

MOLECULAR PHYLOGENY OF LIVING XENARTHANS AND THE IMPACT OF CHARACTER AND TAXON SAMPLING ON THE PLACENTAL TREE ROOTING.

Frédéric Delsuc,* Mark Scally,†‡ Ole Madsen,§ Michael J. Stanhope,†|| Wilfried W. de Jong,§¶
François M. Catzeflis,* Mark S. Springer,‡ and Emmanuel J. P. Douzery*

*Laboratoire de Paléontologie, Paléobiologie et Phylogénie, Institut des Sciences de l'Evolution,
Université Montpellier II, Montpellier, France

†Queen's University of Belfast, Biology and Biochemistry, 97 Lisburn Road, Belfast, BT9 7BL,
United Kingdom

‡Department of Biology, University of California, Riverside, California 92521, USA

§Department of Biochemistry, University of Nijmegen, PO Box 9101, 6500 HB Nijmegen, The
Netherlands

||Bioinformatics, GlaxoSmithKline, 1250 South Collegeville Road, UP1345, Collegeville
Pennsylvania 19426, USA

¶Institute for Biodiversity and Ecosystems Dynamics, University of Amsterdam, The
Netherlands

Author for correspondence and reprints:

Emmanuel J. P. Douzery

Laboratoire de Paléontologie, Paléobiologie et Phylogénie – CC064

Institut des Sciences de l'Evolution UMR 5554 / CNRS, Université Montpellier II

Place Eugène Bataillon, 34 095 MONTPELLIER Cedex 05 – France

Tel.: 33 4 67 14 48 63 / FAX: 33 4 67 14 36 10

E-mail: douzery@isem.univ-montp2.fr

Running title: Molecular phylogeny of living xenarthrans

ABSTRACT

Extant xenarthrans (armadillos, anteaters and sloths) are among the most derived placental mammals ever evolved. South America was the cradle of their evolutionary history. During the Tertiary, xenarthrans experienced an extraordinary radiation while South America remained isolated from other continents. The 13 living genera are relicts of this earlier diversification and represent one of the four major clades of placental mammals. Sequences of the three independent protein-coding nuclear markers $\alpha 2B$ adrenergic receptor (ADRA2B), breast cancer susceptibility (BRCA1), and von Willebrand Factor (VWF) were determined for 12 of the 13 living xenarthran genera. Comparative evolutionary dynamics of these nuclear exons using a likelihood framework revealed contrasting patterns of molecular evolution. All codon positions of BRCA1 were shown to evolve in a strikingly similar manner, and third codon positions appeared less saturated within placentals than those of ADRA2B and VWF. Maximum likelihood and Bayesian phylogenetic analyses of a 47 placental data set rooted by three marsupial outgroups resolved the phylogeny of Xenarthra with some evidence for two radiation events in armadillos and provided a strongly supported picture of placental interordinal relationships. This topology was fully compatible with recent studies dividing placentals into the Southern Hemisphere clades Afrotheria and Xenarthra and a monophyletic Northern Hemisphere clade (Boreoeutheria) composed of Laurasiatheria and Euarchontoglires. Partitioned likelihood statistical tests of the position of the root, under different character partition schemes, identified three almost equally likely hypotheses for early placental divergences: a basal Afrotheria, an Afrotheria + Xenarthra clade or a basal Xenarthra (Epitheria hypothesis). We took advantage of the extensive sampling realized within Xenarthra to assess its impact on the location of the root on the placental tree. By resampling taxa within Xenarthra, the conservative Shimodaira-Hasegawa likelihood based test of alternative topologies was shown to be sensitive to both character and taxon sampling.

Keywords: Xenarthran phylogeny – Taxon and character sampling – Placentals – Evolutionary dynamics – Nuclear markers – Maximum Likelihood and Bayesian phylogenetics.

INTRODUCTION

Living xenarthrans are represented by three morphologically distinct lineages: armored armadillos, toothless anteaters, and phyllophagous tree-sloths. The 30 living species of the order Xenarthra (Wetzel 1985; Vizcaíno 1995) are relics of an impressive South-American Tertiary radiation that gave birth to giant extinct forms like glyptodonts and ground-dwelling sloths (Patterson and Pascual 1972). The monophyly of the order is strongly supported by numerous morphological (Engelmann 1985; Patterson, Segall, and Turnbull 1989; Patterson et al. 1992; Rose and Emry 1993; Gaudin 1999) and molecular (de Jong et al. 1985; van Dijk et al. 1999; Delsuc et al. 2001) synapomorphies. Morphological (Engelmann 1985; Patterson et al. 1992) and molecular (Delsuc et al. 2001; Madsen et al. 2001; Murphy et al. 2001a) studies unambiguously supported the division of Xenarthra into two suborders: Cingulata represented by armadillos (Dasypodidae), and Pilosa represented by anteaters (Vermilingua: Myrmecophagidae) plus sloths (Folivora: Megalonychidae and Bradypodidae). However, whereas the phylogeny of living pilosans is well resolved (Gaudin and Branham 1998; Delsuc et al. 2001; Greenwood et al. 2001), this is not the case for cingulates. Armadillos are the most diversified xenarthrans and comprise 21 extant species classified in eight genera (Wetzel 1985; Vizcaíno 1995). This ecologically and morphologically diverse group contains enigmatic taxa like the dwarf subterranean pink-fairy armadillos (genus *Chlamyphorus*) and the endangered giant armadillo (*Priodontes maximus*). Some striking adaptations can be found in armadillos. Three-banded armadillos (*Tolypeutes*) developed a unique anti-predator strategy by entirely rolling up into a ball, and nine-banded armadillos (*Dasypus*) are the only vertebrates to reproduce by obligate monozygotic polyembryony (Loughry et al. 1998). Notwithstanding these fascinating peculiarities, phylogenetic studies on armadillos are still scarce (Guth 1961; Engelmann 1985; Patterson, Segall, and Turnbull 1989; Cetica et al. 1998; Delsuc et al. 2001) and relationships among the eight living genera remain unclear.

Solving the phylogenetic position of the order Xenarthra within Mammalia is of primary importance to understand the morphological and biogeographical processes that shaped the early stages of placental evolution. Indeed, despite their highly specialized morphology, xenarthrans retain anatomical and physiological characters thought to be primitive for placental mammals (McKenna 1975). This composite morphology, mixing ancestral and derived characters, has made the order's position within placentals very difficult to assess (Engelmann 1985; Gaudin et al. 1996). Based on the retention of numerous "archaic" features, morphologists (Gregory 1910; McKenna 1975; Novacek 1992; Shoshani and McKenna 1998) have long proposed that Xenarthra represents the sister-group to all other eutherians thereby named Epitheria. However, morphological synapomorphies defining epitherians are weak and their phylogenetic distribution is equivocal (Gaudin et al. 1996). Early studies of complete mitochondrial genomes did not support a basal position for Xenarthra (Arnason, Gullberg, and Janke 1997). More recent analyses—including a larger taxon sampling—suggested a sister-group relationship between the nine-banded armadillo (*Dasypus novemcinctus*) and representatives of the African clade (Waddell et al. 1999; Cao et al. 2000; Mouchaty et al. 2000a). However, complete mitochondrial genome analyses appear to be affected by insufficient taxon sampling—especially within Xenarthra—likely responsible for long-branch attraction and rooting artifacts (Waddell et al. 1999), and suffer from saturation in the deepest parts of the tree (Cao et al. 2000; Springer et al. 2001).

By contrast two recent independent analyses based on the large concatenation of mainly nuclear genes for a broad taxon sampling of eutherian mammals have shown that the order Xenarthra on its own represents one of the major clades of placentals (Madsen et al. 2001; Murphy et al. 2001a, b). These studies provided convincing evidence for an arrangement of placental orders into four major clades: I) Afrotheria, II) Xenarthra, III) Euarchontoglires and IV) Laurasiatheria, emphasizing the crucial influence of tectonic events in their early differentiation. A close relationship between the Northern Hemisphere clades III and IV has been proposed (i. e.

"Boreoeutheria" ; Springer and de Jong 2001) with either Afrotheria or Xenarthra as the most basal clade, suggesting a Southern Hemisphere origin for eutherian mammals (Eizirik, Murphy, and O'Brien 2001; Madsen et al. 2001). In fact, the relationships between these four clades directly depend on the unstable position of the root. Recent application of the Bayesian approach (Yang and Rannala 1997; Huelsenbeck et al. 2001) to this problem using a large character sampling supported a basal position of Afrotheria (Murphy et al. 2001b). However, in this study, Xenarthra suffers from poor taxonomic representation (only three taxa) relative to the other three major clades that include eight (Afrotheria), 11 (Euarchontoglires), and 20 (Laurasiatheria) taxa. Here we present a study including the broadest taxonomic representation so far considered in a molecular approach to xenarthran phylogeny. We constructed a supermatrix of 47 placental taxa and three marsupial outgroups, including 12 of the 13 living xenarthran genera, for three genetically independent protein-coding nuclear genes: $\alpha 2B$ Adrenergic receptor (ADRA2B), Breast Cancer Susceptibility exon 11 (BRCA1), and von Willebrand Factor exon 28 (VWF) representing a total of 5130 aligned nucleotide sites. Choice of these nuclear markers was guided by their wide use for inferring the phylogeny of placental mammals (Springer et al. 1997; Stanhope et al. 1998a, b; Madsen et al. 2001) and their resolving power demonstrated to be higher than for mitochondrial markers at the mammalian inter-ordinal level (Springer et al. 1999; Springer et al. 2001). Furthermore, their coding status provides the opportunity to compare the phylogenetic signal contained in both nucleotides and amino acids. The extensive sampling realized within Xenarthra permits us to resolve intraordinal relationships and to investigate the impact of an increased taxon sampling on the root position of the placental tree. Increased nucleotide sampling has a dominant impact on phylogenetic accuracy (Poe and Swofford 2000; Rosenberg and Kumar 2001), but increased taxon sampling has also been shown to facilitate phylogenetic inference (Lecointre et al. 1993; Hillis 1996; Rannala et al. 1998). Indeed, breaking potentially long branches by adding taxa within one of the two most basal clades might help to stabilize the placental ingroup topology in its deepest parts. By the use of a maximum likelihood

and Bayesian framework this study aims to: i) compare the evolutionary dynamics and phylogenetic content of these three protein-coding nuclear markers evolving under different selective pressures, ii) resolve the phylogeny of living xenarthrans with special reference to armadillos, and iii) evaluate the effect of an increased taxon sampling within Xenarthra for finding the root of the placental tree.

MATERIAL & METHODS

Taxonomic sampling

Thirteen xenarthran species representing all living genera—except the rare and cryptic subterranean genus *Chlamyphorus*—were sampled. We chose the data set encompassing all representatives of placentals sequenced to date for the three nuclear coding genes ADRA2B, BRCA1 and VWF. Three marsupials (*Macropus*, *Didelphis* and *Vombatus*) were used as outgroups to locate the root of the placental tree (see Supplementary material).

Data acquisition

Xenarthran samples preserved in 95% ethanol were stored in the mammalian tissue collection of the Institut des Sciences de l'Evolution de Montpellier (Catzefflis 1991). Total DNAs were extracted for *Dasypus novemcinctus* (nine-banded armadillo), *Dasypus kappleri* (great long-nosed armadillo), *Chaetophractus villosus* (larger hairy armadillo), *Euphractus sexcinctus* (six-banded armadillo), *Zaedyus pichi* (pichi), *Tolypeutes matacus* (southern three-banded armadillo), *Cabassous unicinctus* (southern naked-tailed armadillo), *Priodontes maximus* (giant armadillo), *Cyclopes didactylus* (pygmy anteater), *Tamandua tetradactyla* (collared anteater), *Myrmecophaga tridactyla* (giant anteater), *Bradypus tridactylus* (pale-throated three-toed sloth), and *Choloepus didactylus* (southern two-toed sloth).

The single exon gene ADRA2B was amplified and sequenced using the primers designed by Springer et al. (1997) and the following additional pairs of forward (A) and reverse (B) primers: A4 (5'-GCCATCGCGGCNGYCRYCACCTTCCTCATC-3'), B4 (5'-GCTGCGYTTGGCAATCAGGTAGAGTCG-3') and B5 (5'-GCGCCCAGGCTGTAGCTGAAGAAGAA-3'). The exon 28 of VWF was amplified according to Delsuc et al. (2001). PCR products for ADRA2B and VWF were purified from 1% agarose gels using Amicon Ultrafree-DA columns (Millipore), and sequenced on both strands using automatic sequencing (Big Dye Terminator cycle kit) on an ABI 310 (PE Applied Biosystems). BRCA1 gene sequences were obtained as described elsewhere (Teeling et al. 2000; Madsen et al. 2001). The 2900 bp region of exon 11 was amplified in three overlapping segments each around 1000 bp in size as described in Teeling et al. (2000) and Madsen et al. (2001). PCR primers for each segment were as follows: [1] Forward UF1 (5'-GTTTCAAACCTTGCATGTGGAGCC-3'), Reverse XR11 (5'-GCAGATTCTTTTCCAATGATTCTG-3'); [2] Forward GF8 (5'-GGCCTTCATCCTGAGGATTTTATCAA-3'), Reverse R19 (5'-TG YAAATACTGAGTATCAAGTTCAC-3'); [3] Forward XBF17 (5'-TATGGCACTCARGAYAGTATCTCATT-3'), Reverse BRCA1B (5'-GTTGGAAGCAGGGAAGCTCTTCATC-3'). Sequencing was performed using PCR primers and additional internal primers in segment [1] Forward F4 (5'-GAAAGTTAATGAGTGGTTTTCCAGAA-3'), Reverse UR7 (5'-CTTCCTCCGATAGGTTTTCCCAA-3') and in segment [3] Forward F25 (5'-AACTAGGTAGAAACAGAGGRCCTA-3'), Reverse R26 (5'-TTAGGYCCTCTGTTTCTACCTAGTT-3').

The 26 xenarthran sequences new to this study have been deposited in the EMBL data bank. Taxonomy and accession numbers referring to all sequences used in this study are provided as Supplementary Material.

Sequence alignment

Sequences were manually aligned with the ED editor of the MUST package (Philippe 1993). We excluded a glutamic acid repeat region of ADRA2B and also a 21 bp region of BRCA1 that is repeated up to four times according to Madsen et al. (2001). The concatenated placental data set is 5130 bp long with ADRA2B (1152 bp), BRCA1 (2788 bp) and VWF (1190 bp). All introduced gaps were treated as missing data in subsequent analyses. Alignments are available upon request.

Phylogenetic analyses

All phylogenetic analyses were conducted under the Maximum Likelihood (ML) and Bayesian approaches. ML was chosen because it is known to be less sensitive to potential long-branch attraction artifacts and it takes into account the underlying molecular evolutionary process (Swofford et al. 2001). The Bayesian approach allows the analysis of large phylogenetic data sets under complex evolutionary models (Huelsenbeck et al. 2001). Using a likelihood framework also permits the statistical comparison of competing hypotheses and topologies obtained from different partition and rooting schemes (Huelsenbeck and Rannala 1997; Whelan, Liò, and Goldman 2001).

Results from Modeltest 3.06 (Posada and Crandall 1998) based on the Akaike Information Criterion (AIC) indicate that the General Time Reversible model (GTR, Yang 1994) plus a gamma (Γ) distribution of parameter α (Yang 1996a), and a proportion I of invariable sites was the best model for each of the three data sets. However, to render ML analyses comparable between PAUP* version 4.0b8 (Swofford 1998) and PAML version 3.0c (Yang 1997) which does not allow the use of invariable sites, we choose the GTR + Γ_8 model with eight discrete rate categories to represent the continuous Γ distribution.

In order to avoid excessive computation time, ML analyses with PAUP* version 4.0b8 (Swofford 1998) were conducted using a loop approach to estimate the best tree and the optimal likelihood parameters. First, substitution rate matrix and among site substitution rate heterogeneity parameters were optimized on a Neighbor-Joining (NJ) topology derived from ML distances obtained using a GTR + Γ_8 model ($\alpha = 0.5$). Second, a ML heuristic search was conducted by Tree Bisection Reconnection (TBR) branch swapping to identify the optimal tree under these GTR + Γ_8 parameter estimates. Third, likelihood parameters were re-estimated on this new topology. Four, a new heuristic tree search was run under the re-estimated GTR + Γ_8 parameters. This loop procedure was performed until stabilization of both topology and parameters was attained (after three cycles). The Bayesian approach to phylogenetic reconstruction (Yang and Rannala 1997; Huelsenbeck et al. 2001) was implemented using MrBayes 2.01 (Huelsenbeck and Ronquist 2001). Metropolis-coupled Markov chain Monte Carlo (MCMCMC) sampling was performed with four chains that were run for 500 000 generations, using default model parameters as starting values.

Robustness of nucleotide-derived trees was estimated by Bootstrap Percentages (BP) (Felsenstein 1985) computed by PAUP* using the optimal ML estimated parameters, with NJ starting trees and TBR branch swapping. The number of TBR rearrangements was unlimited for the Xenarthra data set (1000 bootstrap replications), and set to 1000 per replicate for the Placental data set (100 bootstrap replications). Bayesian posterior probabilities were picked from the 50% majority rule consensus of trees sampled every 20 generations after removing trees obtained before chains reached apparent stationarity ("burn-in" determined by empirical checking of likelihood values).

Evaluation of the saturation of nucleotide substitutions

The nucleotide substitution saturation of the phylogenetic markers was evaluated with the graphical method used by Philippe and Forterre (1999). The inferred number of substitutions between each pair of sequences was estimated from the ML tree as the sum of the lengths of all branches linking these two sequences. The saturation level was estimated by plotting the number of observed differences as a function of the ML inferred number of substitutions for all 1225 pairwise comparisons for 50 sequences. In these graphics, the straight line ($Y = X$) represents the situation for which the number of inferred substitutions equals the number of observed differences, i.e., there is no detected homoplasy in the data. The nucleotide substitution saturation is evidenced when the number of inferred substitutions increased, whereas the number of observed differences remained constant (plateau shape).

Statistical tests of the position of the root: impact of character partitions and taxon sampling

Extensive studies (Madsen et al. 2001; Murphy et al. 2001a, b; this paper) identified four major clades of placentals: Afrotheria, Xenarthra, Euarchontoglires, and Laurasiatheria. Assuming the respective monophyly of these clades—as suggested by the ML analyses of the data—there are 15 possible ways to connect the four clades into a bifurcating topology rooted by marsupials. These 15 alternative topologies were compared by the ML test of Kishino and Hasegawa (1989), with introduction of the Shimodaira and Hasegawa (1999) correction for multiple comparisons of topologies defined *a posteriori* (SH test).

To evaluate the impact of character sampling on the location of the root, we performed SH tests using different character partitions of the combined data set. As ADRA2B, BRCA1, and VWF locate to different chromosomes and their protein products markedly differ in function, these nuclear genes likely evolve under differential selective pressures. To accommodate expected differences in molecular evolution, we followed the approach of Yang (1996b) for analyzing

multiple genes. Thus, each codon position of the three genes has been attributed one independent GTR + Γ_8 model for which the different sets of likelihood parameters (base frequencies, substitution rate parameters, the α parameter of the Γ distribution, and branch lengths) were estimated. This allows us to take into account both differences in substitution pattern and rate heterogeneity between both the three codon positions and the three different genes. This partitioned ML model was used in PAML to compute log-likelihoods and confidence probability values for the 15 topologies corresponding to the possible locations of the root between the four major clades of placentals. The four following sets of characters were analyzed: (1) first and second codon positions of the three genes (ABV 1+2; 3421 sites; 6 partitions), (2) first and second codon positions of ADRA2B and VWF plus all codon positions of BRCA1 (A2BV2; 4350 sites; 7 partitions), (3) all codon positions of the three genes (ABV 1+2+3; 5130 sites; 9 partitions), and (4) concatenated amino acids of the three proteins (ABV AA; 1709 sites). Because of computational time constraints, a single JTT (Jones, Taylor and Thornton 1992) plus Γ_8 model was assumed for the amino acid concatenated data set.

To assess the impact of increased taxon sampling within Xenarthra on the location of the root, we conducted SH tests after resampling of xenarthran species belonging to its three subgroups: sloths (2 taxa), anteaters (3), and armadillos (8). To do this, we built all taxon combinations with each one of the 13 available xenarthrans (13 combinations), with one from all three xenarthran subclades (48 combinations of three taxa), with two xenarthrans per subclades (84 combinations of six taxa), with three (except for sloths) xenarthrans per subclade (56 combinations of 8 taxa), and with all xenarthrans (the original combination including 13 taxa). Because of computational time limitations, PAUP* was used to perform SH tests assuming a single GTR + Γ_8 model on the different concatenated data sets.

RESULTS

Molecular evolution of the three nuclear markers

Base compositions

The three nuclear genes exhibit marked differences in base composition. ADRA2B and VWF appear quite similar in being GC-rich with mean overall values of 64.9% and 61.9%, respectively. To the contrary, BRCA1 is rather AT rich with a mean overall value of 59.6%. As expected, these differences in base composition are especially contrasted on third codon positions. Indeed, ADRA2B and VWF show mean GC3 values of more than 80% while this value is only 30% for BRCA1. This GC3 enrichment is particularly pronounced in xenarthrans that possess values of more than 90% for ADRA2B and VWF. This phenomenon implies that the ML base composition stationarity assumption is satisfied for BRCA1, but not for several taxa when considering all codon positions of ADRA2B and VWF ($P < 0.05$; χ^2 test of TREE-PUZZLE 4.02; Strimmer and von Haeseler 1996). As expected from their particularly high GC3 content, almost all xenarthran taxa deviated from the mean base composition for the ADRA2B and VWF 50 taxa data sets. However, removal of third codon positions of ADRA2B and VWF in the combination analysis greatly reduces the base composition heterogeneity, only four taxa (*Echinops*, Hystricidae, *Mus*, *Tonatia*) still departing from stationarity. Given these peculiarities, we also considered a data set reduced to the 13 xenarthran taxa to assess their phylogenetic relationships. This reduced data set was homogenous in terms of base composition when all codon positions of the three genes are included.

Evolutionary dynamics

Comparisons of ML GTR + Γ_8 parameters estimated for each codon position reveal a particular behavior of BRCA1 relative to ADRA2B and VWF (Figure 1). Indeed, for the two latter genes,

we observed almost the same pattern of substitutions at first and second codon positions. These positions are characterized by an excess of transitions over transversions especially marked at second codon positions, but seem to exhibit comparable relative evolutionary rates. Furthermore, these positions share similarly small values of the gamma shape parameter ($\alpha < 0.65$) indicating strong rate heterogeneity at these positions within each gene. As expected from genetic code redundancies, third codon positions of ADRA2B and VWF evolve about three times faster than their respective first and second codon positions. High gamma shape parameter values ($\alpha > 2.92$) indicate that these fast rates have a rather homogenous distribution along third codon positions of these exons. In striking contrast to this typical coding gene pattern, all codon positions of BRCA1 seem to behave similarly and evolve at about the same rate (Figure 1). Indeed, there is almost no difference between the three codon positions in terms of transition/transversion ratio, gamma shape parameter and relative evolutionary rate. They are all characterized by high values of the gamma shape parameter ($\alpha > 2.09$) and an evolutionary rate about three times higher than the slowest evolving character partition, i. e. the second codon positions of ADRA2B.

Given the marked differences between the three genes and between their codon positions, we combined them following the partitioned likelihood approach of Yang (1996b), rather than constraining substitution rate parameters and branch lengths to be the same for the three markers.

A significant increase in log-likelihood is observed when using partitioned likelihood models according to codon positions. For example, in the case of the ABV 1+2+3 set of characters, the Akaike Information Criterion (AIC) decreases from 175 910.58 ($= 2 \times 87\,849.29 + 2 \times 106$) to 174 875.84 ($= 2 \times 86\,483.92 + 2 \times 106 \times 9$) when an independent set of GTR + Γ_8 and branch length parameters is attributed to each codon position. This result indicates that the partitioned likelihood parameter-rich model incorporating nine independent sets of 106 free parameters—three base frequencies, five GTR rates, one Gamma shape, and 97 branch lengths—more appropriately describes the underlying evolutionary process than a single GTR + Γ_8 model for the whole concatenation. Additional analyses based on the AIC (not shown) indicate that the

better fit of the partitioned models according to genes and codon positions is due, in decreasing order of impact, to variable evolutionary rates along branches (cf. individual phylograms provided in the Supplementary Material), to variable rates between sites (as measured by the Γ distribution), to variable base compositions, and to variable GTR substitution rates. Given the huge computation time constraints to search for the best tree under partitioned likelihood models, we however conducted heuristic searches using a single GTR + Γ_8 model under PAUP*, but evaluated alternative rooting hypotheses under partitioned models with PAML.

Nucleotide substitution saturation analyses

Saturation plots of the pairwise observed differences between the 50 sequences as a function of the pairwise number of substitutions inferred on the ML tree (Philippe and Forterre 1999) are presented in Figure 2 for third codon positions of each gene. The slope of the regression line between the numbers of observed differences and inferred substitutions is an indication of the relative saturation level of the characters considered: the greater the number of inferred substitutions (abscissa) is for a given number of differences (ordinate), the higher the level of saturation. Under this graphical representation, third codon positions of BRCA1 (slope = 0.44) appear less saturated than third codon positions of ADRA2B (slope = 0.28) and VWF (slope = 0.19) for which there is considerable dispersion of pairwise comparison points. This pattern indicates that multiple substitutions are more frequent at these positions and that a beginning of saturation is reached even for comparisons between quite closely related taxa. For the 141 pairwise comparisons involving the three marsupials against the 47 placentals (triangles in Figure 2), the plots reach a plateau for ADRA2B and VWF (slopes of 0.004 and 0.05, respectively), but not for BRCA1 (slope = 0.31). So, third codon positions of BRCA1 show a weak tendency to form a plateau, being almost linearly distributed just below the $Y = X$ straight line. These saturation analyses strengthen the peculiarity of BRCA1 third codon positions, which appear to

be less saturated than those of ADRA2B and VWF. By contrast, saturation analyses for first and second codon positions of the three genes did not reveal significant saturation, graphical patterns being very similar to the one of BRCA1 third codon positions (data not shown). Therefore, third codon positions of BRCA1 do not seem to exhibit strong saturation in nucleotide substitutions and are more likely to have retained some deep phylogenetic signal than those of ADRA2B and VWF.

Congruence between individual genes

Topological congruence between the individual genes was evaluated by crossed SH tests in which the highest-likelihood topologies obtained with individual and combined data sets were compared against each other (Table 1). Each individual data set significantly rejects the ML topology of the other two, with the exception of the BRCA1 ML topology under the ADRA2B data set. However, none of the three individual data sets significantly rejects the ML topology obtained with the combined data set ($0.18 < P < 0.61$), suggesting that combining individual genes leads to a phylogenetic estimate compatible with the signal contributed by each individual gene. ML trees derived from each of the three individual data sets that were combined are supplied as supplementary material. Thus, we decided to concatenate the three data sets in a total evidence approach designed to maximize the number of character analyzed.

Phylogenetic results

Xenarthrans

The maximum likelihood phylogram depicting intra-xenarthran relationships is presented in Figure 3. It shows an almost fully resolved topology apart from two notable exceptions within the armadillos (Dasypodidae). Some of the nodes (Folivora, Vermilingua, Cingulata and *Dasypus*) are defined by exclusive amino acid deletions occurring in the ADRA2B and BRCA1

proteins. It is also noticeable that all armadillos included in this study have a very reduced glutamic acid repeat region in ADRA2B as compared to most other placental mammals (data not shown). Within Pilosa, the monophyly of sloths (Folivora), anteaters (Vermilingua), and Myrmecophaginae (*Myrmecophaga* + *Tamandua*) receives 100% ML bootstrap (BP_{ML}) support and a Bayesian posterior probability (P_{bay}) of 1.00. The concatenation of the three genes also strongly supports (BP_{ML} = 100; P_{bay} = 1.00) the existence of three distinct lineages within Cingulata: [*Dasybus*], [*Priodontes*, *Cabassous*, *Tolypeutes*] and [*Chaetophratus*, *Euphractus*, *Zaedyus*]. The genus *Dasybus*, here represented by two divergent species (*D. novemcinctus* and *D. kappleri*), appears well separated from the two other groups, which are closely related (BP_{ML} = 100; P_{bay} = 1.00). However, relationships within these two clades (Tolypeutini and Euphractini) are characterized by extremely short internal branches, low ML bootstrap supports and moderate Bayesian posterior probabilities.

Placentals

Figure 4 presents the maximum likelihood phylogram obtained for the relationships among placental mammals using the A2BV2 data set. Despite the shortness of its deepest nodes, this tree strongly supports (BP_{ML} > 98; P_{bay} = 1.00) four major placental clades: I) Afrotheria, II) Xenarthra, III) Euarchontoglires and IV) Laurasiatheria. The grouping of super-clades III and IV into Boreoeutheria is also robustly evidenced (BP_{ML} = 98; P_{bay} = 1.00). In this ML topology, Afrotheria appears as the first offshoot of the placental tree, with moderate support (BP_{ML} = 61; P_{bay} = 0.81). However, the Afrotheria + Xenarthra hypothesis is preferred when removing all third codon positions (BP_{ML} = 57; P_{bay} = 0.46). Relationships among afrotherians are rather well resolved. Our results strongly support the division of Afrotheria in two main clades: Paenungulata (Proboscidea, Sirenia and Hyracoidea; BP_{ML} = 100; P_{bay} = 1.00) and a clade grouping Tubulidentata, Macroscelidea, Chrysochloridae and Tenrecidae (BP_{ML} = 99; P_{bay} = 1.00). Whereas the relationships among paenungulates remain unresolved, there is strong support

for a sister-group relationship between Chrysochloridae and Tenrecidae ($BP_{ML} = 94$; $P_{bay} = 1.00$) and the early emergence of Tubulidentata is evidenced ($BP_{ML} = 99$; $P_{bay} = 1.00$). Within Boreoeutheria, our data robustly support the monophyly of rodents ($BP_{ML} = 100$; $P_{bay} = 1.00$) within the super-clade III, but their grouping with lagomorphs into Glires is only moderately supported ($BP_{ML} = 60$; $P_{bay} = 0.97$). The inter-relationships of Primates, Scandentia and Dermoptera are not resolved by our data. Within Laurasiatheria, we retrieved the monophyly of Eulipotyphla (“core insectivores”), Cetartiodactyla, Perissodactyla and Chiroptera ($BP_{ML} = 100$; $P_{bay} = 1.00$) and a robust sister-group relationship between Carnivora and Pholidota is evidenced ($BP_{ML} = 94$; $P_{bay} = 1.00$). However, the relationships between these five monophyletic groups remain unclear, with the exception of an early divergence separating Eulipotyphla from the others ($BP_{ML} = 89$; $P_{bay} = 1.00$).

Statistical tests of the position of the root

Effect of character sampling

To evaluate the stability of the root of the placental tree in the three data sets, we compared by the SH test the 15 possible bifurcating topologies depicting ingroup relationships between the four major clades (Table 2). The results of the partitioned SH tests appear to depend on the analyzed characters. Indeed, when only first and second codon positions (ABV 1+2) are considered, a basal Xenarthra + Afrotheria clade appears as the most likely hypothesis, whereas the early emergence of Afrotheria is favored when considering either only first and second codon positions of ADRA2B and VWF plus all codon positions of BRCA1 (A2BV2), all codon positions of the three genes (ABV 1+2+3), or amino acids (ABV AA). However, the first three hypotheses (basal Afrotheria, basal Afrotheria + Xenarthra, and Epitheria) are not significantly different from a statistical perspective and are almost indistinguishable on the basis of their likelihood scores when considering only positions 1+2 or amino acids (Table 2). All other

alternative hypotheses but two are significantly worse at the 5% level whatever data set is considered (Table 2). The two remaining alternatives break the monophyly of Boreoeutheria by placing either Euarchontoglires or Laurasiatheria at the base of a tree in which Xenarthra and Afrotheria group together. Despite the fact that they are not significantly rejected by the SH test (P values ranging from 0.106 to 0.203), they involved a severe drop in log-likelihood relative to the three main hypotheses (Table 2). The monophyly of Boreoeutheria is strongly suggested, thus it is highly likely that the position of the root falls along a branch connecting Afrotheria, Xenarthra and Boreoeutheria.

Effect of taxon sampling

In order to test the effect of increasing taxon sampling within Xenarthra on the location of the root, we compared the results of SH tests for the three main competing hypotheses previously identified (basal Afrotheria, Afrotheria + Xenarthra, and Epitheria) using all possible trees including respectively one (13 combinations), three (48), six (84), eight (56) and 13 (1) xenarthran taxa. Table 3 summarizes the percentage of times that each of these three alternatives appears as the best hypothesis. As previously observed (Table 2) when considering all 13 xenarthrans for the first and second codon positions of the three genes, the Afrotheria + Xenarthra hypothesis appeared as the best hypothesis instead of the basal Afrotheria hypothesis. This result is dependent on the number of species chosen to represent Xenarthra. Indeed, if only one xenarthran is considered, the basal Afrotheria hypothesis is the most likely in 77% of the comparisons when considering only positions 1+2. Increasing the xenarthran taxonomic representation inverted the tendency and results converged towards the Afrotheria + Xenarthra hypothesis (Table 3). In contrast, the situation was clearer when using the two longer data sets. In those cases, a basal position of Afrotheria was preferred in all comparisons but two (when using the sloths *Choloepus* or *Bradypus* with all codon positions). Interestingly, despite the fact that it

was not rejected at the 5% significance level, Epitheria was never scored with the highest likelihood hypothesis over all 606 topologies explored.

To appreciate the extent of likelihood variation between comparisons including a variable number of xenarthran taxa, we graphically represented the SH test P values observed for the Epitheria hypothesis, as a function of the number of xenarthrans considered (Figure 5). This graph illustrates the interacting effects of increasing both taxa and character sampling. Two effects can be evidenced. First, the longer the sequences, the lower the support for the Epitheria hypothesis. Second, the larger the species sampling, the smaller the P value dispersion range. However, it seems that increasing the number of analyzed characters leads to the expansion of P value ranges by contrasting the differences between estimates. For example, when a single xenarthran is sampled for the all-codon-positions data set, the P values range from a highly non-significant value of 0.631 with the sloth *Bradypus*, to a marginally significant value of 0.054 with the armadillo *Cabassous*. Thus, including only a small number of selected xenarthran representatives could be misleading in terms of the results of hypotheses testing, especially when a large number of nucleotides is considered.

DISCUSSION

Comparative evolutionary dynamics and phylogenetic content of the three nuclear markers

Detailed analyses of the evolutionary dynamics of the three nuclear protein-coding genes revealed marked differences in terms of base compositions and substitution patterns. Each of these genes is located on a different chromosome in the human genome—chromosomes 2, 12, and 17 for ADRA2B, VWF, and BRCA1, respectively—and observed differences in base composition might be related to the isochore structure of the genome (Bernardi 2001). Regarding molecular evolution, our results expand those of Adkins et al. (2001) who reported the

observation of a similar rate of evolution for the three codon positions in BRCA1 of rodents. All three codon positions of this gene exhibit moderate rates of substitution and striking among-sites rate homogeneity ($\alpha > 2.09$). This unusual pattern of evolution might reflect peculiar selective constraints acting on this large spectrum regulatory protein (Deng and Brodie 2000). Moreover, saturation analyses illustrated that third codon positions of BRCA1 retain a less-saturated phylogenetic signal for comparisons within placentals and between placentals and marsupials relative to ADRA2B and VWF (Figure 2). The saturation of ADRA2B and VWF third codon positions might also reflect the fact that their very high GC content restricts the acceptable nucleotide states to mainly G or C. The case of BRCA1 argues for a careful preliminary examination of the phylogenetic properties of molecular character partitions, before deciding to include or exclude them in subsequent analyses. Moreover, the marked differences observed between the three phylogenetic markers and the significant gain in log-likelihood legitimize the application of partitioned likelihood models when comparing alternative hypotheses using their concatenation (Yang 1996b).

Placental mammal phylogeny

In many aspects the phylogenetic picture of placental mammal relationships that we obtained is fully compatible with the challenging studies of Madsen et al. (2001) and Murphy et al. (2001a, b). Our study based on the detailed ML analysis at the DNA and amino acid levels of 5130 bp of three nuclear protein-coding genes yielded conclusions consistent with those obtained on 16397 bp by Murphy et al. (2001b). They strongly confirmed the existence of the four major placental clades Xenarthra, Afrotheria, Euarchontoglires, and Laurasiatheria and add support to the grouping of the Northern Hemisphere clades Euarchontoglires and Laurasiatheria in the so-called Boreoeutheria (Springer and de Jong 2001).

Xenarthra

We here present the first comprehensive study of xenarthran molecular phylogeny based on the analysis of three nuclear markers for 12 of the 13 living genera. Analyses of the placental data sets (Figure 4) unambiguously supported the monophyly of the order Xenarthra and its division in the two suborders Cingulata (armadillos) and Pilosa (anteaters and sloths). This result confirms those obtained on both morphological (Engelmann 1985; Patterson et al. 1992; Gaudin 1999) and molecular (van Dijk et al. 1999; Delsuc et al. 2001; Madsen et al. 2001; Murphy et al. 2001a) characters. Our results also led to a clear picture of Xenarthra inter-relationships with almost all nodes robustly evidenced (Figure 3).

Relationships within anteaters and sloths (Pilosa) are fully congruent with previous analyses (Delsuc et al. 2001) and confirm both the respective monophyly of anteaters (Vermilingua) and sloths (Folivora), and the early emergence of the pygmy anteater (*Cyclopes*) within Vermilingua (Figure 3). These results emphasize the very deep split between larger body sized anteaters (*Tamandua* and *Myrmecophaga*) and the pygmy *Cyclopes* reflected by numerous morphological peculiarities related to its strictly arboreal way of life (Gaudin and Branham 1998).

Regarding armadillos (Cingulata, Dasypodidae), our results clearly identify three distinct lineages corresponding to the subfamilies defined by McKenna and Bell (1997): Dasypodinae (*Dasypus kappleri* and *D. novemcinctus*), Euphractinae (*Chaetophractus*, *Euphractus* and *Zaedyus*) and Tolypeutinae (*Priodontes*, *Cabassous* and *Tolypeutes*). The evidence strongly supports the early emergence of Dasypodinae whereas Tolypeutinae and Euphractinae unequivocally cluster together (Figure 3). Such a relationship has already been proposed based on the study of spermatozoa (Cetica et al. 1998) and contradicts the early morphological assessment of Engelmann (1985). However, relationships within the subfamilies Euphractinae and Tolypeutinae remain unclear. Our results suggest the grouping of *Euphractus* and *Zaedyus* to the exclusion of *Chaetophractus* within Euphractinae. Such a grouping is consistently retrieved in most of the analyses but appears with only poor support in the total evidence ML tree (Figure

3). This lack of resolution is not unexpected because the three genera are morphologically so similar that their inter-relationships have always been considered as unresolved (Engelmann 1985; Patterson, Segall, and Turnbull 1989). More surprising is the trifurcation obtained within Tolyteutinae. Indeed, a close relationship between *Cabassous* and *Priodontes* was expected based on their comparable external morphologies (Engelmann 1985; Wetzel 1985) and spermatozoa (Cetica et al. 1998). By contrast, in our ML phylogeny, *Priodontes* emerges in a basal position relative to *Cabassous* and *Tolyteutes* but with low support (Figure 3). Thus, relationships between these three genera are still unclear. However, *Tolyteutes* possesses a highly distinct morphology shaped by the anatomical constraints induced by its capacity to entirely roll into a ball, hence rendering difficult the interpretation of homology in post-cranial characters. In this context, it is interesting to note that characters of the ear region tend to support a sister-group relationship between *Tolyteutes* and *Cabassous* as proposed by our data (Patterson, Segall, and Turnbull 1989). The apparent lack of resolution observed within these two clades despite the high number of nucleotides analyzed raises the question of the rapidity by which the splitting events occurred. The shortness of internal nodes associated with low support values suggest that rapid cladogenesis left only short time intervals for molecular synapomorphies to accumulate in these two groups. It is hoped that considering more rapidly evolving molecules, such as nuclear introns or mitochondrial genes, will help to resolve the only two remaining uncertainties in Xenarthra phylogeny.

Afrotheria

Despite its "highly provocative" nature for morphologists (Novacek 2001), there is at present little doubt about the naturalness of the African clade for which support consistently increases as molecular evidences accumulate (de Jong, Leunissen, and Wistow 1993; Douzery and Catzeflis 1995; Lavergne et al. 1996; Springer et al. 1997, 1999; Stanhope et al. 1998a, b; Mouchaty et al. 2000b; Madsen et al. 2001; Murphy et al. 2001a, b; van Dijk et al. 2001). Our results provide

strong ML bootstrap support for nodes within Afrotheria that were previously difficult to resolve (Madsen et al. 2001; Murphy et al. 2001a). Indeed, whereas the retrieval of the well-defined Paenungulata (elephants, sirenians and hyraxes) was expected, a second major afrotherian clade regrouping Tubulidentata (aardvark), Macroscelidea (elephant shrews), Chrysochloridae (golden moles) and Tenrecidae (tenrecs) is robustly confirmed. Within this clade, the grouping of the two ex-insectivoran families Chrysochloridae and Tenrecidae is also well supported. Moreover, in agreement with Murphy et al. (2001b), our results support the early emergence of Tubulidentata leading to the monophyly of an African insectivore-like clade (Macroscelidea + Afrosoricida). It is finally noteworthy that, as in previous analyses of concatenated nuclear genes (Madsen et al. 2001; Murphy et al. 2001a, b), relationships among paenungulates remain unsettled despite the high numbers of nucleotides analyzed. The resolution of this tricky question and the verification of the proposed sister-group relation between Macroscelidea and Afrosoricida might benefit from an expanding taxon sampling within Afrotheria.

Boreoeutheria

The association of Euarchontoglires and Laurasiatheria appears strongly supported in all ML analyses (Figure 4) and its monophyly is suggested by most ML statistical tests (Table 2). Such an arrangement, fundamentally opposing Northern and Southern placental groups, argues for a more crucial importance of plate tectonic events in the early diversification of placental mammals than previously recognized (Eizirik, Murphy, and O'Brien 2001; Madsen et al. 2001; Murphy et al. 2001a, b). It also reveals the occurrence of parallel adaptive radiations on separated continental masses that are likely responsible for the long-standing difficulties encountered by systematic morphologists to reconstruct placental mammal evolution (Madsen et al. 2001; Scally et al. 2001).

The Euarchontoglires clade containing the orders Primates, Scandentia (tree shrews), Dermoptera (flying lemurs), Lagomorpha (rabbits, hares and pikas) and Rodentia (rodents) is consistently

retrieved in all analyses with high support values (Figure 4). The monophyly of rodents is unambiguously supported and their association with lagomorphs into Glires received moderate support. This arrangement, which has for a long time been favored by morphological studies (Gregory 1910; Simpson 1945), has proved difficult to retrieve in molecular studies, but was finally confirmed by larger taxon sampling (Murphy et al. 2001a; Huchon et al. 2002). However, it has been questioned whether the Murphy et al. (2001a) data set is adequate to resolve the relationships of rabbits, rodents and primates or needs more genes with longer sequences (Rosenberg and Kumar 2001).

Similar to Afrotheria and Xenarthra, the clade Laurasiatheria (Waddell, Okada, and Hasegawa 1999) is always supported by 100% bootstrap support values or Bayesian posterior probabilities of 1.00 in our analyses. The existence of this clade is corroborated by analyses of complete mitochondrial genomes, in which the basal position of the hedgehog is likely considered to be an artifact due to its peculiar nucleotide composition (Cao et al. 2000; Mouchaty et al. 2000b; Nikaido et al. 2001). Within this clade, the proposal of an early divergence of Eulipotyphla from the remaining laurasiatherians (Waddell, Okada, and Hasegawa 1999; Murphy et al. 2001b) finds support in our data. The monophyly of Chiroptera is unambiguously supported by our data as is the paraphyly of Microchiroptera (Hipposideros, Megaderma, Tonatia, Myotis, Tadarida) and its subsequent implications for the evolution of echolocation and flight in bats (Hutcheon, Kirsch, and Pettigrew, 1998; Teeling et al. 2000, 2002; Springer et al. 2001). An interesting input provided by our data resides in the reliable support obtained for the sister-group relationship of Pholidota (pangolins) and Carnivora (Figure 4). This grouping, called Ferae by Shoshani and McKenna (1998), was suggested by early molecular studies (de Jong et al. 1985; Shoshani 1986; de Jong, Leunissen, and Wistow 1993). In light of our results and those of Murphy et al. (2001b), it seems likely that the presence of a well-developed ossified tentorium in both carnivores and pangolins represents a true synapomorphy uniting these two orders (Shoshani and McKenna 1998). Cetartiodactyla is indubitably monophyletic based on our data (Figure 4) and other

multiple sets of molecular evidence (Gatesy et al. 1999; Montgelard et al. 1997; Shimamura et al. 1997). In addition, the artiodactyl affinities of whales have been confirmed by recent fossil discoveries (Gingerich et al. 2001; Thewissen et al. 2001). Among Cetartiodactyla, the relationships are entirely congruent with those obtained from studies of other genes (Gatesy et al. 1999; Montgelard et al. 1997; Ursing and Arnason 1998; Madsen et al. 2001; Murphy et al. 2001a, b) and SINE insertions (Shimamura et al. 1997; Nikaido, Rooney, and Okada 1999). Nevertheless, relationships among these five major Laurasiatherian clades remain unclear.

Character and taxon sampling and the root of the placental tree

Despite extensive taxon sampling, our data did not allow the exact position of the root of the placental tree to be determined. Indeed, like Madsen et al. (2001) and Murphy et al. (2001a), we identified three almost equally likely possible root locations corresponding to the following topological arrangements: basal Afrotheria, basal Xenarthra + Afrotheria and basal Xenarthra (as sister to Epitheria). Recently, Murphy et al. (2001b) proposed the root along the branch leading to Afrotheria. One of our goals was to evaluate to what extent increasing the taxon representation of a previously poorly sampled basal clade (Xenarthra) might help to stabilize the position of the root on the placental tree.

Although not statistically conclusive, a detailed examination of the impact of both character and taxon sampling on root location tests (SH tests) allows for the identification of clear trends. Thus, it is noteworthy that Epitheria never appears as the most likely hypothesis suggesting that Xenarthra does not represent the earliest offshoot of the placental tree as previously thought (McKenna 1975). The location of the root between the two remaining hypotheses is sensitive to the molecular characters sampled. Indeed, the basal Afrotheria + Xenarthra hypothesis tends to be slightly favored when only first and second codon positions are used, whereas the basal Afrotheria solution is preferred when adding BRCA1 third codon positions or third codon positions of the three genes. It is however not clear whether this results from signal due to the

growing number of sampled characters (3421 to 5130) or to noise coming from some saturated third codon positions. Moreover, it is striking to note that in amino acid analyses the three competing hypotheses for the position of the root (basal Afrotheria, Afrotheria + Xenarthra, and Epitheria) only differ by no more than 0.75 units of log-likelihood (Table 2). Increasing the number of genes that are analyzed, in association with a critical examination of different codon partitions or amino acids, will be important in further investigations of the rooting problem. It is also important to maintain adequate taxon sampling. Indeed, our results show that including only a single xenarthran representative could be misleading in terms of SH test results, especially when the number of analyzed characters increased. Actually, considering at least one xenarthran per major lineage (armadillos, anteaters, and sloths) might help to stabilize their position in the tree. It is likely that increasing the sampling within Xenarthra contributed to reduce the effect of potential base composition biases or long branch attraction artifacts associated with the distantly related marsupial outgroups. We here suggest that additional sampling among Marsupials and the addition of a representative of Monotremes might also contribute to reach more reliability for assessing the root in placentals. Larger taxon sampling for different genes might help to find a reliable rooting of the ingroup by mitigating the peculiarities of certain taxa and thus increasing the signal/noise ratio (Hillis 1996; Rannala et al. 1998). It is hoped that the future consideration of selected positions of additional genes for a balanced taxon sampling of major clades will resolve the remaining uncertainties about the root position of the placental tree. This question is of primary interest for understanding and reconstructing the early history of placental mammals, in which tectonic and biogeographic events seem to have played a crucial role (Murphy et al. 2001b).

SUPPLEMENTARY MATERIAL

An electronic appendix to this paper providing taxonomy plus accession numbers of the sequences used in this study and the ML trees inferred from the separate analysis of ADRA2B,

BRCA1, and VWF can be found on the Molecular Biology and Evolution web site (<http://www.molbioevol.org>).

ACKNOWLEDGMENTS

We would like to warmly thank the following individuals and institutions for their invaluable contribution in providing us with xenarthran tissues: Jorgue Omar García (Complejo Ecológico Municipal de Presidencia Roque Sáenz Peña, Chaco, Argentina), Guillermo Anibal Lemus, Pablo Carmanchahi, Tammie L. Bettinger (Cleveland Metroparks Zoo, US), Danny Devillier, Eric Hansen, Jesus Mavarez, and Jean-Christophe Vié and his team "Faune Sauvage" in French Guiana. FD particularly appreciated the help and friendship of Mariella Superina, Guillermo Perez Jimeno, Sergio Vizcaíno, Rodolpho Rearte, Leandro Perez and Pablo Cetica in Argentina. Christophe Douady and Heather Amrine kindly gave access to unpublished sequences of *Scalopus aquaticus* (ADRA2B) and *Vombatus ursinus* (VWF), respectively. We also want to thank Marie Sémon for her contribution to this work. Dorothee Huchon, three anonymous reviewers, and the Associate Editor Hervé Philippe provided helpful suggestions on the manuscript. INFOBIOGEN provided computation facilities. This work has been supported by the TMR Network "Mammalian phylogeny" (contract FMRX – CT98 – 022) of the European Community, a NSF grant (DEB-9903810), the "Génopôle Montpellier Languedoc-Roussillon", and the "Action Bioinformatique" of the CNRS. FD acknowledges the financial support of a MENRT grant (contract 99075) associated with the DEA "Biologie de l'Evolution et Ecologie" (Montpellier, France). This is the contribution ISEM 2002-015 of the Institut des Sciences de l'Evolution de Montpellier (UMR 5554 - CNRS). This work is dedicated to the memory of Pr. André Adoutte [1947-2002].

REFERENCES

- ADKINS, R. M., E. L. GELKE, D. ROWE, and R. L. HONEYCUTT. 2001. Molecular phylogeny and divergence time estimates for major rodent groups: evidence from multiple genes. *Mol. Biol. Evol.* 18:777-791.
- ARNASON, U., A. GULLBERG, and A. JANKE. 1997. Phylogenetic analyses of mitochondrial DNA suggest a sister group relationship between Xenarthra (Edentata) and Ferungulates. *Mol. Biol. Evol.* 14:762-768.
- BERNARDI G. 2001. Misunderstandings about isochores. Part 1. *Gene* 276:3-13.
- CAO, Y., M. FUJIWARA, M. NIKAIDO, N. OKADA, and M. HASEGAWA. 2000. Interordinal relationships and timescale of eutherian evolution as inferred from mitochondrial genome data. *Gene* 259:149-158.
- CATZEFLIS, F. 1991. Animal tissue collections for molecular genetics and systematics. *Trends Ecol. Evol.* 6:168.
- CETICA, P. D., A. J. SOLARI, M. S. MERANI, J. C. DE ROSAS, and M. H. BURGOS. 1998. Evolutionary sperm morphology and morphometry in armadillos. *J. Submicroscop. Cytol. Pathol.* 30:309-314.
- DE JONG, W. W., A. ZWEERS, K. A. JOYSEY, J. T. GLEAVES, and D. BOULTER. 1985. Protein sequence analysis applied to Xenarthran and Pholidote phylogeny. Pp. 65-76 in G. G. Montgomery, eds. *The Evolution and Ecology of Armadillos, Sloths and Vermilinguas*. Smithsonian Institution, Washington.
- DE JONG, W. W., J. A. M. LEUNISSEN, and G. J. WISTOW. 1993. Eye lens crystallins and the phylogeny of placental orders: evidence for a Macroscelid-Paenungulate clade? Pp. 5-12 in M. J. N. a. M. C. M. E. F. S. Sazlay, eds. *Mammal Phylogeny (Placentals)*. Springer-Verlag, New York.

- DELSUC, F., F. M. CATZEFLIS, M. J. STANHOPE, and E. J. DOUZERY. 2001. The evolution of armadillos, anteaters and sloths depicted by nuclear and mitochondrial phylogenies: implications for the status of the enigmatic fossil *Eurotamandua*. *Proc. R. Soc. Lond. B* 268:1605-1615.
- DENG, C.-X., and S. G. BRODIE. 2000. Roles of BRCA1 and its interacting proteins. *Bioessays* 22:728-737.
- DOUZERY, E., and F. M. CATZEFLIS. 1995. Molecular evolution of the mitochondrial 12S rRNA in Ungulata (Mammalia). *J. Mol. Evol.* 41:622-636.
- EIZIRIK, E., W. J. MURPHY, and S. J. O'BRIEN. 2001. Molecular dating and biogeography of the early placental mammal radiation. *J. Hered.* 92:212-219.
- ENGELMANN, G. F. 1985. The phylogeny of the Xenarthra. Pp. 51-63 in G. G. Montgomery, eds. *The Evolution and Ecology of Armadillos, Sloths and Vermilinguas*. Smithsonian Institution, Washington.
- FELSENSTEIN, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39:783-791.
- GATESY, J., M. C. MILINKOVITCH, V. G. WADDELL, and M. S. STANHOPE. 1999. Stability of cladistic relationships between Cetacea and higher-level Artiodactyl taxa. *Syst. Biol.* 48:6-20.
- GAUDIN, T. J., J. R. WIBLE, J. A. HOPSON, and W. D. TURNBULL. 1996. Reexamination of the morphological evidence for the Cohort Epitheria (Mammalia, Eutheria). *J. Mammal. Evol.* 3:31-78.
- GAUDIN, T. J., and D. G. BRANHAM. 1998. The phylogeny of the Myrmecophagidae (Mammalia, Xenarthra, Vermilingua) and the relationship of *Eurotamandua* to the Vermilingua. *J. Mammal. Evol.* 5:237-265.
- GAUDIN, T. J. 1999. The morphology of xenarthrous vertebrae (Mammalia: Xenarthra). *Fieldiana: Geology*, n. s. 41:1-38.
- GUTH, C. 1961. La région temporelle des Edentés. Ph. D. Thesis of the University of Paris. France.

- GINGERICH, P. D., M. HAQ, I. S. ZALMOUT, I. H. KHAN, and M. S. MALKANI. 2001. Origin of whales from early artiodactyls: hands and feet of Eocene Protocetidae from Pakistan. *Science* 293:2239-2242.
- GREENWOOD, A. D., J. CASTRESANA, G. FELDMAEIR-FUCHS, and S. PÄÄBO. 2001. A molecular phylogeny of two extinct sloths. *Mol. Phylogenet. Evol.* 18:94-103.
- GREGORY, W. K. 1910. The orders of mammals. *Bull. Am. Mus. Nat. Hist.* 27:1-524.
- HILLIS, D. M. 1996. Inferring complex phylogenies. *Nature* 383:130-131.
- HUCHON, D., O. MADSEN, M. J. J. B. SIBBALD, K. AMENT, M. J. STANHOPE, F. CATZEFLIS, W. W. DE JONG, and E. J. P. DOUZERY. 2002. Rodent phylogeny and a timescale for the evolution of Glires: evidence from an extensive taxon sampling using three nuclear genes. *Mol. Biol. Evol.*, in press.
- HUELSENBECK, J. P., and B. RANNALA. 1997. Phylogenetic methods come of age: testing hypotheses in an evolutionary context. *Science* 276:227-32.
- HUELSENBECK, J. P., and F. RONQUIST. 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17:754-755.
- HUELSENBECK, J. P., F. RONQUIST, R. NIELSEN, and J. P. BOLLBACK. 2001. Bayesian inference of phylogeny and its impact on evolutionary biology. *Science* 294:2310-2314.
- HUTCHEON, J. M., J. A. KIRSCH, and J. D. PETTIGREW. 1998. Base-compositional biases and the bat problem. III. The questions of microchiropteran monophyly. *Philos. Trans. R. Soc. Lond. B* 353:607-617.
- JONES, D. T., W. R. TAYLOR, and J. M. THORNTON. 1992. The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci.* 8:275-282.
- KISHINO, H., and M. HASEGAWA. 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.

- LAVERGNE, A., E. DOUZERY, T. STICHLER, F. M. CATZEFLIS, and M. S. SPRINGER. 1996. Interordinal Mammalian relationships: evidence for Paenungulate monophyly is provided by complete mitochondrial 12S rRNA sequences. *Mol. Phylogenet. Evol.* 6:245-258.
- LECOINTRE, G., H. PHILIPPE, H. L. VAN LE, and H. LE GUYADER. 1993. Species sampling has a major impact on phylogenetic inference. *Mol. Phylogenet. Evol.* 2:205-224.
- LOUGHRY, W. J., P. A. PRODÖHL, C. M. MCDONOUGH, and J. C. AVISE. 1998. Polyembryony in armadillos. *Am. Scient.* 86:274-279.
- Madsen, O., M. Scally, C. J. Douady, D. J. Kao, R. W. DeBry, R. Adkins, H. M. Amrine, M. J. Stanhope, W. W. de Jong, and M. S. Springer. 2001. Parallel adaptative radiations in two major clades of placental mammals. *Nature* 409:610-614.
- MCKENNA, M. C. 1975. Toward a phylogenetic classification of the Mammalia. Pp. 21-46 in W. P. Luckett, and F. S. Szalay, eds. *Phylogeny of the primates*. Plenum Press, New York.
- MCKENNA, M. C., and S. K. BELL. 1997. *Classification of Mammals above the species level*. Columbia University Press, New York.
- MONTGELARD, C., F. M. CATZEFLIS, and E. DOUZERY. 1997. Phylogenetic relationships of artiodactyls and Ceteacans as deduced from the comparison of cytochrome b and 12S rRNA mitochondrial sequences. *Mol. Biol. Evol.* 14:550-559.
- MOUCHATY, S. K., A. GULLBERG, A. JANKE, and U. ARNASON. 2000a. Phylogenetic position of the Tenrecs (Mammalia: Tenrecidae) of Madagascar based on analysis of the complete mitochondrial genome sequence of *Echinops telfairi*. *Zool. Script.* 29:307-317.
- MOUCHATY, S. K., A. GULLBERG, A. JANKE, and U. ARNASON. 2000b. The phylogenetic position of the Talpidae within eutheria based on analysis of complete mitochondrial sequences. *Mol. Biol. Evol.* 17:60-67.
- MURPHY, W. J., E. EIZIRIK, W. E. JONHSON, 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 et al. (11 co-authors). 2001b. Resolution of the early placental mammal radiation using Bayesian phylogenetics. *Science* 294:2348-2351.
- NIKAIDO, M., A. P. ROONEY, and N. OKADA. 1999. Phylogenetic relationships among cetartiodactyls based on insertions of short and long interspersed elements: hippopotamuses are the closest extant relatives of whales. *Proc. Natl. Acad. Sci. USA* 96:10261-10266.
- NIKAIDO, M., K. KAWAI, Y. CAO, M. HARADA, S. TOMITA, N. OKADA, and M. HASEGAWA. 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.
- NOVACEK, M. J. 1992. Mammalian phylogeny: shaking the tree. *Nature* 356:121-125.
- NOVACEK, M. J. 2001. Mammalian phylogeny: Genes and supertrees. *Curr. Biol.* 11:R573-R575.
- PATTERSON, B., AND R. PASCUAL. 1972. The fossil mammal fauna of South-America. Pp. 247-309 in A. Keast, F. C. Erk, and B. P. Glass, eds. *Evolution, mammals and southern continents*. State University of New-York Press, Albany.
- PATTERSON, B., W. SEGALL, and W. D. TURNBULL. 1989. The ear region in xenarthrans (= Edentata, Mammalia). Part I. Cingulates. *Fieldiana: Geology*, n. s. 18:1-46.
- PATTERSON, B., W. SEGALL, W. D. TURNBULL, and T. J. GAUDIN. 1992. The ear region in xenarthrans (= Edentata, Mammalia). Part II. Sloths, Anteaters, Palaeonodons, and a Miscellany. *Fieldiana: Geology*, n. s. 24:1-79.
- PHILIPPE, H. 1993. MUST: a computer package of management utilities for sequences and trees. *Nucl. Acids Res.* 21:5264-5272.
- PHILIPPE, H., and P. FORTERRE. 1999. The rooting of the universal tree of life is not reliable. *J. Mol. Evol.* 49:509-523.
- POE, S., and D. L. SWOFFORD. 1999. Taxon sampling revisited. *Nature* 398:299-300.
- POSADA, D., and K. A. CRANDALL. 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatics* 19:817-818.

- RANNALA, B., J. P. HUELSENBECK, Z. H. YANG, and R. NIELSEN. 1998. Taxon sampling and the accuracy of large phylogenies. *Syst. Biol.* 47:702-710.
- ROSE, K. D., and R. J. EMRY. 1993. Relationships of Xenarthra, Pholidota, and Fossil "Edentates": the morphological evidence. Pp. 81-102 in F. S. Szalay, M. J. Novacek, and M. C. McKenna, eds. *Mammal Phylogeny (Placentals)*. Springer-Verlag, New York.
- ROSENBERG, M. S., and S. KUMAR. 2001. Incomplete taxon sampling is not a problem for phylogenetic inference. *Proc. Natl. Acad. Sci. USA* 98:10751-6.
- SCALLY, M., O. MADSEN, C. J. DOUADY, W. W. DE JONG, M. J. STANHOPE, and M. S. SPRINGER. Molecular evidence for the major clades of placental mammals. *J. Mammal. Evol.* 8:239-277.
- SHIMAMURA, M., H. YASUE, K. OHSHIMA, H. ABE, H. KATO, T. KISHIRO, M. GOTO, I. MUNESHIKA, and N. OKADA. 1997. Molecular evidence from retroposons that whales form a clade within even-toed ungulates. *Nature* 388:666-670.
- SHIMODAIRA, H., and M. HASEGAWA. 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol. Biol. Evol.* 16:1114-1116.
- SHOSHANI, J. 1986. Mammalian phylogeny: comparison of morphological and molecular results. *Mol. Biol. Evol.* 3:222-242.
- SHOSHANI, J., and M. C. MCKENNA. 1998. Higher taxonomic relationships among extant Mammals based on morphology, with selected comparisons of results from molecular data. *Mol. Phylogenet. Evol.* 9:572-584.
- SIMPSON, G. G. 1945. The principles of classification and a classification of mammals. *Bull. Am. Mus. Nat. Hist.* 85:1-350.
- Springer, M. S., G. C. Cleven, O. Madsen, W. W. de Jong, V. G. Waddell, H. M. Amrine, and M. J. Stanhope. 1997. Endemic African mammals shake the phylogenetic tree. *Nature* 388:61-64.
- SPRINGER, M. S., H. M. AMRINE, A. BURK, and M. J. STANHOPE. 1999. Additional support for Afrotheria and Paenungulata, the performance of mitochondrial versus nuclear genes, and the impact of data partitions with heterogeneous base composition. *Syst. Biol.* 48:65-75.

- SPRINGER, M. S., R. W. DEBRY, C. DOUADY, H. M. AMRINE, O. MADSEN, W. W. DE JONG, and M. J. STANHOPE. 2001. Mitochondrial versus nuclear gene sequences in deep-level mammalian phylogeny reconstruction. *Mol. Biol. Evol.* 18:132-143.
- SPRINGER, M. S., and W. W. DE JONG. 2001. Which mammalian supertree to bark up? *Science* 291:1709-1711.
- SPRINGER, M. S., E. C. TEELING, O. MADSEN, M. J. STANHOPE, and W. W. DE JONG. 2001. Integrated fossil and molecular data reconstruct bat echolocation. *Proc. Natl. Acad. Sci. USA* 98:6241-6246.
- STANHOPE, M. J., V. G. WADDELL, O. MADSEN, W. W. DE JONG, S. BLAIR HEDGES, G. C. CLEVEN, D. KAO, and M. S. SPRINGER. 1998a. Molecular evidence for multiple origins of Insectivora and for a new order of endemic African insectivore mammals. *Proc. Natl. Acad. Sci. USA* 95:9967-9972.
- STANHOPE, M. J., O. MADSEN, V. G. WADDELL, G. C. CLEVEN, W. W. DE JONG, and M. S. SPRINGER. 1998b. Highly congruent molecular support for a diverse superordinal clade of endemic african mammals. *Mol. Phylogenet. Evol.* 9:501-508.
- STRIMMER, K., and A. VON HAESELER. 1996. Quartet puzzling: a quartet maximum-likelihood method for reconstructing tree topologies. *Mol. Biol. Evol.* 13:964-969.
- SWOFFORD, D. L. 1998. PAUP*. Phylogenetic Analysis Using Parsimony (* and Other Methods). Version 4.0b8. Sinauer Associates, Sunderland, Massachusetts.
- SWOFFORD, D. L., P. J. WADDELL, J. P. HUELSENBECK, P. G. FOSTER, P. O. LEWIS, and J. S. ROGERS. 2001. Bias in phylogenetic estimation and its relevance to the choice between parsimony and likelihood methods. *Syst. Biol.* 50:525-539.
- 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.

- TEELING, E. C., O. MADSEN, R. A. VAN DEN BUSSCHE, W. W. DE JONG, M. J. STANHOPE, and M. S. SPRINGER. 2002. Microbat paraphyly and the convergent evolution of a key innovation in Old World rhinolophoid microbats. *Proc. Natl. Acad. Sci. USA* 99:1431-1436.
- THEWISSEN, J. G., E. M. WILLIAMS, L. J. ROE, and S. T. HUSSAIN. 2001. Skeletons of terrestrial cetaceans and the relationship of whales to artiodactyls. *Nature* 413:277-281.
- URSING, B. M., and U. ARNASON. 1998. Analyses of mitochondrial genomes strongly support a Hippopotamus-whale clade. *Proc. R. Soc. Lond. B* 265:2251-2255.
- VAN DIJK, M. A. M., E. PARADIS, F. CATZEFLIS, and W. W. DE JONG. 1999. The virtues of gaps: Xenarthran (Edentate) monophyly supported by a unique deletion in alphaA-crystallin. *Syst. Biol.* 48:94-106.
- VAN DIJK, M. A., O. MADSEN, F. CATZEFLIS, M. J. STANHOPE, W. W. DE JONG, and M. PAGEL. 2001. Protein sequence signatures support the African clade of mammals. *Proc. Natl. Acad. Sci. USA* 98:188-193.
- VIZCAÍNO, S. F. 1995. Identificación específica de las "mulitas", género *Dasypus* L. (Mammalia, Dasypodidae), del noroeste argentino. Descripción de una nueva especie. *Mastozoología Neotropical* 2:5-13.
- WADDELL, P. J., Y. CAO, J. HAUF, and M. HASEGAWA. 1999. Using novel phylogenetic methods to evaluate mammalian mtDNA, including amino acid-invariant sites-logDet plus site stripping, to detect internal conflicts in the data, with special reference to the positions of hedgehog, armadillo, and elephant. *Syst. Biol.* 48:31-53.
- WADDELL, P. J., N. OKADA, and M. HASEGAWA. 1999. Towards resolving the interordinal relationships of placental mammals. *Syst. Biol.* 48:1-5.
- WETZEL, R. M. 1985. The identification and distribution of recent Xenarthra (= Edentata). Pp. 5-21 in G. G. Montgomery, eds. *The Evolution and Ecology of Armadillos, Sloths and Vermilinguas*. Smithsonian Institution, Washington.

- WHELAN, S., P. LIO, and N. GOLDMAN. 2001. Molecular phylogenetics: state-of-the-art methods for looking into the past. *Trends Genet.* 17:262-272.
- YANG, Z. 1994. Estimating the pattern of nucleotide substitution. *J. Mol. Evol.* 39:105-111.
- YANG, Z. 1996a. Among-site rate variation and its impact on phylogenetic analyses. *Trends Ecol. Evol.* 11:367-372.
- YANG, Z. 1996b. Maximum-likelihood models for combined analyses of multiple sequence data. *J. Mol. Evol.* 42:587-596.
- YANG, Z. 1997. *Phylogenetic Analysis by Maximum Likelihood (PAML)*. University College London, London.
- YANG, Z., and B. RANNALA. 1997. Bayesian phylogenetic inference using DNA sequences: a Markov chain Monte Carlo method. *Mol. Biol. Evol.* 14:717-724.

TABLES

Table 1

Shimodaira and Hasegawa (1999) crossed tests of topological congruence between the three nuclear markers. Log-likelihoods values of ML topologies inferred from the ADRA2B (1+2), BRCA1 (1+2+3), VWF (1+2), and combined data sets were computed using each of the four data matrices and then compared to the corresponding highest log-likelihood value (in bold). The difference in log-likelihood derived from these crossed comparisons and corresponding confidence P values of the SH test (between brackets) are also indicated. A star (*) indicates that the tested topology is significantly worse at the 5% level than the best ML topology for the corresponding data set.

	Character sets			
ML topologies	ADRA2B (1+2)	BRCA1 (1+2+3)	VWF (1+2)	Combined
ADRA2B (1+2)	-7655.03	230.17 (P < 0.001*)	58.06 (P = 0.002*)	247.63 (P < 0.001*)
BRCA1 (1+2+3)	24.93 (P = 0.10)	-49207.84	32.80 (P = 0.04*)	5.92 (P = 0.71)
VWF (1+2)	64.85 (P = 0.001*)	343.77 (P < 0.001*)	-10155.76	361.13 (P < 0.001*)
Combined	18.42 (P = 0.18)	11.56 (P = 0.61)	18.10 (P = 0.18)	-68132.36

Table 2

Results of Shimodaira and Hasegawa (1999) tests of alternative hypotheses for different partitions of the placental data set. Log likelihood of the 15 possible bifurcating topologies depicting relationships between the four major placental clades were compared with PAML using independent ML models for each codon position partition. The highest likelihood is given between square brackets. XEN = Xenarthra, AFR = Afrotheria, LAU = Laurasiatheria, EUA = Euarchontoglires. ABV (1+2) = concatenation of first and second codon positions of the three genes; A2BV2 = concatenation of first and second codon positions of VWF and ADRA2B plus all codon positions of BRCA1; ABV (1+2+3) = concatenation of all codon positions of the three genes; ABV (AA) = concatenation of amino acids for the three proteins. $\Delta \ln L / S. E.$ = ratio between the log-likelihood difference of a phylogenetic alternative relative to the best hypothesis and its standard error. P_{SH} = probability of the SH test; a star (*) indicates that the tested hypothesis is significantly rejected at the 5% level. ^{KH} indicates that the hypothesis tested is significantly rejected at the 5% level by the Kishino-Hasegawa test.

Topologies	ABV (1+2)		A2BV2		ABV (1+2+3)		ABV (AA)	
	$\Delta \ln L / S. E.$	P_{SH}	$\Delta \ln L / S. E.$	P_{SH}	$\Delta \ln L / S. E.$	P_{SH}	$\Delta \ln L / S. E.$	P_{SH}
BASAL XENARTHRA + AFROTHERIA	[-49411.37]	Best	1.87/3.78	0.845	2.47/4.65	0.826	0.54/3.76	0.90
BASAL AFROTHERIA	0.44/3.22	0.83	[-67295.90]	Best	[-86483.92]	Best	[-54801.47]	Best
EPITHERIA	1.78/2.47	0.72	28.10/8.61 ^{KH}	0.09	26.87/8.60 ^{KH}	0.10	0.75/3.55	0.84
(LAU,(EUG,(XEN,AFR)))	14.80/6.52 ^{KH}	0.20	22.30/9.63 ^{KH}	0.11	21.15/10.09 ^{KH}	0.13	18.76/9.73 ^{KH}	0.21
(EUG,(LAU,(XEN,AFR)))	14.85/6.51 ^{KH}	0.20	22.32/9.64 ^{KH}	0.11	21.22/10.07 ^{KH}	0.13	18.28/9.83 ^{KH}	0.22

(AFR,(LAU,(XEN,EUG)))	45.07/14.11 ^{KH}	0.002 [*]	59.98/15.69 ^{KH}	<0.001 [*]	57.99/14.84 ^{KH}	0.001 [*]	60.56/17.01 ^{KH}	0.001 [*]
(AFR,(EUG,(XEN,LAU)))	45.13/14.10 ^{KH}	0.002 [*]	59.97/15.70 ^{KH}	<0.001 [*]	57.86/14.88 ^{KH}	0.001 [*]	61.88/16.75 ^{KH}	0.001 [*]
(XEN,(EUG,(LAU,AFR)))	50.58/14.61 ^{KH}	0.001 [*]	70.17/17.28 ^{KH}	<0.001 [*]	71.41/16.88 ^{KH}	<0.001 [*]	63.62/18.35 ^{KH}	0.001 [*]
(XEN,(LAU,(AFR,EUG)))	50.85/14.51 ^{KH}	0.001 [*]	70.18/17.27 ^{KH}	<0.001 [*]	71.42/16.89 ^{KH}	<0.001 [*]	65.03/18.08 ^{KH}	0.001 [*]
(EUG,(AFR,(XEN,LAU)))	53.70/15.74 ^{KH}	0.001 [*]	72.55/18.46 ^{KH}	0.001 [*]	73.66/17.95 ^{KH}	<0.001 [*]	73.13/19.77 ^{KH}	0.001 [*]
(LAU,(AFR,(XEN,EUG)))	53.93/15.71 ^{KH}	0.001 [*]	72.74/18.41 ^{KH}	0.001 [*]	73.87/17.91 ^{KH}	<0.001 [*]	71.97/19.94 ^{KH}	0.001 [*]
((AFR,EUG),(XEN,LAU))	53.69/15.74 ^{KH}	0.001 [*]	72.57/18.47 ^{KH}	0.001 [*]	73.64/17.95 ^{KH}	<0.001 [*]	73.23/19.72 ^{KH}	0.001 [*]
((XEN,EUG),(LAU,AFR))	54.07/15.68 ^{KH}	0.001 [*]	72.78/18.40 ^{KH}	0.001 [*]	73.93/17.88 ^{KH}	<0.001 [*]	71.26/20.03 ^{KH}	0.001 [*]
(EUG,(XEN,(LAU,AFR)))	54.11/15.67 ^{KH}	0.001 [*]	72.85/18.38 ^{KH}	0.001 [*]	73.95/17.88 ^{KH}	<0.001 [*]	71.39/20.09 ^{KH}	0.001 [*]
(LAU,(XEN,(AFR,EUG)))	54.33/15.58 ^{KH}	0.001 [*]	72.85/18.38 ^{KH}	0.001 [*]	73.96/17.88 ^{KH}	<0.001 [*]	73.10/19.76 ^{KH}	0.001 [*]

Table 3

Comparison of the likelihood of the three main placental rooting hypotheses depending on the number of xenarthran taxa included in the data set. Values indicate the percentages of times that an hypothesis appeared with the highest likelihood in the comparisons. ABV (1+2) = concatenation of first and second codon positions of the three genes; A2BV2 = concatenation of first and second codon positions of VWF and ADRA2B plus all codon positions of BRCA1; ABV (1+2+3) = concatenation of all codon positions of the three genes. N represents the number of possible trees in each case.

Hypotheses	1 XEN (N = 13)	3 XEN (N = 48)	6 XEN (N = 84)	8 XEN (N = 56)	13 XEN (N = 1)
ABV (1+2)					
Basal Afrotheria	77	21	21	2	0
Xenarthra + Afrotheria	23	79	79	98	100
Epitheria	0	0	0	0	0
A2BV2					
Basal Afrotheria	100	100	100	100	100
Xenarthra + Afrotheria	0	0	0	0	0
Epitheria	0	0	0	0	0

ABV (1+2+3)

Basal Afrotheria	85	100	100	100	100
Xenarthra + Afrotheria	15	0	0	0	0
Epitheria	0	0	0	0	0

FIGURE LEGENDS

Figure 1

Comparative evolutionary dynamics of the three nuclear genes using the 50 mammalian taxa data set. Partitioned maximum likelihood parameters (GTR + Γ_8) for each of the three codon positions of each gene were computed using PAML 3.0c (Yang 1997). Graphic representation shown from left to right: GTR rate matrix parameters for transitions AG and CT (gray bars) and transversions AC, AT, CG and GT=1 (white bars), α parameter of the Γ_8 distribution (horizontally hatched bars), and rate between partitions relative to the slowest partition, *i.e.* position 2 of ADRA2B = 1 (diagonally hatched bars).

Figure 2

Saturation plots of the numbers of observed differences as a function of the numbers of inferred substitutions for each pair of sequences at third codon positions of the three nuclear genes. ML pairwise distances were inferred from branch lengths calculated on the best ML topology found with the concatenation of the three genes (Figure 4) and using GTR + Γ_8 parameters estimated on third codon positions of each gene. The $Y = X$ straight line shows the case where there is no homoplasy in the data. Triangles represent Marsupialia/Placentalia pairwise comparisons, diamonds intra-Xenarthra comparisons and open-circles other comparisons. Note that the scale of the ordinate axis (number of observed differences) for BRCA1 is twice the scale for the other genes.

Figure 3

Maximum likelihood phylogram of Xenarthra relationships using the all codon position concatenation of the three nuclear markers (5113 sites, of which 1481 are variable). Anteaters

and sloths (Pilosa) were used as outgroup to infer relationships among armadillos (Cingulata, Dasypodidae). ML bootstrap proportions (100 replications) and Bayesian posterior probabilities are indicated at each node. Unique amino acid deletions in the proteins supporting different clades are also indicated on the tree (triangles: ADRA2B, squares: BRCA1). GTR + Γ_8 ML parameters estimated with in PAUP*: rate matrix (AC: 0.91, AG: 4.64, AT: 0.72, CG: 1.92, CT: 4.75, GT: 1.00), $\alpha = 0.60$.

Figure 4

Maximum likelihood phylogram of placental mammal relationships using the concatenation of first and second codon positions of ADRA2B and VWF plus all codon positions of BRCA1 (4350 sites, of which 3362 variable). The three marsupials Didelphis, Macropus, and Vombatus were used as outgroup. ML bootstrap proportions (1000 replications) and Bayesian posterior probabilities are indicated at each node. GTR + Γ_8 ML parameters estimated with PAUP*: rate matrix (AC: 1.31, AG: 4.72, AT: 0.77, CG: 1.46, CT: 4.65, GT: 1.00), $\alpha = 1.23$.

Figure 5

Graphic representation of Shimodaira and Hasegawa (1999) test probability for the Epitheria hypothesis as a function of the number of xenarthran taxa sampled. This probability was computed for all possible trees including respectively one, three, six, eight and 13 xenarthrans for three data sets: ABV (1+2) [diamonds] = concatenation of first and second codon positions of the three genes (3421 sites); A2BV2 [circles] = concatenation of first and second codon positions of VWF and ADRA2B plus all codon positions of BRCA1 (4350 sites); ABV (1+2+3) [triangles] = concatenation of all codon positions of the three genes (5130 sites). Horizontal dashed lines indicate the P values of the SH test obtained for each data set considering all xenarthran taxa (13). These values are 0.404, 0.175, 0.144 respectively for the ABV (1+2), A2BV2 and ABV

(1+2+3) data sets. The continuous horizontal line represents the 5% significance level of the test below which the alternative is significantly worse than the best topology.

Figure 1

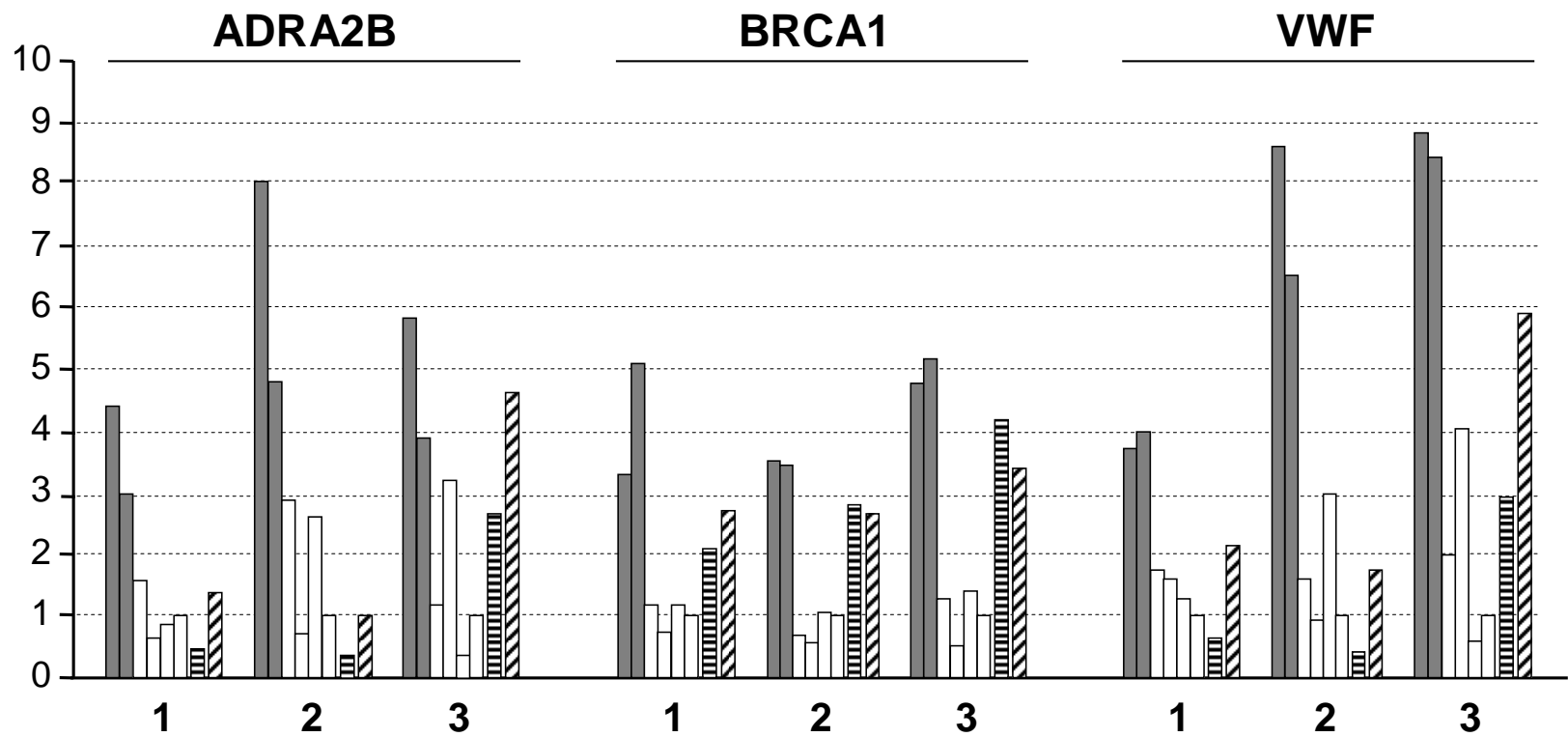


Figure 2

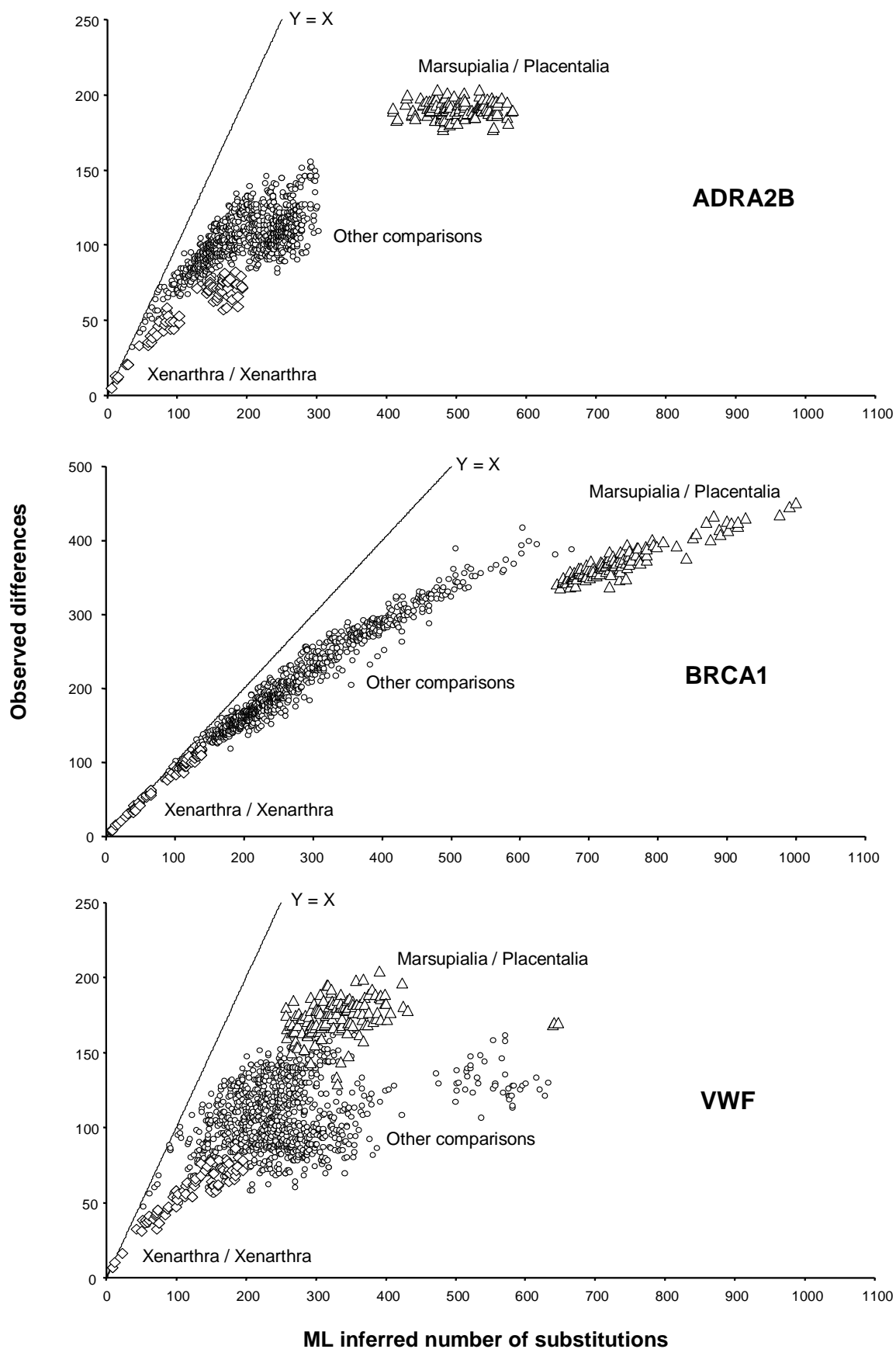


Figure 3

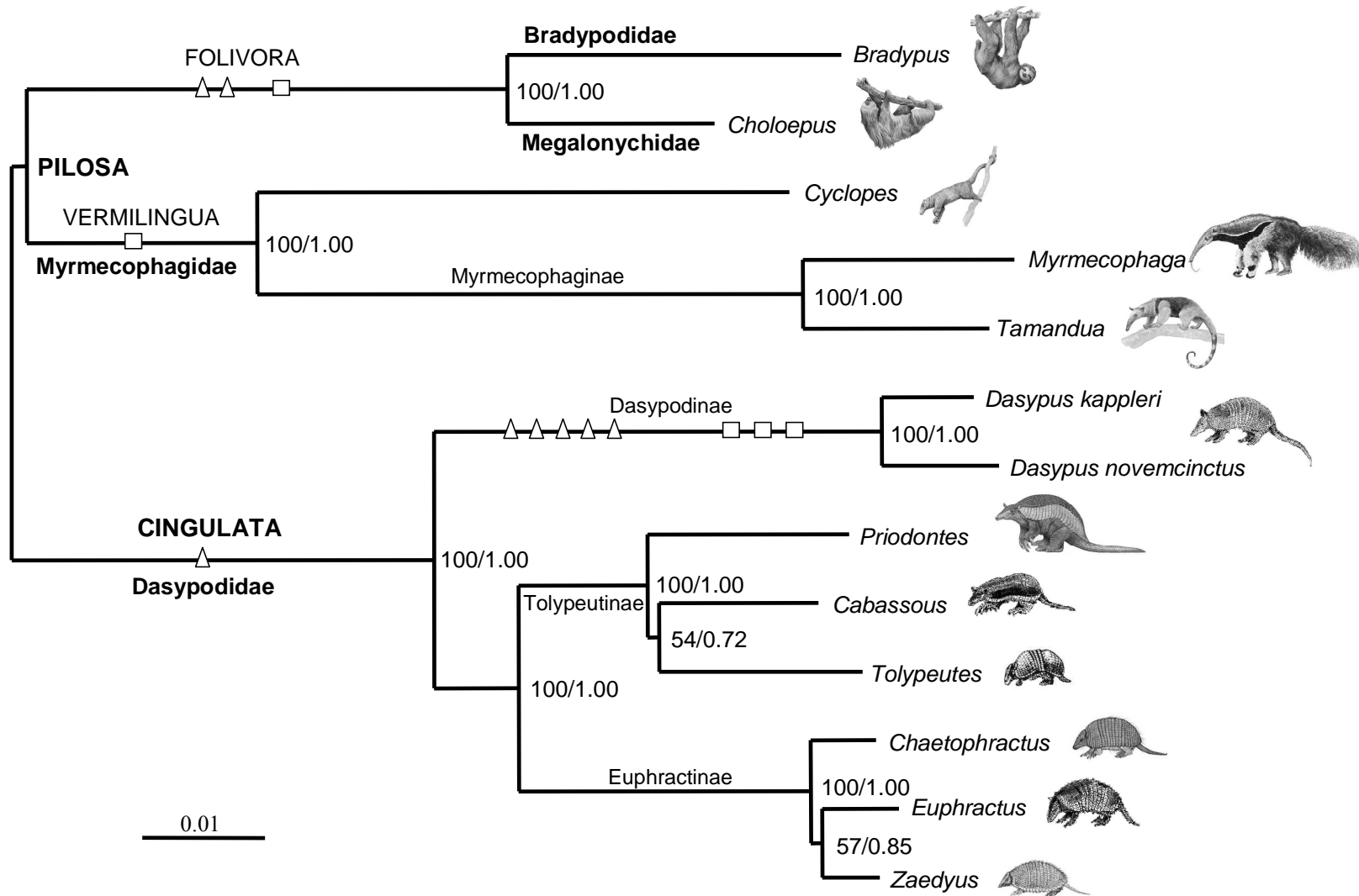


Figure 4

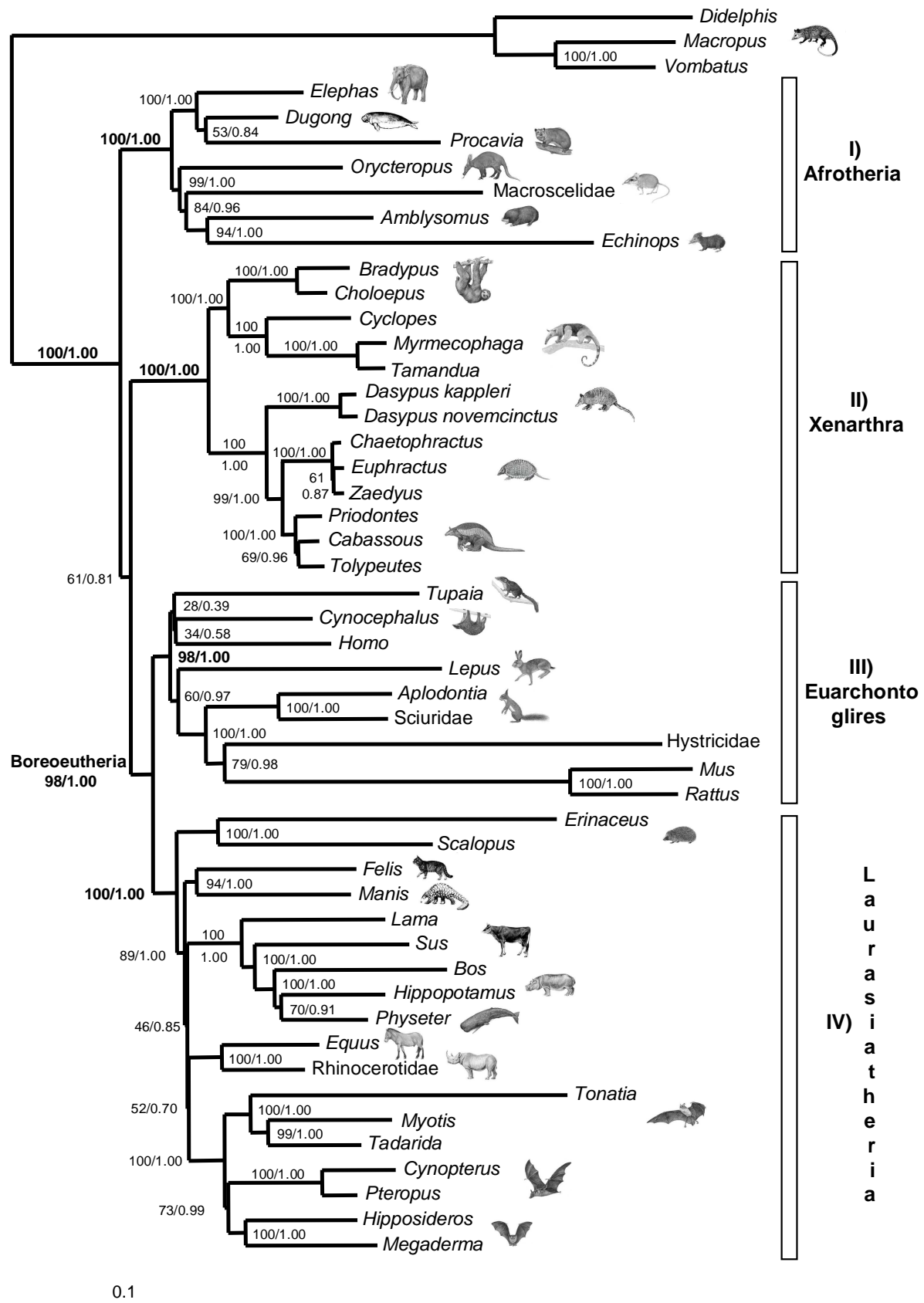


Figure 5

