

The Phylogenetic Position of the Talpidae Within Eutheria Based on Analysis of Complete Mitochondrial Sequences

Suzette K. Mouchaty, Anette Gullberg, Axel Janke, and Ulfur Arnason

Department of Genetics, Division of Evolutionary Molecular Systematics, University of Lund, Lund, Sweden

The complete mitochondrial (mt) genome of the mole *Talpa europaea* was sequenced and included in phylogenetic analyses together with another lipotyphlan (insectivore) species, the hedgehog *Erinaceus europaeus*, and 22 other eutherian species plus three outgroup taxa (two marsupials and a monotreme). The phylogenetic analyses reconstructed a sister group relationship between the mole and the fruit bat *Artibeus jamaicensis* (order Chiroptera). The *Talpa/Artibeus* clade constitutes a sister clade of the cetferungulates, a clade including Cetacea, Artiodactyla, Perissodactyla, and Carnivora. A monophyletic relationship between the hedgehog and the mole was significantly rejected by maximum parsimony and maximum likelihood. Consistent with current systematic schemes, analyses of complete cytochrome *b* genes including the shrew *Sorex araneus* (family Soricidae) revealed a close relationship between Talpidae and Soricidae. The analyses of complete mtDNAs, along with the findings of other insectivore studies, challenge the maintenance of the order Lipotyphla as a taxonomic unit and support the elevation of the Soricomorpha (with the families Talpidae and Soricidae and possibly also the Solenodontidae and Tenrecidae) to the level of an order, as previously proposed in some morphological studies.

Introduction

According to most current taxonomic schemes (e.g., MacPhee and Novacek 1993), the insectivore order Lipotyphla includes three suborders: Erinaceomorpha, with the family Erinaceidae (hedgehogs, gymnures); Chrysochloromorpha, with the family Chrysochloridae (golden moles); and Soricomorpha, with the families Solenodontidae (solenodonts), Tenrecidae (Madagascan tenrecs, African water shrews), Soricidae (shrews), and Talpidae (moles, shrew-moles, desmans). Some taxonomists (e.g., McKenna and Bell 1997) have retained the same basic arrangement but ranked the six families as orders, elevating the Lipotyphla to the superordinal level.

Based on detailed morphological analysis, Butler (1988) concluded that monophyly of Lipotyphla was supported by six shared derived characters (synapomorphies), but only two of these characters (hindgut simplification, reduction of the pubic symphysis) were considered synapomorphic by MacPhee and Novacek (1993). Although relationships among lipotyphlan families have been uncertain because of the morphological diversity of the group (Butler 1972), morphological and paleontological evidence is commonly interpreted as supporting two relationships: (1) ancient origin and clear morphological distinction of the Erinaceidae and (2) a close relationship between Talpidae and Soricidae (Butler 1988; Carroll 1988; MacPhee and Novacek 1993).

The relationship of the Lipotyphla to the other orders of eutherian mammals has remained a major problem in mammalian phylogeny (Butler 1972; Novacek 1992). The combination of numerous plesiomorphic "primitive" traits and specialized adaptations of the lipotyphlans has confounded attempts to conclusively de-

termine the phylogenetic position of Lipotyphla by morphological comparison, and recent cladistic analyses of morphological data have left Lipotyphla in an unresolved polytomy near the base of the eutherian tree (Novacek, Wyss, and McKenna 1988; Novacek 1992). Phylogenetic analyses of large mitochondrial data sets support an early divergence of the hedgehog lineage (Erinaceidae) from the rest of Eutheria (Krettek, Gullberg, and Arnason 1995; D'Erchia et al. 1996; Janke, Xu, and Arnason 1997; Penny and Hasegawa 1997). However, resolution of the phylogenetic position of Lipotyphla as a whole in these analyses, with the hedgehog as its only representative, depends on the monophyly of Lipotyphla as a taxonomic unit.

Although recent morphological reviews (Butler 1988; MacPhee and Novacek 1993) concur on the composition of Lipotyphla, recent molecular studies (Laverge et al. 1996; Springer et al. 1997; Stanhope et al. 1998) have indicated that Lipotyphla are polyphyletic, with the tenrecs and golden moles being members of a clade composed mainly of taxa endemic to the African continent, to the exclusion of the non-African lipotyphlan families. An important difference between the results of Stanhope et al. (1998), which were based on nuclear genes and combined trees, and results based on large mitochondrial data sets (Krettek, Gullberg, and Arnason 1995; D'Erchia et al. 1996; Arnason, Gullberg, and Janke 1997; Janke, Xu, and Arnason 1997) is the position of this non-African lipotyphlan clade. Stanhope et al. (1988) found that the mole and the hedgehog formed the sister group to Artiodactyla (represented by the cow), while the mitochondrial analyses have consistently placed the hedgehog as the outgroup to all other eutherians.

In this study, we determined the sequence of the complete mtDNA genome of the mole *Talpa europaea* (family Talpidae) and reconstructed the phylogenetic relationship of the mole to the hedgehog and 22 other eutherian taxa. We also sequenced the complete mitochondrial cytochrome *b* gene of the shrew *Sorex ara-*

Key words: Phylogenetics, insectivore, Lipotyphla, mole, *Talpa europaea*, mitochondrial DNA.

Address for correspondence and reprints: Ulfur Arnason, Department of Genetics, Division of Evolutionary Molecular Systematics, University of Lund, Sölvegatan 29, S-223 62 Lund, Sweden. E-mail: ulfur.arnason@gen.lu.se.

Mol. Biol. Evol. 17(1):60–67, 2000

© 2000 by the Society for Molecular Biology and Evolution. ISSN: 0737-4038

neus (Soricidae) in order to establish the affinities between the families Talpidae and Soricidae.

Materials and Methods

An enriched mtDNA fraction was isolated from liver and kidney samples of *T. europaea* following the procedure described in Arnason, Gullberg, and Widegren (1991). The specimen was collected in Dalby, Sweden. The DNA was digested separately with *BclI*, *EcoRI*, *HindIII*, *SpeI*, and *XhoII* and cloned in the phage vectors M13mp18 or M13mp19. Portions of the 12S rRNA and NADH5 genes were amplified by PCR prior to cloning. MtDNA was isolated from a shrew (*S. araneus*) collected in Madesjö, near Nybro, Sweden, by Maarit Jaarola (specimen 5590). The cytochrome *b* gene of the shrew was amplified by PCR and then cloned. All sequencing was performed on single-stranded DNA with the dideoxy method using ³⁵S-dATP and both universal and specific oligonucleotide primers. PCR-derived sequences represent the consensus of three different clones. The mitochondrial genome sequence of the mole has been deposited under accession number Y19192, while the cytochrome *b* sequence of the shrew has accession number AJ245893. Users of these sequences are kindly requested to refer to the present paper in addition to the accession numbers.

The mole was analyzed along with the following species: the platypus *Ornithorhynchus anatinus* (Janke et al. 1996); the opossum *Didelphis virginiana* (Janke et al. 1994); the wallaroo *Macropus robustus* (Janke, Xu, and Arnason 1997); the hedgehog *Erinaceus europaeus* (Krettek, Gullberg, and Arnason 1995); the mouse *Mus musculus* (Bibb et al. 1981); the rat *Rattus norvegicus* (Gadaleta et al. 1989); the guinea pig *Cavia porcellus* (D'Erchia et al. 1996); the rabbit *Oryctolagus cuniculus* (Gissi, Gullberg, and Arnason 1998); the gibbon *Hylobates lar* (Arnason, Gullberg, and Xu 1996); the human, *Homo sapiens* (Arnason, Xu, and Gullberg 1996); the armadillo *Dasypus novemcinctus* (Arnason, Gullberg, and Janke 1997); the fruit bat *Artibeus jamaicensis* (Pumo et al. 1998); the harbor seal *Phoca vitulina* (Arnason and Johnsson 1992); the dog *Canis familiaris* (Kim et al. 1998); the domestic cat *Felis catus* (Lopez et al. 1996); the horse *Equus caballus* (Xu and Arnason 1994); the donkey *E. asinus* (Xu, Gullberg, and Arnason 1996); the Indian rhinoceros *Rhinoceros unicornis* (Xu, Janke, and Arnason 1996) the white rhinoceros *Ceratotherium simum* (Xu and Arnason 1997); the pig *Sus scrofa* (Ursing and Arnason 1998a); the cow *Bos taurus* (Anderson et al. 1982); the sheep *Ovis aries* (Hiendleder et al. 1998); the hippopotamus, *Hippopotamus amphibius* (Ursing and Arnason 1998b); the fin whale, *Balaenoptera physalus* (Arnason, Gullberg, and Widegren 1991); and the blue whale, *Balaenoptera musculus* (Arnason and Gullberg 1993).

The phylogenetic analyses were based on the concatenated sequences of 12 mitochondrial protein-coding genes, excluding the L-strand-encoded NADH6 gene, the composition of which deviates from that of the H-

strand-encoded genes. After removing gaps and ambiguous sites adjacent to gaps, the resulting alignment contained 9,870 nt, corresponding to 3,290 amino acids (aa). The analyses were based on both amino acid and nucleotide sequences of first plus second codon positions, with the first positions of leucine codons coded as Y (pyrimidine). Three different methods of phylogenetic reconstruction were used: maximum parsimony (MP; Fitch 1971), neighbor joining (NJ; Saitou and Nei 1987), and maximum likelihood (ML; Felsenstein 1981). The PUZZLE, version 4.01 (Strimmer and von Haeseler 1996), and MOLPHY (Adachi and Hasegawa 1996a) programs were used for ML analyses, in which we applied the mtREV-24 rate matrix (Adachi and Hasegawa 1996b) and the HKY model of sequence evolution (Hasegawa, Kishino, and Yano 1985). The same models were used to generate distance matrices for NJ analysis. The ML analysis was performed assuming rate homogeneity and rate heterogeneity with four categories of variable sites and one category of invariable sites. The MP and NJ analyses were carried out using the PHYLIP package (Felsenstein 1991). The support for different positions of the mole and the hedgehog within the eutherian tree was investigated both separately and in combination by comparing the log likelihood values obtained from ML analyses of the various topologies.

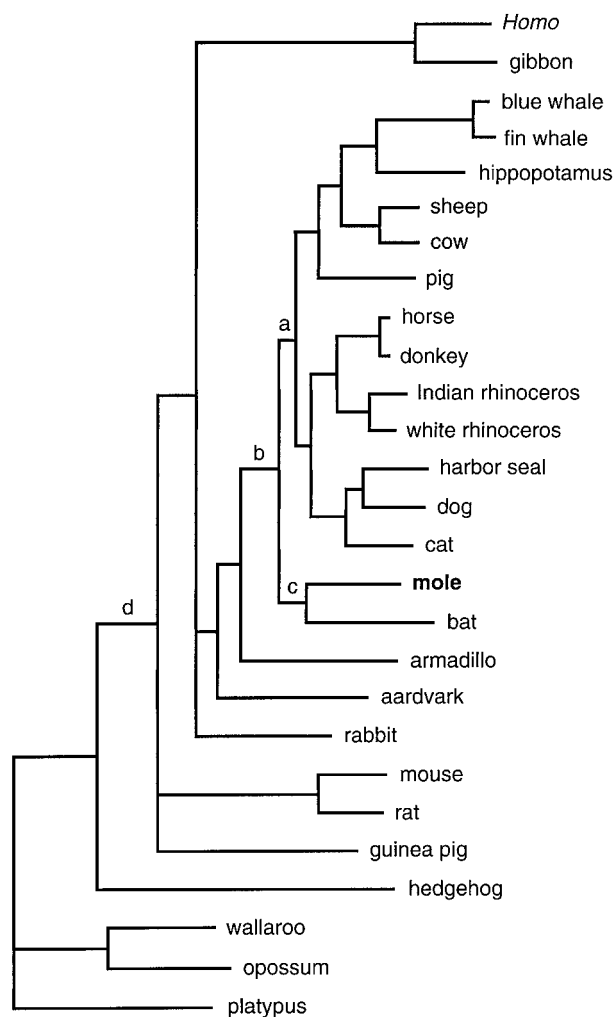
Phylogenetic analyses of the cytochrome *b* gene included *S. araneus*, along with the taxa included in the study of complete mtDNAs. The length of the cytochrome *b* alignment was 1,122 nt, corresponding to 374 aa, after removal of 18 nt (6 aa) of ambiguous homology at the 3' end of the gene. The cytochrome *b* data were analyzed in the same way as the large mitochondrial alignment. Partial sequences of the cytochrome *b* genes of a number of soricids have been reported by Ohdachi et al. (1997) and Fumagalli et al. (1999). Analyses of these sequences placed the soricids on a common branch as the sister group of the mole (data not shown).

Congruence with previously published phylogenetic hypotheses based on nuclear data sets, exon 28 of the von Willebrand factor gene (vWF, 318 aa), and the alpha-2 B-adrenergic receptor gene (A2AB, 378 aa), reported by Stanhope et al. (1998), was tested with the nonparametric test suggested by Templeton (1983). The test included the seven taxa represented by both vWF and A2AB: the hedgehog, the rat, the rabbit, the human, the mole, the horse, and the cow. The best tree topology for each data set was determined by MP, NJ, and ML. The analyses were performed on all codon positions, as well as on first plus second codon positions. The support of the nuclear sequences for the best mitochondrial tree was also tested.

Results

Characteristics of the mtDNA of *Talpa europaea*

The length of the mole mitochondrial genome described in this paper is 16,884 nt. However, the size of the genome varies in proportion to the number of 16-nt repeats in the control region; here, we report 19 repeats of ACAGGCGTATACACCC in the 1,422-nt-long con-



0.1 substitutions per site

FIG. 1.—Phylogenetic position of the mole *Talpa europaea* in a maximum-likelihood tree from the concatenated amino acid sequences of 12 mitochondrial protein-coding genes. Support values for labeled branches are shown in table 1. For the scientific names of the taxa included, see *Materials and Methods*.

tol region. The two additional clones sequenced had 19 and 24 repeats, respectively. The gene order of the genome conforms with that of other eutherians. The start codon of NADH6 is ATT; all other genes have a methionine (ATG, ATA) start codon. TAA formed the stop codon in eight genes, while incomplete stop codons (TA or T) were found in COIII, NADH2, NADH3, and NADH4; these codons may be converted into complete stop codons by posttranscriptional polyadenylation (Ojala, Montoya, and Attardi 1981). The stop codon of cytochrome *b* is AGA. The nucleotide and amino acid sequence compositions of the mole do not differ significantly from those of other mammals according to a 5% level χ^2 test (PUZZLE). The hedgehog fails the same test for nucleotide composition but passes it ($P = 0.174$) in the corresponding test for amino acid composition.

Table 1
Bootstrap and Quartet Support Values for Selected Branches of the Tree Shown in Figure 1

	a	b	c	d
NJ1, 2	89	100	55	98
aa	96	100	84	99
MP1, 2	62	98	63	87
aa	67	99	74	70
ML/QP . . .1, 2	nr	99	nr	71
aa	88	98	92	85

NOTE.—The neighbor-joining (NJ) and maximum-parsimony (MP) values represent 1,000 bootstrap replicates of nonsynonymous nucleotide changes at first and second codon positions (1, 2); the corresponding amino acid (aa) values were based on 100 replicates. The maximum-likelihood/quartet puzzling (ML/QP) support values were calculated on the basis of 1,000 puzzling steps (Strimmer and von Haeseler 1996). The ML/QP analysis of nucleotide sequences did not conclusively resolve (nr) branches a and c. These positions were therefore investigated in more detail by ML analysis of individual trees (table 2).

Phylogenetic Analysis and the Divergence Between *Talpa*, Chiroptera, and the Cetferungulates

Consistent with a previous study (Krettek, Gullberg, and Arnason 1995), the hedgehog represented the most basal eutherian taxon; however, the mole was not found to be the sister group of the hedgehog, as would be expected if Lipotyphla were monophyletic. Instead, the mole was identified as the sister group of the Chiroptera. The position of the *Talpa*/Chiroptera clade was immediately basal to the cetferungulates (fig. 1). This phylogenetic relationship was strongly supported by all three analytical approaches used (MP, NJ, and ML) and was not affected by exclusion of the hedgehog. Table 1 summarizes the bootstrap and quartet puzzle (QP) support values for the branches relevant to the positions of the mole and the hedgehog in the eutherian tree.

The topology of the tree shown in figure 1 and different topologies involving the hedgehog and the mole were examined in an exhaustive search with the number of taxa constrained to 10 operational taxonomic units. Due to computational limitations, this required the removal of one taxon from the analysis. Since the guinea pig is not crucial for the definition of the Lipotyphla, this species was not included in the exhaustive search. Under the assumption of rate homogeneity, the ML log likelihood values for the second best positions involving the mole and the hedgehog were all >2 SE worse (table 2) than the topology shown in figure 1. More distant positions became rejectable with progressively higher significance, and, as evident in table 2, monophyletic grouping of the mole and the hedgehog, either as basal in the eutherian tree (topology 4) or as the sister group of the Chiroptera (topology 6), could be rejected with high statistical significance. It is noteworthy that a transfer of the hedgehog from the base of the eutherian tree to the more apical position as the sister group of the mole is less costly than the reciprocal transfer of the mole to a position as the sister group of the hedgehog. The same observation was made in separate analyses excluding the mole and the hedgehog, respectively (table 2, trees 7–10 and 11–13). It is probable that the hedgehog lineage, with its early origin among the eu-

Table 2
Maximum-Likelihood Analyses of Different Positions of the Mole (MOLE) and the Hedgehog (HED) in the Mammalian Tree

	Topology		pBoot	ΔlnL _{RH} SE	Steps ± SD
	ΔlnL ± SE	ΔlnL ± SE			
1. OG, (HED, (MYO, (PRIM, (LAG, (TUB, (XEN, (MOLE, CHI), (PIG, (RUM, (HIP, CET))), (PER, CAR))))))	(-52,626.2)	(-45,432.1)	0.9520	(8,287)	
2. OG, (HED, (MYO, (PRIM, (LAG, (TUB, (XEN, (MOLE, CHI), (PIG, (RUM, (HIP, CET))), (PER, CAR))))))	(-40.9 ± 20.3)	(-19.4 ± 9.8)	0.0180	+11 ± 8.4	
3. OG, (HED, (MYO, (PRIM, (LAG, (TUB, (XEN, (MOLE, CHI), (PIG, (RUM, (HIP, CET))), (PER, CAR))))))	(-38.8 ± 20.9)	(-17.5 ± 10.4)	0.0280	+21 ± 7.9	
4. OG, (HED, (MOLE), (MYO, (PRIM, (LAG, (TUB, (XEN, (MOLE, CHI), (PIG, (RUM, (HIP, CET))), (PER, CAR))))))	(-389.0 ± 60.2)	(-171.6 ± 31.9)	0.0000	+116 ± 19.2	
5. OG, (MYO, (HED, (PRIM, (LAG, (TUB, (XEN, (MOLE, CHI), (PIG, (RUM, (HIP, CET))), (PER, CAR))))))	(-49.0 ± 18.6)	(-11.6 ± 6.4)	0.0020	+20 ± 9.8	
6. OG, (MYO, (PRIM, (LAG, (TUB, (XEN, (MOLE, HED), (PIG, (RUM, (HIP, CET))), (PER, CAR))))))	(-191.3 ± 43.9)	(-32.9 ± 17.9)	0.0000	+65 ± 17.3	
7. OG, (MYO, (PRIM, (LAG, (TUB, (XEN, (MOLE, CHI), (PIG, (RUM, (HIP, CET))), (PER, CAR))))))	(-49,452.5)	(-42,790.1)	0.9670	(7,572)	
8. OG, (MYO, (PRIM, (LAG, (TUB, (XEN, (MOLE, CHI), (PIG, (RUM, (HIP, CET))), (PER, CAR))))))	(-41.9 ± 20.3)	(-19.7 ± 9.8)	0.0120	+11 ± 8.5	
9. OG, (MYO, (PRIM, (LAG, (TUB, (XEN, (MOLE, CHI), (PIG, (RUM, (HIP, CET))), (PER, CAR))))))	(-41.5 ± 20.7)	(-18.3 ± 10.4)	0.0210	+23 ± 7.9	
10. OG, (MOLE, (MYO, (PRIM, (LAG, (TUB, (XEN, (MOLE, CHI), (PIG, (RUM, (HIP, CET))), (PER, CAR))))))	(-339.6 ± 56.3)	(-164.8 ± 30.2)	0.0000	+117 ± 18.3	
11. OG, (HED, (MYO, (PRIM, (LAG, (TUB, (XEN, (MOLE, CHI), (PIG, (RUM, (HIP, CET))), (PER, CAR))))))	(-51,168.7)	(-44,397.5)	0.9980	(8,024)	
12. OG, (MYO, (HED, (PRIM, (LAG, (TUB, (XEN, (MOLE, CHI), (PIG, (RUM, (HIP, CET))), (PER, CAR))))))	(-48.1 ± 18.5)	(-12.3 ± 6.6)	0.0020	+19 ± 9.6	
13. OG, (MYO, (PRIM, (LAG, (TUB, (XEN, (HED, CHI), (PIG, (RUM, (HIP, CET))), (PER, CAR))))))	(-158.4 ± 45.3)	(-26.0 ± 20.2)	0.0000	+49 ± 18.8	

NOTE.—The best maximum-likelihood tree was based on amino acid sequences using the mtREV24 model of sequence evolution and exhaustive search excluding the guinea pig to reduce the number of operational taxonomic units (OTUs) to 10, constraining the cetferungulates, primates, myomorph rodents, and outgroups to individual OTUs. OG = Outgroup (Marsupialia, wallaroo, opossum; Monotremata, platypus), HED = Erimacidae (hedgehog), MYO = Myomorphia (rat, mouse), PRIM = Primates (*Homo*, gibbon), LAG = Lagomorpha (rabbit), TUB = Tubulidentata (aardvark), XEN = Xenarthra (armadillo), MOLE = Talpidae (mole), CHI = Chiroptera (bat), PIG = pig, RUM = Ruminantia (cow, sheep), HIP = Hippopotamus, CET = Cetacea (fin whale, blue whale), PER = Perissodactyla (donkey, horse, white rhinoceros, Indian rhinoceros), and CAR = Carnivora (cat, dog, harbor seal). The value in angle brackets is the log likelihood value of the best tree. The differences of the log likelihood values of the alternative trees from that of the best tree, ΔlnL, are shown with their standard errors (Kishino and Hasegawa 1989) and estimated bootstrap probabilities, pBoot (Kishino, Miyata, and Hasegawa 1990). Likelihood estimations were done assuming rate homogeneity (MOLPHY 2.3) or rate heterogeneity (ΔlnL_{RH}) with four classes of variable sites and one class of invariable sites (PUZZLE, version 4.01). The likelihood analysis has been complemented by the Templeton (1983) test as implemented in the PHYLIP program package and shows the difference in substitutions (steps) and their standard deviations relative to the number of inferred nucleotide substitutions of the best tree (shown in angle brackets).

therians, has accumulated more homoplasies than the more recent mole lineage, thereby resulting in different costs for changing the position of the two taxa in the tree. This interpretation would be consistent with the differing number of substitutions in the best trees including only the mole or only the hedgehog.

The cytochrome *b* sequence of *S. araneus* is 80.3% and 92.3% identical to the mole sequence at the nucleotide and amino acid levels, respectively. The amino acid difference between the mole and the shrew (7.7%) is less than, for example, the intraperissodactyl difference between Equidae and Rhinocerotidae (8%), the intracarnivoran difference between the cat and the dog (11%), or the intrahominoid difference between Hylobatidae and Hominidae (13%). All three analytical approaches (MP, NJ, and ML) clearly recognized a monophyletic origin of the mole and the shrew with support values between 87% and 99%. The cytochrome *b* analyses also reconstructed a sister group relationship between the bat and the mole/sorex clade even though the analysis of a single gene did not permit resolution of the relationships among more distantly related taxa.

ML distances calculated using the mtREV-24 model of amino acid sequence evolution (Adachi and Hasegawa 1996b) were used to estimate the divergence times between the mole and the bat, as well as the origin of the mole/bat clade. The molecular clock was calibrated with the two references A/C-60, the divergence between ruminant Artiodactyla and Cetacea 60 MYA (Arnason and Gullberg 1996), and E/R-50, the divergence between Equidae and Rhinocerotidae 50 MYA (Xu, Janke, and Arnason 1996; Arnason, Gullberg, and Janke 1998). Both references yielded consistent datings, according to which the mole and the bat lineages split ≈74 MYA, while the corresponding dating for the divergence between the mole/bat clade and the cetferungulates was ≈79 MYA. In a previous study, the cetferungulate origin was defined as occurring at the divergence of edentates and cetferungulates (Arnason, Gullberg, and Janke 1997). The present study further constrains the cetferungulate origin by placing it at the divergence between *Talpa*/Chiroptera and the cetferungulate clade.

The branches defining the positions of the guinea pig and the rabbit have been collapsed in figure 1, as they were not consistently identified by all data sets and analyses. The limited resolution in these parts of the tree have been discussed previously (Arnason, Gullberg, and Janke 1997, 1999; Janke, Xu, and Arnason 1997) and therefore are not detailed here. The other relationships shown in figure 1 have been described and dated in previous studies (Janke et al. 1994; Xu, Janke, and Arnason 1996; Janke, Xu, and Arnason 1997).

ML analysis of nucleotide sequences using the HKY model of sequence evolution and rate heterogeneity with eight classes of variable sites or four classes of variable sites plus one class of invariable sites did not resolve the phylogenetic position of the hedgehog relative to other eutherian orders, as the lnL differences for the various positions of this taxon differed only marginally. It is probable that this poor resolution is related

Table 3
Templeton Test of Individual Nuclear Sequences

Sequence	Topology	Steps \pm SD
A2AB.....	1. (MYO, ((PRIM, LAG), (ART, (PER, (MOLE, HED))))))	(575)
	2. (MYO, ((PRIM, PER), (LAG, (ART, (MOLE, HED))))))	+20 \pm 6.0
	3. (MYO, ((PRIM, LAG), (PER, (ART, (MOLE, HED))))))	0 \pm 2.8
	4. (HED, (MYO, (PRIM, (LAG, (MOLE, (ART, PER))))))	+15 \pm 7.4
vWF.....	1. (MYO, ((PRIM, LAG), (ART, (PER, (MOLE, HED))))))	+25 \pm 8.3
	2. (MYO, ((PRIM, PER), (LAG, (ART, (MOLE, HED))))))	(643)
	3. (MYO, ((PRIM, LAG), (PER, (ART, (MOLE, HED))))))	+13 \pm 6.7
	4. (HED, (MYO, (PRIM, (LAG, (MOLE, (ART, PER))))))	+43 \pm 11.7

NOTE.—Details as in table 2. Topology 3 represents the relationship favored by Stanhope et al. (1998, fig. 2). Topology 4 corresponds to the results of phylogenetic analyses of the concatenated sequences of 12 mitochondrial protein-coding genes (present study).

to the significantly deviating nucleotide composition of the hedgehog. The same analysis using amino acid sequences, however, confirmed a basal position of the hedgehog (table 2).

Comparison of Different Data Sets

In contrast to the findings presented here, which are based on an alignment with a length of almost 10,000 nt, analyses of nuclear sequences (exon 28 of the vWF gene and the A2AB gene) and of combined nuclear and mitochondrial ribosomal (12S and 16S rDNA) sequences support a mole + hedgehog clade well within the eutherian tree (Stanhope et al. 1998). To investigate possible reasons for this inconsistency, we examined the relative phylogenetic signals in the sequences reported by Stanhope et al. (1998). This examination (table 3) revealed significant (2σ) inconsistencies between the phylogenetic signals of the two nuclear sequences and between the vWF and the mitochondrial rDNA data sets, indicating that these genes are subject to differing evolutionary processes or constraints (Bull et al. 1993; de Queiroz, Donoghue, and Kim 1995).

The best tree topologies for each of the vWF and A2AB data sets were established using MP, NJ, and ML. The support for the best tree from one data set was then tested against the support for the best tree from the other data set using the Wilcoxon signed-ranks test (Templeton 1983). As is evident in table 3, analyses of A2AB support a phylogenetic position of the hedgehog in a clade with the mole, the cow, and the horse, but the branching order is poorly resolved. The vWF data set contains a strong signal supporting the mole + hedgehog clade and would thus appear to be responsible for this topology in analyses of the combined data sets. However, vWF also provides strong support for excluding the horse from a (rabbit, (cow, (mole, hedgehog))) clade, which is the basis of the incongruence between the two nuclear data sets and between the vWF and the mitochondrial rDNA data sets. In cases such as these, combining the data sets may give misleading results (de Queiroz, Donoghue, and Kim 1995). Removal of unidentified nucleotides (14%) from the vWF sequence of the hedgehog had no effect on the results.

Discussion

The analyses presented here have identified the hedgehog as the most basal eutherian taxon. This is con-

sistent with the results of previous analyses of concatenated mitochondrial protein-coding sequences, performed using a somewhat different species representation and without the mole (Krettek, Gullberg, and Arnason 1995; Arnason, Gullberg, and Janke 1997; Penny and Hasegawa 1997; Cao et al. 1998; Waddell et al. 1999). It is noteworthy that one of these studies (Penny and Hasegawa 1997) used the LogDet method, an approach which is insensitive to deviations in molecular composition. This consistency among results from different analyses and different taxa samplings makes it unlikely that the basal position of the hedgehog is due to the use of any particular type of analysis or to a random sampling effect.

A nonbasal position of the hedgehog was shown in a tree reported by Sullivan and Swofford (1997). The findings which were based on a rate heterogeneity model and ML analysis of mitochondrial nucleotide sequences were inconclusive, however, as to the definite position of the hedgehog. The study included no amino acid analysis, and the possible influence of the significantly biased base composition of the hedgehog relative to the other eutherians was not discussed. When rate heterogeneity is taken into account in the ML analysis of amino acid sequences, as in the present study, the hedgehog remains the most basal eutherian taxon. Under the same model of sequence evolution, the monophyly of the mole (Talpidae) and the hedgehog (Erinaceidae) could be excluded with high statistical support. The amino acid composition of the hedgehog does not differ significantly from that of other taxa, even though it remains the most deviating; all analyses of this data set placed the hedgehog basal among the eutherians.

Analyses of two nuclear (vWF and A2AB) and mitochondrial rDNA sequences have supported a sister group relationship between the hedgehog and the mole (Stanhope et al. 1998), but trees reconstructed from the analyses showed some inconsistencies, and several ordinal relationships which have received conclusive support in studies of complete mitochondrial genomes remained unresolved. Thus, the analyses of Stanhope et al. (1998) disrupted some phylogenetic relationships, such as the cetferungulate clade, which have been strongly supported in analyses of complete mtDNAs (Arnason et al. 1996; D'Erchia et al. 1996; Xu, Janke, and Arnason 1996; Arnason, Gullberg, and Janke 1997, 1998; Janke, Xu, and Arnason 1997; Ursing and Arna-

son 1998a, 1998b). Our examination of the vWF and A2AB data sets showed that they favored different phylogenies, with the vWF data set containing a strong signal for the mole + hedgehog clade. However, at the same time, vWF did not support other relationships, such as the affinity between Artiodactyla and Perissodactyla, which have been identified in other molecular studies. Our examination of these findings (see table 3) indicate that the vWF and A2AB genes are under different evolutionary constraints, which may result in gene phylogenies deviating from the evolutionary history of the species harboring these sequences. Comparable studies of mitochondrial protein-coding genes (Cao et al. 1994, 1998) have shown that among the H-strand-encoded genes, only one, NADH1, provides a phylogeny which with respect to one taxon, the rodents, deviates markedly from reconstructions based on the concatenated sequences of all mitochondrial protein-coding genes. Even so, the deviation in this case was still less pronounced than that occurring in the nuclear data sets used by Stanhope et al. (1998).

The phylogenetic position of Lipotyphla has remained obscure in morphological studies because of the combination of numerous “primitive” traits and specialized adaptations of lipotyphlan families. Cladistic analyses of morphological characters have recognized Lipotyphla as both monophyletic and basal in the eutherian tree (Novacek and Wyss 1986; Novacek, Wyss, and McKenna 1988; Novacek 1989, 1990). The present molecular analyses of 12 mitochondrial protein-coding genes do not support lipotyphlan monophyly. Instead, a phylogenetic position of the hedgehog at a basal position in the eutherian tree and of the mole/bat as the sister group of the cetferungulates was strongly supported by all three methods of phylogeny reconstruction used, MP, NJ, and ML (table 1). Alternative placements of the mole were without statistical support (table 2). Inclusion or exclusion of the hedgehog did not affect the position of the mole. The results challenge the retention of the Lipotyphla as a monophyletic systematic unit. Furthermore, the findings support phylogenetic reevaluation of several morphological characteristics, as proposed in a recent provocative study of eutherian evolution (Werdelin and Nilsson 1999).

Morphologists have long recognized a clear distinction between moles and hedgehogs, advocating separation of the Lipotyphla into two primary lineages, the Erinaceomorpha, with the family Erinaceidae, and the Soricomorpha, traditionally including the shrews (Soricidae) and the talpid moles (Gregory 1910; Butler 1972, 1988; McKenna 1975). As discussed by MacPhee and Novacek (1993) the morphological evidence in support of lipotyphlan monophyly is weak, and the molecular results indicating lipotyphlan polyphyly are therefore not contradicted by morphological conclusions. The most important result of the present study is the strong statistical support for the phylogenetic position of Talpidae distant from the Erinaceidae in the eutherian tree. Furthermore, our results support the traditional view of a close phylogenetic relationship between Soricidae and Talpidae in the Soricidea (Gregory 1910; Hutchinson

1968; Butler 1972, 1988). McDowell (1958) and McKenna and Bell (1997) have suggested a sister group relationship between Talpidae and Erinaceidae. The morphological support for this relationship is not strong, however, and the characters listed by McDowell have been interpreted as being plesiomorphic (Butler 1988).

The molecular data presented here are consistent with the placement of the families Talpidae and Soricidae within the Soricomorpha (Gregory 1910; McKenna 1975; Butler 1972, 1988; MacPhee and Novacek 1993) and the ordinal-level classification of the Erinaceomorpha and Soricomorpha (McKenna 1975). The study has placed the Erinaceomorpha and the Soricomorpha at very different positions in the eutherian phylogenetic tree, a finding that is inconsistent with lipotyphlan monophyly.

Acknowledgments

We thank Bengt and Kerstin Olsson for the mole specimen and Karl Fredga and Maarit Jaarola for the shrew specimen. Scott K. Davis, Bengt Olle Bengtsson, and W. Patrick Lockett provided valuable comments on the manuscript. S.K.M. was the recipient of a Fulbright fellowship for the academic year 1996–1997. This work was supported by the Erik Philip-Sörensen Foundation, by the Swedish Natural Sciences Research Council and by contract ERBCHR XCT 930254 from the European Commission. The alignments used will be made available to other researchers on request.

LITERATURE CITED

- ADACHI, J., and M. HASEGAWA. 1996a. MOLPHY version 2.3: programs for molecular phylogenetics based on maximum likelihood. *Comput. Sci. Monogr.* **28**:1–150.
- . 1996b. Model of amino acid substitution in proteins encoded by mitochondrial DNA. *J. Mol. Evol.* **42**:459–468.
- ANDERSON, S., M. H. L. DE BRUIJN, A. R. COULSON, I. C. EPERON, F. SANGER, and I. G. YOUNG. 1982. Complete sequence of bovine mitochondrial DNA. *J. Mol. Biol.* **156**: 683–717.
- ARNASON, U., and A. GULLBERG. 1993. Comparison between the complete mtDNA sequences of the blue and the fin whale, two species that can hybridize in nature. *J. Mol. Evol.* **37**:312–322.
- . 1996. Cytochrome *b* nucleotide sequences and the identification of five primary lineages of extant cetaceans. *Mol. Biol. Evol.* **13**:407–417.
- 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.
- . 1998. Molecular timing of primate divergences as estimated by two non-primate calibration points. *J. Mol. Evol.* **47**:718–727.
- . 1999. The mitochondrial DNA molecular of the aardvark, *Orycteropus afer*, and the position of the Tubulidentata in the eutherian tree. *Proc. R. Soc. Lond. B Biol. Sci.* **22**:339–345.
- ARNASON, U., A. GULLBERG, A. JANKE, and X. XU. 1996. Pattern and timing of evolutionary divergences among hominoids based on analyses of complete mtDNAs. *J. Mol. Evol.* **43**:650–661.

- ARNASON, U., A. GULLBERG, and B. WIDEGREN. 1991. The complete nucleotide sequence of the mitochondrial DNA of the fin whale, *Balaenoptera physalus*. *J. Mol. Evol.* **33**:556–568.
- ARNASON, U., A. GULLBERG, and X. XU. 1996. A complete mitochondrial DNA molecule of the white-handed gibbon, *Hylobates lar*, and comparison among individual mitochondrial genes of all hominoid genera. *Hereditas* **124**:185–189.
- ARNASON, U., and E. JOHNSON. 1992. The complete mitochondrial DNA sequence of the harbor seal, *Phoca vitulina*. *J. Mol. Evol.* **34**:493–505.
- ARNASON, U., X. XU, and A. GULLBERG. 1996. Comparison between the complete mitochondrial sequences of *Homo* and the common chimpanzee based on non-chimaeric sequences. *J. Mol. Evol.* **41**:952–957.
- BIBB, M. J., R. A. VAN ETTEN, C. T. WRIGHT, M. W. WALBERG, and D. A. CLAYTON. 1981. Sequence and gene organization of mouse mitochondrial DNA. *Cell* **26**:167–180.
- BULL, J. J., J. P. HUELSENBECK, C. W. CUNNINGHAM, D. L. SWOFFORD, and P. J. WADDELL. 1993. Partitioning and combining data in phylogenetic analyses. *Syst. Biol.* **42**:384–397.
- BUTLER, P. M. 1972. The problem of insectivore classification. Pp. 253–265 in K. A. JOYSEY and T. S. KEMP, eds. *Studies in vertebrate evolution*. Winchester Press, New York.
- . 1988. Phylogeny of the insectivores. Pp. 117–141 in M. J. BENTON, ed. *The phylogeny and classification of the tetrapods, Vol. 2. Mammals*. Clarendon Press, Oxford, England.
- CAO, Y., J. ADACHI, A. JANKE, S. PÄÄBO, and M. HASEGAWA. 1994. Phylogenetic relationships among eutherian orders estimated from inferred sequences of mitochondrial proteins: instability of a tree based on a single gene. *J. Mol. Evol.* **39**:519–527.
- CAO, Y., A. JANKE, P. J. WADDELL, M. WESTERMAN, O. TAKENAKA, S. MURATA, N. OKADA, S. PAABO, and M. HASEGAWA. 1998. Conflict amongst individual mitochondrial proteins in resolving the phylogeny of eutherian orders. *J. Mol. Evol.* **47**:307–322.
- CARROLL, R. L. 1988. *Vertebrate paleontology and evolution*. Freeman, New York.
- D'ERCHIA, A. M., C. GISSI, G. PESOLE, C. SACCONI, and U. ARNASON. 1996. The guinea-pig is not a rodent. *Nature* **381**:597–599.
- DE QUEIROZ, A., M. J. DONOGHUE, and J. KIM. 1995. Separate versus combined analysis of phylogenetic evidence. *Annu. Rev. Ecol. Syst.* **26**:657–681.
- FELSENSTEIN, J. 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. *J. Mol. Evol.* **17**:368–376.
- . 1991. PHYLIP (phylogeny inference package). Version 3.4. Distributed by the author, Department of Genetics, University of Washington, Seattle.
- FITCH, W. M. 1971. Toward defining the course of evolution, minimum change from a specific tree topology. *Syst. Zool.* **20**:406–415.
- FUMAGALLI, L., P. TABERLET, D. T. STEWART, L. GIELLY, J. HAUSSER, and P. VOGEL. 1999. Molecular phylogeny and evolution of *Sorex* shrews (Soricidae: Insectivora) inferred from mitochondrial DNA sequence data. *Mol. Phylogenet. Evol.* **11**:222–235.
- GADALETA, G., G. PEPE, G. DE CANDIA, C. QUAGLIARIELLO, E. SBISA, and C. SACCONI. 1989. The complete nucleotide sequence of the *Rattus norvegicus* mitochondrial genome: signals revealed by comparative analysis between vertebrates. *J. Mol. Evol.* **28**:497–516.
- GISSI, C., A. GULLBERG, and U. ARNASON. 1998. The complete mitochondrial DNA sequence of the rabbit, *Oryctolagus cuniculus*. *Genomics* **50**:161–169.
- GREGORY, W. K. 1910. The orders of mammals. *Bull. Am. Mus. Nat. Hist.* **27**:1–524.
- HASEGAWA, M., H. KISHINO, and T. YANO. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.* **22**:160–174.
- HIENDLEDER, S., H. LEWALSKI, R. WASSMUTH, and A. JANKE. 1998. The complete mitochondrial DNA sequence of the domestic sheep (*Ovis aries*) and comparison with the other major ovine haplotype. *J. Mol. Evol.* **47**:441–448.
- HUTCHINSON, J. H. 1968. Fossil Talpidae from the later Tertiary of Oregon. *Bull. Mus. Nat. Hist. Univ. Oreg.* **11**:1–117.
- JANKE, A., G. FELDMAIER-FUCHS, W. K. THOMAS, A. VON HAESELER, and S. PÄÄBO. 1994. The marsupial mitochondrial genome and the evolution of placental mammals. *Genetics* **137**:243–256.
- JANKE, A., N. J. GEMMELL, G. FELDMAIER-FUCHS, A. VON HAESELER, and S. PÄÄBO. 1996. The mitochondrial genome of a monotreme—the platypus (*Ornithorhynchus anatinus*). *J. Mol. Evol.* **42**:153–159.
- JANKE, A., X. XU, and U. ARNASON. 1997. The complete mitochondrial genome of the wallaroo (*Macropus robustus*) and the phylogenetic relationship among Monotremata, Marsupialia and Eutheria. *Proc. Natl. Acad. Sci. USA* **94**:1276–1281.
- KIM, K. S., S. E. LEE, H. W. JEONG, and J. H. HA. 1998. The complete nucleotide sequence of the domestic dog (*Canis familiaris*) mitochondrial genome. *Mol. Phylogenet. Evol.* **10**:210–220.
- 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.
- KISHINO, H., T. MIYATA, and M. HASEGAWA. 1990. Maximum likelihood inference of protein phylogeny and the origin of chloroplasts. *J. Mol. Evol.* **31**:153–160.
- KRETTEK, A., A. GULLBERG, and U. ARNASON. 1995. Sequence analysis of the complete mitochondrial DNA molecule of the hedgehog, *Erinaceus europaeus*, and the phylogenetic position of the Lipotyphla. *J. Mol. Evol.* **41**:952–957.
- LAVERGNE, A., E. DOUZERY, T. STICHLER, F. M. CATZEFLIS, and M. S. SPRINGER. 1996. Interordinal mammalian relationships: evidence for paenungulate monophyly provided by complete mitochondrial 12S rRNA sequences. *Mol. Phylogenet. Evol.* **6**:245–258.
- LOPEZ, J. V., M. CULVER, S. CEVARIO, and S. J. O'BRIEN. 1996. Complete nucleotide sequences of the domestic cat (*Felis catus*) mitochondrial genome and a transposed mtDNA repeat (Numt) in the nuclear genome. *Genomics* **33**:229–246.
- MCDOWELL, S. B. 1958. The Greater Antillean insectivores. *Bull. Am. Mus. Nat. Hist.* **115**:113–214.
- McKENNA, M. C. 1975. Toward a phylogeny and classification of the Mammalia. Pp. 21–46 in W. P. LUCKETT and F. S. SZALAY, eds. *Phylogeny of the Primates: a multidisciplinary approach*. Plenum Press, New York.
- McKENNA, M. C., and S. K. BELL. 1997. *Classification of mammals above the species level*. Columbia University Press, New York.
- MACPHEE, R. D. E., and M. J. NOVACEK. 1993. Definition and relationships of Lipotyphla. Pp. 13–31 in F. S. SZALAY, M. J. NOVACEK, and M. C. McKENNA, eds. *Mammal phylogeny: placentals*. Springer-Verlag, New York.
- NOVACEK, M. J. 1989. Higher mammal phylogeny: the morphological-molecular synthesis. Pp. 421–435 in B. FERN-

- HOLM, K. BREMER, and J. JÖRNVALL, eds. The hierarchy of life. Elsevier, Amsterdam.
- . 1990. Morphology, palaeontology, and the higher clades of mammals. Pp. 507–543 in H. H. GENOWAYS, ed. Current mammalogy. Vol. 2. Plenum Press, New York.
- . 1992. Mammalian phylogeny: shaking the tree. *Nature* **356**:121–125.
- NOVACEK, M. J., and A. R. WYSS. 1986. Higher-level relationships of the recent eutherian orders: morphological evidence. *Cladistics* **2**:257–287.
- NOVACEK, M. J., A. A. WYSS, and M. C. MCKENNA. 1988. The major groups of eutherian mammals. Pp. 31–71 in M. J. BENTON, ed. The phylogeny and classification of the Tetrapods. Clarendon Press, Oxford, England.
- OHDACHI, S., R. MASUDA, H. ABE, J. ADACHI, N. E. DOKUCHAEV, V. HAUKISALMI, and M. C. YOSHIDA. 1997. Phylogeny of Eurasian soricine shrews (Insectivora, Mammalia) inferred from the mitochondrial cytochrome b gene sequences. *Zool. Sci.* **14**:527–532.
- OJALA, D., J. MONTOYA, and G. ATTARDI. 1990. tRNA punctuation model of RNA processing in human mitochondria. *Nature* **290**:470–474.
- PENNY, D., and M. HASEGAWA. 1997. The platypus put in its place. *Nature* **387**:549–550.
- PUMO, D. E., P. S. FINAMORE, W. R. FRANEK, C. J. CARLETON, S. TARZAMI, and D. BALZARANO. 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.
- SAITOU, N., and M. NEI. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**:406–425.
- 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.
- STANHOPE, M. J., V. G. WADDELL, O. MADSEN, W. W. DE JONG, S. B. HEDGES, G. C. CLEVEN, D. KAO, and M. S. SPRINGER. 1998. Molecular evidence for multiple origins of Insectivora and for a new order of endemic African insectivore mammals. *Proc. Natl. Acad. Sci. USA* **17**:9967–9972.
- STRIMMER, K., and A. VON HAESELER. 1996. Quartet puzzling: a quartet maximum-likelihood method for reconstructing tree topologies. *Mol. Biol. Evol.* **13**:964–969.
- SULLIVAN, J., and D. L. SWOFFORD. 1997. Are guinea pigs rodents? The importance of adequate models in molecular phylogenies. *J. Mamm. Evol.* **4**:77–86.
- TEMPLETON, A. R. 1983. Phylogenetic inference from restriction endonuclease cleavage site maps with particular reference to humans and apes. *Evolution* **37**:221–244.
- URSING, B. M., and U. ARNASON. 1998a. The complete mitochondrial DNA sequence of the pig (*Sus scrofa*). *J. Mol. Evol.* **47**:302–306.
- . 1998b. Analyses of mitochondrial genomes strongly support a hippopotamus-whale clade. *Proc. R. Soc. Lond. B Biol. Sci.* **265**:2251–2255.
- 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.
- WERDELIN, L., and A. NILSSONNE. 1999. The evolution of the scrotum and testicular descent in mammals: a phylogenetic view. *J. Theor. Biol.* **196**:61–72.
- XU, X., A. GULLBERG, and U. ARNASON. 1996. The complete mitochondrial DNA (mtDNA) of the donkey and mtDNA comparisons among four closely related mammalian species-pairs. *J. Mol. Evol.* **43**:438–446.
- XU, X., and U. ARNASON. 1994. The complete mitochondrial DNA sequence of the horse, *Equus caballus*; extensive heteroplasmy of the control region. *Gene* **148**:357–362.
- . 1997. The complete mitochondrial DNA sequence of the white rhinoceros, *Ceratotherium simum*, and comparison with the mtDNA of the Indian rhinoceros, *Rhinoceros unicornis*. *Mol. Phylogenet. Evol.* **7**:189–194.
- XU, X., A. JANKE, and U. ARNASON. 1996. The complete mitochondrial DNA sequence of the greater Indian rhinoceros, *Rhinoceros unicornis*, and the phylogenetic relationship among Carnivora, Perissodactyla, and Artiodactyla (+ Cetacea). *Mol. Biol. Evol.* **13**:1167–1173.

PEKKA PAMILO, reviewing editor

Accepted September 7, 1999