MOLECULAR PHYLOGENETICS AND EVOLUTION
Vol. 7, No. 2, April, pp. 189–194, 1997
ARTICLE NO. FY960385

The Complete Mitochondrial DNA Sequence of the White Rhinoceros, *Ceratotherium simum,* and Comparison with the mtDNA Sequence of the Indian Rhinoceros, *Rhinoceros unicornis*

Xiufeng Xu and Ulfur Arnason[1]

Division of Evolutionary Molecular Systematics, University of Lund, Sölvegatan 29, S-223 62 Lund, Sweden

Received June 21, 1996; revised September 10, 1996

The complete nucleotide sequence of the mitochondrial genome of the white rhinoceros, *Ceratotherium simum*, was determined. The length of the reported sequence is 16,832 nucleotides. This length can vary, however, due to pronounced heteroplasmy caused by differing numbers of a repetitive motif (5′-CG-CATATACA-3′) in the control region. The 16,832 nucleotide sequence presented here is the longest version of the molecule and contains 35 copies of this motif. Comparison between the complete mitochondrial sequences of the white and the Indian (*Rhinoceros unicornis*) rhinoceroses allowed an estimate of the date of the basal evolutionary divergence among extant rhinoceroses. The calculation suggested that this divergence took place approximately 27 million years before present.

INTRODUCTION

The perissodactyl family Rhinocerotidae includes four recent genera, *Rhinoceros* (Indian and Javan rhinoceroses), *Dicerorhinus* (Sumatran rhinoceros), *Diceros* (black rhinoceros), and *Ceratotherium* (white rhinoceros). *Rhinoceros* and *Dicerorhinus* occur in Asia, whereas *Diceros* and *Ceratotherium* are found in Africa. The family Rhinocerotidae is distinct and well defined, but systematic relationships among the four genera (reviewed by Morales and Melnick, 1994) are controversial. Evolutionary relationships within the Rhinocerotidae have generally been evaluated on the basis of geographical distribution of the different genera or on the number of horns (one or two) characterizing these genera. In the present study we describe the complete mtDNA sequence of a two-horned species, the white rhinoceros, and compare it in detail with that of a one-horned species, the Indian rhinoceros (Xu *et al.*, 1996b). This comparison makes it possible to address the date of the basal rhinocerotid divergence irrespective of whether it is defined by geographical distribution or number of horns. The date of divergence between the two rhinoceroses was estimated using the evolutionary separation of Artiodactyla and Cetacea set at 60 million years before present, MYBP, as a reference (Arnason and Gullberg, 1996). The position of the black rhinoceros, *Diceros bicornis*, relative to the other two species was also examined; however, only a limited amount of sequence data is currently available for this species compared to the Indian and the white rhinoceroses.

MATERIALS AND METHODS

A total amount of 1.5 g of striated muscle from the white rhinoceros, *Ceratotherium simum*, was kindly provided by Dr. Peter Arctander, Department of Zoology, Univ. of Copenhagen, Denmark. An enriched mtDNA fraction was isolated as described by Arnason *et al.* (1991). The enriched mtDNA was digested separately with *Hin*dIII, *Spe*I, *Xba*I, *Bln*I, *Bgl*II, and *Bcl*I. The products were ligated directly into M13 and cloned in *Escherichia coli* JM101. Positive clones were identified through hybridization, using mtDNA fragments from the horse and the donkey as labeled probes. The clones covered the entire molecule except for a 3,971-nucleotide (nt)-long region located between positions 11,622 and 15,592 of the complete sequence. This region was PCR amplified in two separate reactions prior to cloning. Six individual PCR clones from each reaction (all identical) were then used to determine the sequence of the region.

The sequence of the control region was determined by sequencing two natural (not PCR) clones. In addition, the number and sequence of repeat motifs in this region were determined from a total of 58 clones derived from PCR amplification. Sequencing was according to the dideoxy termination technique (Sanger, 1977) with [^{35}S]dATP using both universal and numerous specific oligonucleotide primers.

Handling of sequences and alignments were per-

[1] To whom correspondence and reprint requests should be addressed. Fax: 46 46 147874. E-mail: Ulfur.Arnason@gen.lu.se.

1055-7903/97 $25.00

formed with the GCG (1994) program package. Insertion/deletion (indel) differences were counted as single events irrespective of their lengths. Conservative nucleotide changes (Irwin *et al.*, 1991) included all substitutions at the 1st codon position except leucine transitions, all substitutions at the 2nd codon position, and transversions at the 3rd codon position.

The mtDNA sequence of the white rhinoceros has been deposited at EMBL with Accession No. Y07726. Users of this sequence are kindly requested to refer to the present paper in addition to the accession number of the sequence.

RESULTS

The length of the mtDNA sequence presented here for the white rhinoceros, *Ceratotherium simum*, is 16,832 nt. As with other perissodactyl species for which complete mtDNA sequences are available—horse (Xu and Arnason, 1994), donkey (Xu *et al.*, 1996a), and Indian rhinoceros (Xu *et al.*, 1996b)—the control region of the white rhinoceros contains variable numbers of repeat motifs arranged in tandem. The lengths of individual mtDNA sequences of the white rhinoceros can therefore vary (heteroplasmy). The base composition of the L-strand, excluding the control region, is A, 33.4%; C, 27.9%; G, 12.9%; and T, 25.8%. In the control region of the reported sequence (35 repeats) the corresponding values are A, 33.7%; C, 28.9%; G, 12.0%; and T, 25.4%.

Positions of protein-coding genes were determined by the occurrence of start and stop codons and by analogy with other complete eutherian mtDNA sequences. The start codon of the NADH3 gene is ATT (isoleucine). All other genes have a methionine start codon. Three protein-coding genes—COIII, NADH3, and NADH4—have incomplete stop codons (TA or T). In all three cases, the terminal nt is contiguous to the 5′ terminal nt of a tRNA gene. As discussed by Wolstenholme (1992) and consistent with the findings of Ojala *et al.* (1981), the transcripts of such genes contain a stop codon created by posttranscriptional polyadenylation. Among other eutherians studied so far, only the fin and the blue whales (Arnason *et al.*, 1991; Arnason and Gullberg, 1993) have a complete stop codon in COIII, while the NADH3 gene is terminated by a complete stop codon only in the mouse (Bibb *et al.*, 1981), the rat (Gadaleta *et al.*, 1989), and the hedgehog (Krettek *et al.*, 1995). A complete stop codon has not yet been described in the NADH4 gene of any eutherian.

The control region of the sequence reported here is 1,381 nt long with a continuous run of 35 repeat motifs (5′-CGCATATACA-3′). These repeats are located in the 3′ part of the control region, between positions 16,187 and 16,536 of the complete sequence. Like most repeat motifs in eutherian control regions, the white rhinoceros motif is characterized by a purine/pyrimidine alternation (Ghivizzani *et al.*, 1993; Xu and Arnason, 1994). In order to determine the range of variation among different control regions, we sequenced 60 different clones (two natural plus 58 PCR clones) of the repeat part of the control region. The number of motifs among these clones varied from 10 to 35. Most sequences fall in the middle of the range of repeat numbers (see Fig. 1). The control region of the white rhinoceros, like that of the horse (Xu and Arnason, 1994), is characterized by only one type of repeat motif. Both species are heteroplasmic with respect to the number of repetitive motifs, but unlike the white rhinoceros, the distribution of repetitive motifs in the horse is bimodal with skewed distribution toward high number of motifs (see Fig. 1 in Xu and Arnason, 1994).

The sequence of the repeat motif in the control region of the white rhinoceros is identical to one and similar to the other (5′-CGCACACACA-3′) repeat motif in the control region of the black rhinoceros (Jama *et al.*, 1993).

COMPARISON WITH THE mtDNA OF INDIAN RHINOCEROS

Alignment of the complete mtDNA sequences for the white and the Indian rhinoceroses outside the control regions showed 17 indel differences between the two sequences. One indel (a codon triplet) was in the ATPase8 gene, which is 207 nt long in the white rhinoceros as compared to 204 nt in the Indian rhinoceros, the horse, and the donkey. This indel difference is at the C terminus of the gene. There were six indels in the 12S rRNA gene and eight in the 16S rRNA gene. One indel was observed in each of the tRNA-Leu(UUR)

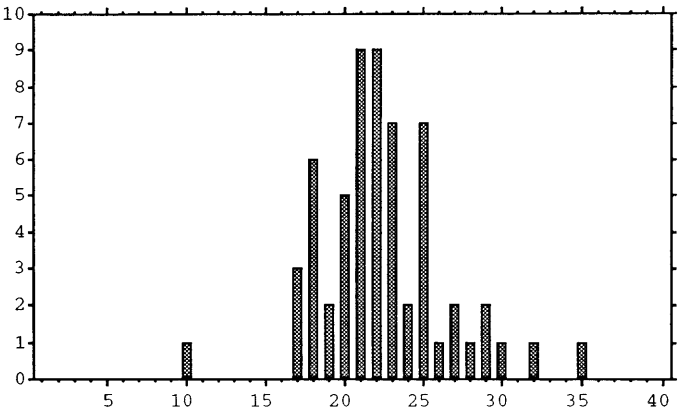

FIG. 1. Distribution of clones, according to their repeat number, in the control region of the mtDNA of the white rhinoceros. The number of repeats per clone was determined by sequencing a total of 60 clones (two directly cloned restriction fragments and 58 PCR clones). The number of repeats per clone is shown on the horizontal axis. The number of clones with a particular repeat number is shown on the vertical axis. The PCR amplification of the control region was undertaken by means of primers that were complementary to the sequence of directly cloned restriction fragments.

TABLE 1

Nucleotide Differences with Respect to Codon Position (1, 2, 3) between Each of the 13 Mitochondrial Peptide-Coding Genes of the White and the Indian Rhinoceroses

			1			2		3				
	Length		Ti		Tv	Ti	Tv	Ti		Tv		
Gene	White	Indian	a	b				a	b	a	b	Indel
NADH1	957	957	10	7	2	3	—	8	50	5	10	
NADH2	1,044	1,044	5	13	3	10	1	6	46	2	15	
COI	1,545	1,545	16	4	1	1	—	16	138	6	25	
COII	684	684	9	4	—	—	—	7	46	—	6	
ATPase8	207	204	—	4	—	9	—	—	13	1	4	1
ATPase6	681	681	10	8	3	6	—	10	52	3	8	
COIII	783	783	6	3	—	5	—	8	59	5	10	
NADH3	345	345	3	3	3	1	1	5	17	1	4	
NADH4L	297	297	5	2	—	2	1	3	20	—	1	
NADH4	1,377	1,377	21	14	6	8	1	16	95	6	9	
NADH5	1,821	1,821	12	25	7	21	7	19	115	3	22	
NADH6	528	528	1	15	1	8	2	9	24	—	12	
Cyt *b*	1,140	1,140	8	6	5	7	1	11	73	4	22	
Total	11,409	11,406	106	108	31	81	14	118	748	36	148	1
Cons. Diff.				139		95				184		
Total Diff.				245		95			1050			1
R. Total Diff.				2.6		1.0			11.1			
R. Cons. Diff.				1.5		1.0			1.9			

Note. nt, nucleotide; Ti, transitions; Tv, transversions; R., ratio; Cons., conservative; a, differences involving leucine in both species; b, differences other than those involving leucine in both species. The indel difference in ATPase8 was not included in the codon position difference.

and the tRNA-Arg gene alignments. These indels were located in the D-loop region of the inferred secondary structure of the tRNA genes. Apart from indel differences, the two molecules (excluding the control regions) differed at 1,666 nt positions (10.8%).

The 13 protein-coding genes of the white and the Indian rhinoceroses were compared with respect to both total nt difference and conservative nt difference (Irwin *et al.*, 1991), Table 1. The nt differences were examined with regards to both the codon position and the type of substitution (transition, transversion). The genes differ at 1,391 positions (12.2%). The codon position ratio for total nt difference is 2.6:1:11.1. The total number of conservative nt differences is 418 (3.7%). The codon position ratio for these differences is 1.5:1:1.9.

Table 2 shows total nt, conservative nt, and aa differences for each of the 13 protein-coding genes of the white and the Indian rhinoceroses. The genes have been arranged according to increasing aa difference. This order and that based on conservative nt difference are reasonably consistent, whereas the order based on total nt difference is markedly different. The relationship among different modes of comparisons has been discussed previously in comparisons of several closely related species pairs (Xu and Arnason, 1996a) and will therefore not be detailed here. The range for total nt difference is between 9.7% (NADH2) and 15.5% (ATPase8), that for conservative nt difference is between 1.5% (COII) and 8.7% (ATPase8), and that for aa difference is between 1.2% (COI) and 18.8% (ATPase8).

TABLE 2

Differences in Number and Percentage between Mitochondrial Protein-Coding Genes of the White and the Indian Rhinoceroses

Gene	Total nt	Cons. nt	Amino acid
COI	207 (13.4)	37 (2.4)	6 (1.2)
COII	72 (10.5)	10 (1.5)	4 (1.8)
COIII	96 (12.3)	23 (2.9)	8 (3.1)
NADH1	95 (9.9)	27 (2.8)	12 (3.8)
NADH4L	34 (11.5)	6 (2.0)	4 (4.1)
Cyt b	137 (12.0)	45 (4.0)	18 (4.7)
NADH4	176 (12.8)	44 (3.2)	30 (6.5)
NADH2	101 (9.7)	44 (4.2)	24 (6.9)
NADH3	38 (11.0)	13 (3.8)	8 (7.0)
ATPase6	100 (14.7)	28 (4.1)	16 (7.1)
NADH5	231 (12.7)	85 (4.7)	57 (9.4)
NADH6	72 (13.6)	38 (7.2)	24 (13.6)
ATPase8	32 (15.5)	18 (8.7)	13 (18.8)
Sum	1391 (12.2)	418 (3.7)	224 (5.9)

Note. Cons. nt, conservative nucleotide difference.

The amplitude between the lowest and highest values for total nt difference is limited compared to that for conservative nt and aa difference, suggesting a high degree of saturation in the dataset for total nt difference.

Table 3 gives details of four pairwise perissodactyl comparisons, horse/donkey, white/Indian rhinoceroses, horse/white rhinoceros, and horse/Indian rhinoceros. The horse and the donkey form the most closely related species-pair, whereas the comparison between the horse and the rhinoceroses represents the deepest perissodactyl divergence. We have previously (Xu *et al.*, 1996a) dated the divergence between the horse and the donkey to ≈9 MYBP, and that between the families Equidae and Rhinocerotidae to ≈50 MYBP (Xu *et al.*, 1996b). With respect to total nt difference a saturation effect is pronounced in the comparison between the horse and the rhinoceroses, leading to a difference that is only 2.3 times greater than that between the horse and the donkey. The differences among the 12S rRNA and the 16S rRNA genes, respectively, in the three divergences (counting the relationship between the horse and the two rhinoceroses as one divergence) differs somewhat from that based on aa and conservative nt differences. While this discrepancy may to some extent be due to stochastic fluctuations, the effect of multiple hits (both transitions and transversions) in fast evolving rRNA sites cannot be ruled out.

The differences in codon position ratios (both total and conservative nt substitution) in the three divergences are consistent with differences in rates and mode of substitution at the three codon positions.

The combined sequences of the tRNA genes of the two rhinoceroses differ by 5.8%. The greatest differences (eight transitions and one transversion, respectively) were registered between the tRNA-Glu and tRNA-Thr genes.

The sequences of two complete tRNA genes, tRNA-Phe and tRNA-Pro, from the black rhinoceros were reported by Jama *et al.* (1993). The tRNA-Phe genes of the white and the black rhinoceroses differ by two transitions, as compared to seven transitions and one transversion between the white and the Indian rhinoceroses. The tRNA-Pro genes of the white and the black rhinoceroses differ by three transitions, one less than that between the white and the Indian rhinoceroses.

The stem region of the origin of replication of the L-strand is identical in the white and the Indian rhinoceroses, while the loop regions differ by three transitions.

ESTIMATING THE DIVERGENCE TIME BETWEEN THE WHITE AND THE INDIAN RHINOCEROSES

The time of the evolutionary divergence between the families Equidae and Rhinocerotidae was recently estimated using the 60 MYBP dating of the split between Artiodactyla and Cetacea (Arnason and Gullberg, 1996) as an external reference. The application of this reference (A/C-60) to the molecular distance between the Equidae and the Rhinocerotidae gave an estimated divergence time of ≈50 MYBP between these two perissodactyl families (Xu *et al.*, 1996b). The aa difference between the combined sequences of the 13 protein-coding genes of the Equidae and the Rhinocerotidae is 9.8%, while the corresponding difference between the genera *Ceratotherium* and *Rhinoceros* is 5.9%. Application of A/C-60 dates the divergence between the two rhinoceroses to 27.2 MYPB (95% confidence limits: ±3.8 MY). This dating conforms with the molecular difference between the Equidae and the Rhinocerotidae, suggesting that the relative rate of evolution has been similar in the two families.

TABLE 3

Comparison of mtDNA in Four Perissodactyl Species-Pairs: Horse/Donkey, White/Indian Rhinoceroses, Horse/White Rhinoceros, and Horse/Indian Rhinoceros

	Horse/donkey	White/Indian rhinos	Horse/white rhino	Horse/Indian rhino
Percentage nt difference	6.9	10.9	15.6	15.6
Percentage conservative nt difference	1.2	3.7	8.1	8.4
Percentage aa difference	1.9	5.9	9.5	9.8
Codon position ratio, total nt substitution	3.5:1.0:25.7	2.6:1.0:11.1	2.4:1.0:9.5	2.3:1.0:8.9
Codon position ratio, conservative nt substitution	1.5:1.0:2.1	1.5:1.0:1.9	1.6:1.0:3.5	1.6:1.0:3.6
Percentage 12S rRNA difference, Total;Ti;Tv;indel	4.9;3.9;0.7;0.3	7.6;6.0;1.0;0.6	12.3;9.0;2.4;0.9	13.1;9.4;2.8;0.9
Percentage 16S rRNA difference, Total;Ti;Tv;indel	3.7;2.9;0.6;0.2	8.1;6.8;0.8;0.5	11.5;6.7;3.6;1.2	11.3;6.4;3.6;1.3
Percentage tRNA difference, Total;Ti;Tv;indel	3.5;2.7;0.3;0.5	5.8;5.0;0.7;0.1	11.2;8.5;2.0;0.7	11.1;8.5;1.8;0.8
Ti/Tv ratio in codon position 1;2;3	25.5;14.0;11.2	6.9;5.8;4.7	2.7;2.6;1.7	2.9;3.6;1.5
Ti/Tv ratio in 12S rRNA; 16S rRNA; tRNA	5.4;5.1;8.2	5.8;8.2;7.6	3.8;1.8;4.3	3.4;1.8;4.6

Note. Ti, transition; Tv, transversion; Percentage nt difference does not include control region. Percentage conservative nt difference is limited to protein-coding genes. Data of protein-coding and tRNA genes are based on the concatenated sequences of 13 protein-coding genes and 22 tRNA genes, respectively.

COMPARISON OF SEQUENCES FROM THE WHITE, THE INDIAN, AND THE BLACK RHINOCEROSES

The sequences of the mitochondrial cytochrome *b* gene (Irwin *et al.*, 1991) and of two tRNA genes (tRNA-Phe and tRNA-Pro) and the control region (Jama *et al.*, 1993) have been reported for the black rhinoceros. In both the control region and the two tRNA genes there are fewer differences between the white and the black rhinoceroses than between either of these species and the Indian rhinoceros. Comparisons of repeat motifs from the control region of the three species are also consistent with a closer relationship between the white and the black rhinoceroses than between either of these species and Indian rhinoceros. With respect to the cytochrome *b* gene, however, the phylogenetic relationship among the three species varies depending on the mode of comparison. The pairwise difference between the cytochrome *b* gene of the three species is shown in Table 4. Protpars analysis identified the white and the Indian rhinoceroses as sister taxa to the exclusion of the black rhinoceros (bootstrap value 81), whereas analysis of the nt sequences joined the white and the black rhinoceroses with a bootstrap value of 86.

DISCUSSION

The family Rhinocerotidae includes four extant genera—two African (*Ceratotherium* and *Diceros*) and two Asian (*Rhinoceros* and *Dicerorhinus*). Systematic relationships among the four genera are controversial, with arguments, to a considerable extent, focussing on the phylogenetic position of the Sumatran rhinoceros, *Dicerorhinus sumatrensis* (an Asian two-horned rhinoceros). Hypotheses for the relationships of these genera have been proposed on the basis of both the number of horns and geographical distribution. According to the number of horns hypothesis (Simpson, 1945) the African genera plus the Sumatran rhinoceros have been joined in a two-horned group. The "geographic split" hypothesis (Groves, 1983), which does not consider the number of horns, joins the two Asian genera into one group, separate from the African rhinoceroses. A third hypothesis (Prothero and Schoch, 1989), which takes into account both number of horns and geographic distribution, separates all extant rhinoceroses into three subtribes—one including the two African genera (*Ceratotherium* and *Diceros*), one including *Dicerorhinus*, and one including *Rhinoceros*. Restriction site mapping of the mitochondrial rRNA genes (Morales and Melnick, 1994) supports a subdivision of the living rhinoceroses according to the number of horns, consistent with Simpson (1945).

In the present study we have studied the basal divergence among extant rhinoceroses, irrespective of the mode of defining this distinction (number of horns or geographical distribution), by examining the molecular difference between the white rhinoceros, a two-horned rhinoceros which lives in Africa, and the Indian rhinoceros, a one-horned rhinoceros which lives in Asia. On the basis of this molecular comparison we estimate the basal rhinocerotid divergence to have been ≈27 MYBP. The external reference used in establishing this divergence time was the A/C-60 standard, the dating of 60 MYBP for the evolutionary divergence between Artiodactyla and Cetacea (Arnason and Gullberg, 1996). As discussed by Xu *et al.* (1996b) and Arnason *et al.* (1996) the divergence between Artiodactyla and Cetacea is probably the most distinct ordinal mammalian divergence that has been defined in both paleontological (Gingerich *et al.*, 1994; Thewissen *et al.*, 1994) and molecular (Arnason and Gullberg, 1996) terms, and the application of A/C-60 to the dating of the divergence between Equidae and Rhinocerotidae gives an estimated divergence time (Xu *et al.*, 1996b) which conforms reasonably well with paleontological (Prothero and Schoch, 1989) findings.

Our dating of the basal divergence among extant rhinoceroses is somewhat earlier than some previously proposed datings of this divergence. Based on studies of variation at allozymic loci, Merenlender *et al.* (1989) calculated that the genetic distance between the white (*Ceratotherium simum*) and the black (*Diceros bicornis*) rhinoceroses was around one-third of that between either of these species and the Indian rhinoceros (*Rhinoceros unicornis*). The fossil records of *Diceros* and *Ceratotherium* indicate that they have coexisted for approximately seven million years and based on these findings Merenlender *et al.* (1989) dated the divergence between these genera and *Rhinoceros* to 26 MYBP. This dating is consistent with our dating of the divergence between the white and Indian rhinoceroses, obtained using A/C-60. It should be noted, however, that the dating by Merenlender *et al.* (1989) is dependent on a much closer relationship between *Ceratotherium* and *Diceros* than is suggested by comparison of the complete cytochrome *b* genes of the black, white, and Indian rhinoceroses (see Table 4).

Based on restriction site mapping, O'Ryan and Har-

TABLE 4

Percentage aa (above Diagonal) and Conservative nt (below Diagonal) Differences among the Cytochrome *b* Genes of Five Perissodactyl Species

Species	1	2	3	4	5
1 Horse	—	0.8	8.4	7.4	8.2
2 Donkey	1.1	—	8.2	7.1	7.9
3 White rhinoceros	8.6	8.6	—	4.7	5.8
4 Indian rhinoceros	7.6	7.5	4.0	—	6.3
5 Black rhinoceros	8.6	8.8	3.6	5.0	—

ley (1993) reported a 6.8% sequence divergence between the mtDNAs of the black and the white rhinoceroses. This figure is low compared with the total difference (9.9%, see Table 4) between the cytochrome *b* genes of these species. Without calibration for differences in evolutionary rates it is difficult to use total difference for calculating times of divergence and it has furthermore been shown (Ohland *et al.*, 1995) that differences calculated on the basis of restriction analysis may differ markedly from actual differences. On the basis of restriction analysis, Morales and Melnick (1994) proposed an evolutionary rate of 0.3% per million years in the mitochondrial rRNA genes of rhinoceroses. Applying this rate to partial sequences of the 12S and 16S rRNA genes (≈1.6 kilobases total length), and using as a reference the divergence between Equidae and Rhinocerotidae dated at 50 MYBP, they calculated that extant two-horned (such as the white rhinoceros) and one-horned (such as the Indian rhinoceros) rhinoceroses separated around 21.5 MYBP. This dating is somewhat later than our estimate of 27 MYBP, obtained using complete mitochondrial DNA molecules.

ACKNOWLEDGMENTS

We are grateful to Peter Arctander for tissue samples, to Axel Janke for valuable comments, and to Kerryn Slack for help with language. The work was supported by the Swedish Natural Sciences Research Council, the Erik Philip-Sörensen Foundation, and EU Contract ERBCHR XCT 930254.

REFERENCES

Arnason, U., Gullberg, A., and Widegren, B. (1991). The complete nucleotide sequence of the mitochondrial DNA of the fin whale, *Balaenoptera physalus*. *J. Mol. Evol.* **33:** 556–568.

Arnason, U., and Gullberg, A. (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.

Arnason, U., and Gullberg, A. (1996). Cytochrome *b* nucleotide sequences and the identification of five primary lineages of extant cetaceans. *Mol. Biol. Evol.* **13:** 407–417.

Arnason, U., Gullberg, A., Janke, A., and Xu, X. (1996). Pattern and timing of evolutionary divergences among hominoids based on analyses of complete mtDNAs. *J. Mol. Evol.*, **43:** 650–661.

Attardi, G., Chomyn, A., Doolittle, R. F., Mariottini, P., and Ragan, C. I. (1986). Seven unidentified reading frames of human mitochondrial DNA encode subunits of the respiratory chain NADH dehydrogenase. *Cold Spring Harbor Symp. Quant. Biol.* **LI:** 103–114.

Bibb, M. J., Van Etten, R. A., Wright, C. T., Walberg, M. W., and Clayton, D. A. (1981). Sequence and gene organization of mouse mitochondrial DNA. *Cell* **26:** 167–180.

Gadaleta, G., Pepe, G., De Candia, G., Quagliariello, C., Sibisa, E., and Saccone, C. (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.

GCG (1994). Program Manual for the Wisconsin Package, Version 8, Genetics Computer Group, 575 Science Drive, Madison, Wisconsin, USA 53711.

Ghivizzani, S. C., Mackay, S. L. D., Madsen, C. S., Laipis, P. J., Hauswirth, W. W. (1993). Transcribed heteroplasmic repeated sequence in the porcine mitochondrial DNA D-loop region. *J. Mol. Evol.* **37:** 36–47.

Gingerich, P. D., Raza, S. M., Arif, M., Zhou, X. Y. (1994). New whale from the Eocene of Pakistan and the origin of cetacean swimming. *Nature* **368:** 844–847.

Groves C. P. (1983). Phylogeny of the living species of rhinoceros. *Z. Zool. Syst. Evolutionsforsch.* **21:** 293–313.

Irwin, D. M., Kocher, T. D., and Wilson, A. C. (1991). Evolution of the cytochrome *b* gene of mammals. *J. Mol. Evol.* **32:** 128–144.

Jama, M., Zhang, Y. P., Aman, R. A., and Ryder, O. A. (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.

Krettek, A., Gullberg, A., Arnason, U. (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.

Merenlender, A. M., Woodruff, D. S., Ryder, O. A., Kock, R., and Vahala, J. (1989). Allozyme variation and differentiation in African and Indian rhinoceros. *J. Hered.* **80:** 377–382.

Morales, J. C., and Melnick, D. J. (1994). Molecular Systematics of the living rhinoceros. *Mol. Phyl. Evol.* **3:** 128–134.

Ohland, D. P., Harley, E. H., and Best, P. B. (1995). Systematics of cetaceans using restriction site mapping of mitochondrial DNA. *Mol. Phylogenet. Evol.* **4:** 10–19.

Ojala, D., Montoya, J., and Attardi, G. (1981). tRNA punctuation model of RNA processing in human mitochondria. *Nature* **290:** 470–474.

O'Ryan, C., and Harley, E. H. (1993). Comparisons of mitochondrial DNA in black and white rhinoceroses. *J. Mamm.* **74:** 343–346.

Prothero, D. R., and Schoch, R. M. (1989). Classification of the Perissodactyla. *In* "The Evolution of Perissodactyla" (D. R. Prothero and R. M. Schoch, Eds.), pp. 530–537, Oxford Univ. Press, New York.

Sanger, F., Nicklen, S., and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74:** 5463–5467.

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., Hussain, S. T., and Arif, M. (1994). Fossil evidence for the origin of aquatic locomotion in archaeocete whales. *Science* **263:** 210–212.

Wolstenholme, D. R. (1992). Animal mitochondrial DNA: structure and evolution. *Int. Rev. Cytol.* **141:** 173–216.

Xu, X., Gullberg, A., and Arnason, U. (1996a). 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., Janke, A., and Arnason, U. (1996b). 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.