-driven nucleotide, amino acid, and codon models. Mol. Phylogenet. Evol. 28: 197–224.

Waddell P.J., Steel M.A. 1997. General time-reversible distances with unequal rates across sites: mixing gamma and inverse Gaussian distributions with invariant sites. Mol. Phylogenet. Evol. 8:398–414.

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

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

Zhou X., Xu S., Zhang P., Yang G. 2011. Developing a series of conservative anchor markers and their application to phylogenomics of Laurasiatherian mammals. Mol. Ecol. Resour. 11:134–140.