The Complete Mitochondrial DNA Sequence of the Greater Indian Rhinoceros, *Rhinoceros unicornis,* and the Phylogenetic Relationship Among Carnivora, Perissodactyla, and Artiodactyla (+ Cetacea)

Xiufeng Xu, Axel Janke, and Ulfur Arnason

Division of Evolutionary Molecular Systematics, University of Lund, Sweden

The sequence (16,829 nt) of the complete mitochondrial genome of the greater Indian rhinoceros, *Rhinoceros unicornis,* was determined. Like other perissodactyls studied (horse and donkey) the rhinoceros demonstrates length variation (heteroplasmy) associated with different numbers of repetitive motifs in the control region. The 16,829-nt variety of the molecule includes 36 identical control region motifs. The evolution of individual peptide-coding genes was examined by comparison with a distantly related perissodactyl, the horse, and the relationships among the orders Carnivora, Perissodactyla, and Artiodactyla (+ Cetacea) were examined on the basis of concatenated sequences of 12 mitochondrial peptide-coding genes. The phylogenetic analyses grouped Carnivora, Perissodactyla, and Artiodactyla (+ Cetacea) into a superordinal clade and within this clade a sister group relationship was recognized between Carnivora and Perissodactyla to the exclusion of Artiodactyla (+ Cetacea). On the basis of the molecular difference between the rhinoceros and the horse and by applying as a reference the Artiodactyl/Cetacean divergence set at 60 million years ago (MYA), the evolutionary divergence between the families Rhinocerotidae and Equidae was dated to ≈50 MYA.

Introduction

Sequence data of mitochondrial DNA (mtDNA), notably the peptide-coding genes, have become a widely used tool for addressing phylogenetic relationships at various levels, especially among mammals. The importance of using comprehensive amounts of mitochondrial sequence data for inferring phylogenetic relationships was demonstrated by Cao et al. (1994), who, on the basis of analyses of complete mtDNAs, showed that different mtDNA genes provided different topologies for the ordinal relationship among Primates, Rodentia, Carnivora, Artiodactyla, and Cetacea. The topic that individual mtDNA genes (cytochrome *b,* COII) may provide different tree topologies has also been addressed by Honeycutt and Adkins (1993) and Honeycutt et al. (1995).

The mtDNA findings of Arnason and Johnsson (1992) and Janke et al. (1994) grouped Carnivora, Artiodactyla, and Cetacea into a superordinal clade, consistent with the nuclear data of Li et al. (1990) and Bulmer, Wolfe, and Sharp (1991). These analyses lacked, however, a suitable ungulate representation for conclusively assessing this relationship. The availability of the complete mtDNA of the horse (Xu and Arnason 1994) amended this shortcoming and a recent phylogenetic study including the horse and the hedgehog (Krettek, Gullberg, and Arnason 1995) identified the relationship among Perissodactyla (horse), Carnivora, and Artiodactyla (+ Cetacea) as essentially that of an unresolved triochotomy. The inclusion of the hedgehog was essential for these conclusions because of the proposal of phylogenetic affinities between Lipotyphla (hedgehog) and Carnivora (Miyamoto and Goodman 1986; MacPhee and Novacek 1993; Wyss and Flynn 1993). The analysis of Krettek, Gullberg, and Arnason (1995) yielded no support to this understanding, however, showing that the Lipotyphla (as represented by the hedgehog) had a basal position among the eutherians included.

We report now the complete mtDNA molecule of the greater Indian rhinoceros (hereafter referred to as Indian rhinoceros), *Rhinoceros unicornis,* which represents a family, Rhinocerotidae, that is distantly related to the Equidae. On the basis of this improved perissodactyl representation and by including also the mtDNA of the domestic cat (Lopez et al. 1996), we reexamine the mtDNA relationships among Carnivora, Perissodactyla, and Artiodactyla + Cetacea and propose a molecular dating of the divergence between the families Rhinocerotidae and Equidae.

Materials and Methods

Mitochondrial DNA was isolated from a frozen kidney sample of Indian rhinoceros ("Miris") that died in Zoologischer Garten Berlin AG, Germany. The sample was kindly provided by Dr. Reinhard Göltenboth. The isolation of mtDNA followed the same procedure as described in Arnason, Gullberg, and Widegren (1991). The enriched mtDNA was digested separately with *Bgl* II, *Bln* I, *Hin*dIII and *Spe* I. The products were ligated directly into M13 and cloned in *E. coli* JM109. Positive clones were identified by hybridization using mtDNA fragments of the horse and donkey as labeled probes. Sequencing of cloned fragments was performed on single-stranded DNA applying the dideoxy termination technique with [^{35}S]dATP. The work was performed manually using both universal and numerous specific oligonucleotide primers. The entire mtDNA molecule was covered by natural (not PCR) clones. Complementary to the natural clones the repetitive portion of the control region was sequenced after M13 cloning of PCR-amplified fragments in order to determine the numbers of repetitive motifs.

Key words: greater Indian rhinoceros, mtDNA, Perissodactyla, Rhinocerotidae, phylogeny.

Address for correspondence and reprints: Ulfur Arnason, 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. 13(9):1167–1173. 1996

Table 1
Percent Nucleotide Composition in Different Parts of the Mitochondrial Genome of Indian Rhinoceros

Nucleotide	Complete Molecule	Codon Position			12S rRNA Gene	16S rRNA Gene	TRNA Genes	Control Region
		1st	2nd	3rd				
A.....	33.6	32.1	19.6	42.5	37.9	38.2	33.4	32.5
G.....	12.7	20.4	12.1	5.6	16.7	15.5	19.0	13.6
C.....	27.5	25.2	26.2	33.3	23.5	21.4	17.8	27.2
T.....	26.2	22.3	42.1	18.6	21.9	24.9	29.8	26.7

The mtDNA sequence of the greater Indian rhinoceros has been deposited at the EMBL data bank with accession number X97336. Users of the sequence are kindly requested to refer to the present paper and not only to the accession number of the sequence.

Results and Discussion

General Features of the Mitochondrial Genome of Indian Rhinoceros

The length of the complete mtDNA sequence of the Indian rhinoceros, *Rhinoceros unicornis,* presented here is 16,829 base pairs. Like the mtDNAs of other perissodactyls, the control region of the Indian rhinoceros is characterized by tandemly arranged repetitive motifs. Therefore, the length of the molecule is not absolute. The nt composition of the L-strand and different features of the molecule is given in table 1. The nucleotide composition of different regions of the mitochondrial genome is consistent with that of other mammals (Janke et al. 1994). The underrepresentation of guanine in the L-strand is particularly pronounced at third codon position where all transitions are silent. Underrepresentation of this kind is not observable, however, in the control region.

The control region presently reported is 1,376 nt long with a continuous run of 36 identical repetitive motifs, 5′-CACATGTA. The repetitive motifs are located in the 3′ part of the control region in positions 16194–16481 of the complete sequence. The motif has the purine/pyrimidine alternation that characterizes most mammalian mtDNA motifs so far described (Ghivizzani et al. 1993; Hoelzel, Hancock, and Dover 1993; Xu and Arnason 1994). The number of repeats was determined in a total of 48 clones (23 natural and 25 PCR clones). The lowest number of repeats was 3 and the highest was 36. The characteristics of the repeat motifs in the three perissodactyl families, Rhinocerotidae, Tapiridae, and Equidae will be dealt with in a separate paper.

Two peptide-coding genes, NADH3 and NADH6, do not have a methionine start codon. Comparison with other mammalian mtDNAs suggests that the Indian rhinoceros has ATT (isoleucine) as the start codon of the NADH3 gene and GTG (valine) as the start codon of NADH6. Both these codons appear to be potential start codons in different mammalian mtDNAs (e.g., Xu and Arnason 1994). Three of the 13 peptide-coding genes, COIII, NADH3, and NADH4, are not terminated by a complete stop codon. Among mammals the occurrence of a complete stop codon in COIII has only been reported in the fin (Arnason, Gullberg, and Widegren 1991) and the blue (Arnason and Gullberg 1993) whales, and NADH3 is terminated by a complete stop codon only in the mouse (Bibb et al. 1981) and the rat (Gadaleta et al. 1989). A complete stop codon in NADH4 has not been described so far in any mammalian mtDNA. The COIII, NADH3, and NADH4 genes of the Indian rhinoceros have incomplete stop codons (T in COIII and NADH4, and TA in NADH3), consistent with the findings that the transcripts of such genes contain a stop codon created by posttranscriptional polyadenylation (Ojala, Montoya, and Attardi 1981).

The boundaries of the 22 tRNA genes of the mtDNA of the Indian rhinoceros were determined by analogy with the tRNA genes of different eutherians. Twenty of the tRNA genes of the rhinoceros mtDNA have the standard secondary structure discussed by Kumazawa and Nishida (1993). The features of these tRNAs, as well as those of the structurally atypical tRNA-Ser(AGY) and tRNA-Ser(UCN), conform with those described for other eutherians.

Comparison with the mtDNA of the Horse

In the present study we give an account of the molecular difference between two distantly related species of the order Perissodactyla, the Indian rhinoceros and the horse. The percentage nt composition of the mtDNA (L-strand) outside the control region of rhinoceros/horse is similar: A = 33.8/32.5; C = 27.5/28.4; G = 12.6/13.3; T = 26.1/25.8. The corresponding percentages for the control region are: A = 32.5/27.3; C = 27.2/30.9; G = 13.6/15.5; T = 26.7/26.3. Alignment of the mtDNA molecules of the rhinoceros and the horse, less the control region, shows a sequence difference of 15.6%. In addition to indel (insertion/deletion) differences in genic regions this alignment shows three indels in nongenic regions, one at the junction between tRNA-Ser(UCN) and tRNA-Asp, one between tRNA-Arg and NADH4L, and one in the loop of origin of L-strand replication.

The length of the NADH2 gene of the rhinoceros is 1,044 nt, 3 nt (one codon) more than that of the horse. Other peptide-coding genes have the same length in the two species. The nt differences (both total and conservative) were detailed according to codon position and type of substitution (transition or transversion) (table 2). The 13 peptide-coding genes differ by a total of 1.923 substitutions, 16.9%. The ratio for these differences with respect to codon positions 1, 2, and 3 is 2.3:1:8.9, as compared with 3.5:1:25.7 in the intrageneric comparison between the horse and the donkey (Xu, Gullberg, and Arnason 1996). The total number of conservative nt substitutions is 960 (8.4%). The codon position ratio for these differences is 1.6:1:3.6, as compared with 1.5:1:2.1 in the horse/donkey comparison. The two sets of ratios for conservative nt substitutions show that the accumulation of substitutions in second codon position and the accumulation of nonsynonymous substitutions in first codon position have the same rate in the close (horse/donkey) and the distant (horse/rhinoceros) rela-

Table 2
Nucleotide Difference with Respect to Codon Position (1, 2, 3) Between Each of the 13 Mitochondrial Protein-Coding Genes of Indian Rhinoceros and Horse and the Number of Amino Acid Differences

			1					3					No. of Amino Acid
	Length (bp)		Ti			2		Ti		Tv		In-	Differ-
Gene	Rhino	Horse	a	b	Tv	Ti	Tv	a	b	a	b	del	ences
NADH1	957	957	15	12	7	8	1	7	59	14	33		25
NADH2	1,044	1,041	8	14	19	18	3	8	53	9	46	1	53
COI	1,545	1,545	18	7	2	2	—	12	134	19	49		9
COII	684	684	14	6	2	1	—	6	47	10	22		8
ATPase8	204	204	1	3	8	4	2	2	18	—	7		17
ATPase6	681	681	9	5	7	7	—	6	42	10	18		19
COIII	783	783	6	5	5	4	2	5	56	10	33		14
NADH3	345	345	2	3	3	3	1	3	15	8	9		12
NADH4L	297	297	6	6	2	—	2	4	19	1	7		9
NADH4	1,377	1,377	15	25	12	18	7	14	94	18	51		54
NADH5	1,821	1,821	12	33	16	35	9	16	110	24	79		86
NADH6	528	528	2	19	4	13	4	5	26	—	27		38
Cyt *b*	1,140	1,140	13	13	7	10	3	10	71	16	38		28
Total	11,406	11,403	121	151	94	123	34	98	744	139	419	1	372
Cons. diff.				245		157				588			
Total diff.				366		157			1,400			1	
Ratio total diff.					2.3		1.0			8.9			
Ratio cons. diff.					1.6		1.0			3.6			

Note.—Ti: transitions; Tv: transversion; a: number of differences involving leucine in both species; b: number of differences other than those involving leucine; cons. diff: conservative difference defined as all nonsynonymous nucleotide substitutions in first codon position, all substitutions in second codon position and transversions in third codon position.

tionships and that this rate is considerably slower than that for accumulation of transversions in third codon position. The difference between the two sets of ratios for total nt substitution shows that there is a high degree of transition saturation in third codon position in rhinoceros/horse relative to horse/donkey. The findings suggest that there is a high degree of saturation in first codon position synonymous transitions (leucine) in the rhinoceros/horse comparison.

Table 3
Percent Amino Acid, Conservative Nucleotide (Cons. nt), and Total Nucleotide Differences Between Mitochondrial Peptide-Coding Genes of Indian Rhinoceros and Horse

Gene	Amino Acid	Cons. nt	Total nt
COI	1.8	5.1	15.7
COII	3.5	6.0	15.8
COIII	5.4	7.5	16.1
Cyt *b*	7.4	7.6	15.9
NADH1	7.8	7.8	16.3
ATPase6	8.4	6.9	15.3
NADH4L	9.1	6.1	15.8
NADH3	10.4	7.8	13.6
NADH4	11.8	9.5	18.4
NADH5	14.2	10.8	18.3
NADH2	15.5	10.4	17.2
NADH6	21.7	12.7	18.9
ATPase8	25.0	11.8	22.1

Table 3 shows the results of a pairwise comparison between the 13 mitochondrial peptide-coding genes of the Indian rhinoceros and the horse. The table shows, for each gene, the percent total amino acid (aa) difference, the percent conservative nt difference (Irwin, Kocher, and Wilson 1991), and the percent nt difference. The genes have been arranged according to increasing aa difference. Percent total aa difference was in the range 1.8 (COI) to 25.0 (ATPase8). For conservative nt substitutions the range was 5.1% (COI) to 12.7% (NADH6). For total nt difference the corresponding range was 13.6% (NADH3) to 22.1% (ATPase8). The order of the genes, with respect to increasing difference, is highly similar between the two conservative modes of comparison, aa difference, and conservative nt substitutions. The amplitude for total nt difference is relatively limited among the 13 genes, consistent with (primarily) third codon position mutational saturation, and the order based on percent total nt substitution deviates to some extent from the order based on the other two approaches, which are more conservative.

The results of a pairwise comparison between each of the 22 tRNA genes of the Indian rhinoceros and the horse are shown in table 4. The differences are detailed according to position in secondary structures. There are a total of 169 differences (11.1%), 129 transitions (8.5%), 28 transversions (1.8%), and 12 indels (0.8%).

Table 4
Nucleotide Differences Between the 22 Mitochondrial tRNA Genes of Indian Rhinoceros and Horse Based on Inferred Secondary Structure

	AA Stem		D Stem		AC Stem		T Stem		D Loop			AC Loop		Extra Loop		T Loop			Others	Total		
Gene	Ti	Tv	Ti	Tv	Ti	Tv	Ti	Tv	Ti	Tv	Gap	Ti	Tv	Ti	Tv	Ti	Tv	Gap	(Ti)	Ti	Tv	Gap
tRNA-Phe								3	1		1	1					1			3	4	1
tRNA-Val					4		1	1				1				1	1			6	2	
tRNA-Leu(UUR)	2		1							1				2	1					5	2	
tRNA-Ile	1														2	2				3	2	
tRNA-Gln	4				3	1	2		1			1								11	1	
tRNA-Met									1							1	1			2	1	
tRNA-Trp					1		2		1		2						1	1		4	1	3
tRNA-Ala	3				1		3			1		1								8	1	
tRNA-Asn														1		2				3		
tRNA-Cys	3		1		1						1			1		1		1	1	8		2
tRNA-Tyr	5						1			1						2	1			8	2	
tRNA-Ser(UCN)					1					1										1	1	
tRNA-Asp	2				2					1				1					1	6	1	
tRNA-Lys	3		1		2		2		1		1			2		3	1		1	15	1	1
tRNA-Gly									2	1							1			2	2	
tRNA-Arg	2										1			2		1	1			5	1	1
tRNA-His	2				2		1					1		1	1	1	1			8	2	
tRNA-Ser(AGY)[a]	1				1		2									2		1	2	8		1
tRNA-Leu(CUN)									1										1	2		
tRNA-Glu	1				2		2		1							2	1			8	1	
tRNA-Thr	4	1					2		3		1					1	2	2		10	3	3
tRNA-Pro							2									1				3		
Total	33	1	3		20	1	20	4	12	6	7	5		10	4	20	12	5	6	129	28	12

NOTE.—AA, D, AC and T: the stem region of amino acid acceptor, dihydrouridine, anticodon and TψC, respectively.
[a] There is no D stem in the inferred secondary structure of tRNA-Ser(AGY).

Very few differences (all transitions) occur in the D stem and in the AC loop. The alignment between the tRNA genes for cysteine, lysine, and leucine(CUN) shows a single transition at the junction between the D and AC stems. The difference between the two species is particularly marked in tRNA-Lys (17 differences) and tRNA-Thr (16 differences). The most conservative tRNAs are tRNA-Leu(CUN) (two substitutions), tRNA-Pro (three substitutions), and tRNA-Met (three differences). The transition/transversion (Ti/Tv) ratio for the combined length of the tRNA genes is 4.6.

The tRNA-Pro and tRNA-Phe genes of the black rhinoceros were reported by Jama et al. (1993). In the tRNA-Pro gene the two rhinoceroses differ by five transitions, two more than horse and Indian rhinoceros. In the tRNA-Phe gene the two rhinoceroses differ by seven transitions and one transversion, as compared with three transitions, four transversions and one indel (insertion/deletion) between the Indian rhinoceros and the horse.

Table 5
Difference Between the 12S and 16S rRNA Genes of Indian Rhinoceros and Horse

	Length (nt) (Rhino/Horse)	Percent nt Difference	No. of Transitions (%)	No. of Transversions (%)	No. of Indels (%)
12S rRNA	971/975	13.1	92 (9.4)	27 (2.8)	9 (0.9)
16S rRNA	1,577/1,581	11.3	102 (6.4)	57 (3.6)	21 (1.3)
Total	2,548/2,556	12.0	194 (7.5)	84 (3.3)	30 (1.2)

The 12S and 16S rRNA genes of the Indian rhinoceros and the horse differ at 128 (13.1%) and 180 (11.3%) positions, respectively. The differences are detailed in table 5. For the combined length of the rRNA genes the Ti/Tv ratio is 2.3.

The Phylogenetic Relationships of Carnivora, Perissodactyla, Artiodactyla, and Cetacea

Phylogenetic analyses based on single (Arnason and Johnsson 1992) and combined sequences of all (Janke et al. 1994) mtDNA protein-coding genes have grouped Carnivora and Artiodactyla (+ Cetacea) into a superordinal clade. The statistical support for this relationship was detailed by Janke et al. (1994). The inclusion of a perissodactyl representative, the horse (Krettek, Gullberg, and Arnason 1995), has shown a much closer relationship among Carnivora, Perissodactyla, and Artiodactyla (+ Cetacea) than is generally recognized by classical approaches. In the present study we readdress these relationships by complementing the sampling with an additional perissodactyl, the Indian rhinoceros (family Rhinocerotidae), which is distantly related to the Equidae (horse), and a carnivore, the domestic cat (Lopez et al. 1996), of a family (Felidae) that is distantly related to the seals (family Phocidae). Thus the analysis

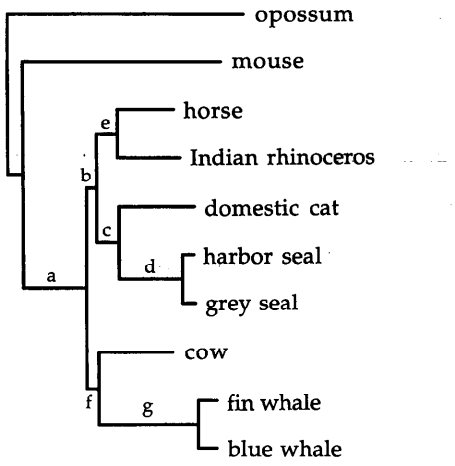

FIG. 1.—Phylogenetic relationships among ferungulates (carnivores, perissodactyls, artiodactyls (+ cetaceans) established by neighbor-joining (NJ) amino acid sequence analysis. The same topology was obtained by maximum-likelihood (ML) analysis of first codon position nonsynonymous substitutions + all second codon position substitutions. Ferungulate branches are designated with lowercase letters (a–g). The bootstrap support for different branches, detailed according to analytical approach, is given in table 6. The analysis grouped Carnivora/Perissodactyla/Artiodactyla (+ Cetacea) into a superordinal clade. Within this clade the carnivores (domestic cat, harbor and grey seals) grouped with the perissodactyls (horse, Indian rhinoceros), and the artiodactyls (cow) grouped with the cetaceans (fin and blue whales).

included the following species: opossum, mouse, cow, fin whale, blue whale, domestic cat, harbor seal, grey seal, Indian rhinoceros, and horse. The phylogenetic analysis was based on the concatenated sequences of 12 peptide-coding mitochondrial genes. The NADH6 gene was not included because of the different nt composition of this gene relative to the other peptide-coding genes.

The phylogenetic analyses were performed both on nt and aa sequences applying different analytical approaches. Maximum-parsimony (MP) and neighbor-joining (NJ) analyses were carried out with the PHYLIP program package (Felsenstein 1991), whereas the maximum-likelihood (ML) analyses were performed with both the PHYLIP (DNAML) and the MOLPHY (Adachi and Hasegawa 1995) program packages (PROTML and NUCML). At the nt level the analysis was based on nonsynonymous substitutions at first codon position, all substitutions at second codon position, and transversions at third codon position. First, second, first + second, and first + second + third codon positions were analyzed separately in order to examine the degree of support provided by each character set for a particular topology. Except for NJ analysis of second codon position all phylogenetic analysis (ML analysis not shown) resulted in the topology shown in figure 1. The bootstrap support for individual branches (a–g) is summarized in table 6.

The phylogenetic analyses supported a sister group relationship between Carnivora and Perissodactyla although the support for this relationship was not significant in all modes of analysis. This relationship was examined further by ML analysis of the aa sequences (PROTML with mtREV matrix). The analysis was undertaken because, relative to nt analysis, this mode of analysis is relatively insensitive to mutational and compositional effects. The three alternative hypotheses of the relationship between the carnivores and the ungulates were tested (table 7). The sister group relationship between the Carnivora and the Perissodactyla received by far the largest support also in this approach, although the traditional view of carnivores being an outgroup to ungulates (+ cetaceans) could not be rejected at the 5% level. The overall good, and in the case of the NJ analysis of aa sequences significant, support for a carnivore/perissodactyl sister group relationship contradicts the classical view of ungulate relationships. The position of the cow was somewhat labile in the MP tree, but, consistent with previous studies (Janke et al. 1994), its grouping with the cetaceans was significantly supported by the NJ analysis. It is possible that the different support for the position of the cow, provided by different approaches, is a reflection of the somewhat slower evolutionary rate of the cow relative to the other species included in the analyses. If this is the case it suggests that NJ analysis is less sensitive to differences in evolutionary rates than is MP analysis.

Table 6
Bootstrap Values for the Branches (a–g) Shown in Figure 1

	a	b	c	d	e	f	g
NJ							
1.....	100	76.6	100	100	100	96.1	100
2.....	—	—	—	—	—	—	—
1, 2.....	100	74.8	100	100	100	100	100
1, 2, 3..	100	76.4	100	100	100	100	100
MP							
1.....	100	59.3	97.9	97.9	100	65.5	100
2.....	100	33.3	73.2	96.0	100	38.2	100
1, 2.....	100	56.7	100	99.6	100	60.4	100
1, 2, 3..	100	77.9	99.9	98.9	100	77.5	100
aa MP....	100	82.9	97.9	99.3	100	46.8	100
aa NJ	100	98.0	100	100	100	94.0	100

NOTE.—Bootstrap values were obtained by different tree constructing methods, neighbor joining (NJ), and maximum parsimony (MP), for different combinations of codon positions (1, 2, 3), and for amino acid (aa) sequences. Except for NJ analysis of second codon position, all reconstructions resulted in the same best tree as shown in figure 1. The bootstrap values in percent were obtained from 1,000 replications for nucleotide sequences and from 500 replications for the aa sequences. Analysis of first codon position did no include synonymous leucine transitions, while that of third codon position included only transversions.

Dating of the Evolutionary Divergence Between Rhinocerotidae and Equidae

Irrespective of the quality of fossil data, evolutionary separations are impossible to date precisely on the

Table 7
Maximum-Likelihood Analysis of Ferungulate Relationships

	ΔlnL	SE	pBoot
((CAR,PER),(ART,CET))	[−22846.4]		90.80
(CAR,(PER,(ART,CET)))	−24.6	±17.5	8.80
(PER,(CAR,(ART,CET)))	−36.6	±15.6	0.40

NOTE.—Maximum-likelihood analysis of different combinations of the relationships among carnivores (CAR), perissodactyls (PER), artiodactyls (ART) and cetaceans (CET). The differences of the log likelihood values (ΔlnL) to that of best tree and their standard error as well as the bootstrap probabilities (pBoot) are shown. The bracketed value gives the log likelihood value of the best tree.

Table 8
Amino Acid Differences and Distances Among and Within Orders

	Dvi	Mmu	Eca	Run	Fca	Pvi	Hgr	Bta	Bph	Bmu
Dvi	—	781	717	732	733	739	741	715	760	764
Mmu ...	0.302	—	643	634	647	666	667	648	702	695
Eca	0.273	0.235	—	239	332	351	351	325	412	410
Run	0.279	0.233	0.078	—	349	438	346	349	424	424
Fca.....	0.278	0.239	0.111	0.117	—	297	297	364	474	477
Pvi.....	0.282	0.247	0.119	0.118	0.098	—	50	364	453	454
Hgr	0.283	0.248	0.119	0.117	0.098	0.016	—	376	459	464
Bta.....	0.272	0.238	0.109	0.117	0.124	0.123	0.128	—	375	371
Bph	0.294	0.264	0.143	0.146	0.166	0.159	0.161	0.128	—	78
Bmu....	0.296	0.261	0.143	0.147	0.168	0.160	0.163	0.126	0.025	—

NOTE.—Order Marsupialia: Dvi, opossum. Order Rodentia: Mmu, mouse. Order Carnivora: Pvi, harbor seal; Hgr, grey seal; Fca, domestic cat. Order Cetacea: Bph, fin whale, Bmu, blue whale. Order Artiodactyla: Bta, cow. Order Perissodactyla: Eca, horse; Run, Indian rhinoceros. Values are based on the concatenated sequences, 3,033 aa, of 11 mitochondrial peptide-coding genes. The lengths do not include gaps or ambiguous aa alignments adjacent to gaps. Above diagonal: total number of differences. Below diagonal: distances according to the ML mtREV model.

basis of the paleontological record. The reason for this is that a certain time of evolutionary separation is necessary before the development of morphological traits that can be recognized among fossil finds, even when the paleontological record is reasonably complete. The effects of incomplete paleontological records for the dating of evolutionary divergences have been addressed by Martin (1993). Although the author primarily addressed primate evolution the conclusions are of a general relevance, suggesting that paleontological finds will generally grossly underestimate the age of evolutionary divergences.

It is conceivable that radical ecological shifts will promote the development of morphological characteristics in such a way that the time between evolutionary separation and the development of these characteristics will be shorter than in cases where no drastic ecological shifts have taken place. Among the mammals the most drastic ecological shifts are probably the transitions from terrestrial to aquatic life. The divergence of the archeocetes from their artiodactyl relatives is not the only event of this kind, but the origin of the archeocetes is probably better documented by paleontological finds than any other similar event. The oldest archeocete fossils are 52–54 million years old (Gingerich et al. 1994; Thewissen, Hussain, and Arif 1994). On the basis of the age of these fossils and analysis of the complete cytochrome *b* gene of more than 30 cetaceans plus several artiodactyls, it has been proposed that Ceatacea and Artiodactyla separated ≈60 MYA (Arnason and Gullberg 1996).

The distances among eight species representing Perissodactyla (rhinoceros, horse), Carnivora (harbor seal, grey seal, domestic cat), Artiodactyla (cow), and Cetacea (fin whale, blue whale) plus the mouse (Rodentia) and the opossum (Marsupialia), are given in table 8. The comparison is based on the concatenated inferred protein sequence of 12 protein-coding mtDNA genes, excluding NADH6, which is encoded by a different strand relative to the other protein-coding genes. Based on an artiodactyl/cetacean separation 60 MYA, and taking into account the faster evolution of cetacean mtDNA, the values in table 8 suggest that the Rhinocerotidae and the Equidae had a last common ancestor 50 MYA (95% confidence limits 43.5–56.5 MYA).

The sequence of the cytochrome *b* gene of the black rhinoceros was reported by Irwin, Kocher, and Wilson (1991). The difference (aa, conservative nt) between the two rhinoceroses suggests that they diverged ≥30 MYA. This dating, however, should be considered as tentative until supported by additional data.

The presently proposed dating for the evolutionary separation between Rhinocerotidae and Equidae is consistent with the ≈50 MYA palaeontological dating (Prothero and Schoch 1989) of the evolutionary separation between the perissodactyl suborders Hippomorpha and Ceratomorpha. The Ceratomorpha includes rhinoceroses and tapirs, and a more precise molecular dating of the separation of the hippomorph Equidae and the ceratomorph Rhinocerotidae should, therefore, also include tapirid representation. The necessity of including tapir data in a comparison of this kind is exemplified by the fact that two molecular studies, one on restriction site mapping of the alpha globin gene cluster (Flint, Ryder, and Clegg 1990) and one on aa sequence data of pancreatic polypeptide (Henry, Lance, and Conlon 1991), do not recognize tapirs and rhinoceroses as sister groups to the exclusion of the Equidae.

Acknowledgments

We express our sincere thanks to Dr. Reinhard Göltenboth, Zoologischer Garten Berlin AG, for tissue samples. The work was supported by the Swedish Natural Sciences Research Council and the Erik Philip-Sörensen Foundation.

LITERATURE CITED

ADACHI, J., and M. HASEGAWA. 1995. MOLPHY: programs for molecular phylogenetics. Version 2.3. Institute of Statistical Mathematics, Tokyo.

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, E. JOHNSSON, and C. LEDJE. 1993. The nucleotide sequence of the mitochondrial DNA molecule of the grey seal, *Halichoerus grypus,* and a comparison with mitochondrial sequences of other true seals. J. Mol. Evol. **37**:323–330.

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., and E. JOHNSSON. 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.

BULMER, M., K. H. WOLFE, and P. M. SHARP. 1991. Synonymous nucleotide substitution rates in mammalian genes: implication for the molecular clock and the relationship of mammalian orders. Proc. Natl. Acad. Sci. USA **88**:5974–5978.

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.

FELSENSTEIN, J. 1991. PHYLYP: phylogeny interference package. Version 3.4. University of Washington, Seattle.

FLINT, J., O. A. RYDER, and J. B. CLEGG. 1990. Comparison of the alpha globin gene cluster structure in Perissodactyla. J. Mol. Evol. **30**:36–42.

GADALETA, G., G. PEPE, G. DE CANDIA, C. QUAGLIARIELLO, E. SBISA, and C. SACCONE. 1989. The complete nucleotide sequence of the *Rattus norvegicus* mitochondrial genome: cryptic signals revealed by comparative analysis between vertebrates. J. Mol. Evol. **28**:497–516.

GHIVIZZANI, S. C., S. L. D. MACKAY, C. S. MADSEN, P. J. LAIPIS, and W. W. HAUSWIRTH. 1993. Transcribed heteroplasmic repeated sequences in the porcine mitochondrial DNA D-loop region. J. Mol. Evol. **37**:36–47.

GINGERICH, P. D., S. M. RAZA, M. ARIF, and X. Y. ZHOU. 1994. New whale from the Eocene of Pakistan and the origin of cetacean swimming. Nature **368**:844–847.

HENRY, J. S., V. A. LANCE, and J. M. CONLON. 1991. Primary structure of pancreatic polypeptide from four species of Perissodactyla, Przewalski's horse, zebra, rhino, tapir. Gen. Comp. Endocrinol. **84**:440–446.

HOELZEL, A. R., J. M. HANCOCK, and G. A. DOVER. 1993. Generation of VNTRs and heteroplasmy by sequence turnover in the mitochondrial control region of two elephant seal species. J. Mol. Evol. **37**:190–197.

HONEYCUTT, R. L., and R. M. ADKINS. 1993. High level systematics of eutherian mammals: an assessment of molecular characters and phylogenetic hypotheses. Annu. Rev. Ecol. Syst. **24**:279–305.

HONEYCUTT, R. L., M. A. NEDBAL, R. M. ADKINS, and L. L. JANECEK. 1995. Mammalian mitochondrial DNA evolution: a comparison of the cytochrome *b* and cytochrome *c* oxidase II genes. J. Mol. Evol. **40**:260–272.

IRWIN, D. M., T. D. KOCHER, and A. C. WILSON. 1991. Evolution of the cytochrome *b* gene of mammals. J. Mol. Evol. **32**:128–144.

JAMA, M., Y. ZHANG, R. A. AMAN, O. R. RYDER. 1993. Sequence of the mitochondrial control region, $tRNA^{Thr}$, $tRNA^{Pro}$ and $tRNA^{Phe}$ genes from the black rhinoceros, *Diceros bicornis.* Nucleic Acids Res. **21**:4392.

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.

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.

KUMAZAWA, Y., and M. NISHIDA. 1993. Sequence evolution of mitochondrial tRNA genes and deep-branch animal phylogenetics. J. Mol. Evol. **37**:380–398.

LI, W.-H., M. GOUY, P. M. SHARP, C. O'HUIGIN, and Y.-W. YANG. 1990. Molecular phylogeny of Rodentia, Lagomorpha, Primates, Artiodactyla and Carnivora and molecular clocks. Proc. Natl. Acad. Sci. USA **87**:6703–6707.

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.

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.

MARTIN, R. D. 1993. Primate origins: plugging the gaps. Nature **363**:223–234.

MIYAMOTO, M. M., and M. GOODMAN. 1986. Biomolecular systematics of eutherian mammals: phylogenetic patterns and classification. Syst. Zool. **35**:230–240.

OJALA, D., J. MONTOYA, and G. ATTARDI. 1981. tRNA punctuation model of RNA processing in human mitochondria. Nature **290**:470–474.

PROTHERO, D. R., and R. M. SCHOCH. 1989. Origin and evolution of the Perissodactyla: summary and synthesis. Pp. 504–529 *in* D. R. PROTHERO and R. M. SCHOCH, eds. The evolution of the perissodactyls. Oxford University Press, New York.

SIMPSON, G. G. 1945. The principles of classification and a classification of mammals. Bull. Am. Mus. Nat. Hist. **85**: 1–350.

THEWISSEN, J. G. M., S. T. HUSSAIN, and M. ARIF. 1994. Fossil evidence for the origin of aquatic locomotion in archaeocete whales. Science **263**:210–212.

WYSS, A. R., and J. J. FLYNN. 1993. A phylogenetic analysis and definition of the carnivora. Pp. 32–52 *in* F. S. SZALAY, M. J. NOVACEK, and M. C. MCKENNA, eds. Mammal phylogeny: placentals. Springer-Verlag, New York.

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.

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. (in press).

RODNEY L. HONEYCUTT, reviewing editor

Accepted July 18, 1996