J Mol Evol (1998) 47:302–306

JOURNAL OF MOLECULAR EVOLUTION

The Complete Mitochondrial DNA Sequence of the Pig (*Sus scrofa*)

Björn M. Ursing, Ulfur Arnason

Department of Genetics, Division of Evolutionary Molecular Systematics, University of Lund, S-223 62 Lund, Sweden

Received: 2 December 1997 / Accepted: 20 February 1998

Abstract. The complete mitochondrial genome sequence of the pig, *Sus scrofa*, was determined. The length of the sequence presented is 16,679 nucleotides. This figure is not absolute, however, due to pronounced heteroplasmy caused by variable numbers of the motif GTACACGTGC in the control region of different molecules. A phylogenetic study was performed on the concatenated amino acid and nucleotide sequences of 12 protein-coding genes of the mitochondrial genome. The analysis identified the pig (Suiformes) as a sister group of a cow/whale clade, making Artiodactyla paraphyletic. The split between pig and cow/whale was molecularly dated at 65 million years before present.

Key words: Mitochondrial DNA — *Sus scrofa* — Suiformes — Artiodactyla — Ruminantia — Cetacea — Ferungulata — Phylogeny

Introduction

Artodactyla is the most speciose order of large mammals. The fossil record of the order is extensive, making Artiodactyla an attractive choice for testing and comparing molecularly and morphologically based phylogenetic hypotheses. The postulated divergence time for the deepest divergences among extant artiodactyls is 55–65 million years before present (MYBP), (Miyamoto et al. 1993). The minimum divergence times for the same divergences, as established by paleontological findings, are ≈55 MY. Thus the difference between the minimum and the postulated divergence times in Artiodactyla is less than in most other deep mammalian divergences. This circumstance makes the order ideal for establishing molecular references that can be applied for dating evolutionary divergences within orders with a poorer fossil record.

Traditionally the order Artiodactyla includes three recent suborders: Suiformes (pigs and hippopotamuses), Tylopoda (camels and lamas), and Ruminantia (cows, deers, antelopes, tragulids, and giraffes). The suborder Ruminantia is commonly recognized as the most recently originating lineage, but the relationship between the Ruminantia and the other two suborders has not been conclusively settled.

Until now the artiodactyls have been represented by only one complete mitochondrial DNA (mtDNA) molecule, that of the cow, *Bos taurus* (Anderson et al. 1982). To explore artiodactyl and artiodactyl/cetacean relationships further, we present the complete mtDNA sequence of the pig, *Sus scrofa.*

Materials and Methods

An enriched mtDNA fraction was isolated from the liver of a male pig delivered by Scan Slaughterhouse, Kävlinge, Sweden. The prepration procedure was the same as described by Arnason et al. (1991). The mtDNA was digested separately with *Spe*I, *Bln*I, *Bcl*III, *Hin*dIII, and *Bgl*II and the product ligated directly into phage M13mp18. A *Spe*I-digested sample was also run on a preparative agarose gel; the bands were excised, electroeluted, and ligated into phage M13mp18. Regions not covered by natural clones were PCR amplified before ligation into M13mp18 and/or M13mp19. Sequencing was performed manually on single-stranded DNA, according to the dideoxy method with ^{35}S-dATP (Sanger 1981), using both universal and numerous specific oligodeoxynucleotide primers. Sequences of PCR-amplified regions constitute a consensus of at least three clones.

The sequence of the mitochondrial DNA of the pig has been deposited at the EMBL database under accession number AJ002189. Users of the sequence are kindly requested to refer to the present publication, and not just to the accession number of the sequence.

Correspondence to: B.M. Ursing; *e-mail:* bjorn.ursing@gen.lu.se

Results

Description of the mtDNA of the Pig

The length of the presently reported mtDNA of the pig is 16,679 nucleotides (nt). The length of the molecule is not absolute, however, due to heteroplasmy caused by differences in the numbers of repeated motifs in the control region. The 16,679-nt variety of the molecule has 20 such motifs. The location of various features of the molecule is given in Table 1. Position 1 of the sequence has been assigned to the 5′ nt of the tRNA-Phe gene.

With the exception of NADH2 and NADH4L, all mitochondrial protein-coding genes of the pig have a methionine start codon, ATG or ATA. The start codon of NADH2 is ATT (isoleucine) and that of NADH4L is GTG (valine). An ATT start codon in the NADH2 gene has been reported previously in primates (Horai et al. 1995; Arnason et al. 1996b). Among mammals a GTG start codon has been reported in the NADH4L gene of blue whale (Arnason et al. 1993) and in the NADH6 gene of Indian rhinoceros (Xu et al. 1996b). The GTG has also been described as a start codon in several metazoan mtDNAs (Okimoto et al. 1990; Wolstenholme 1992).

Four protein-coding genes, NADH3, NADH4, cytochrome *c* oxidase subunit II (COII), and COIII, are terminated with an incomplete stop codon, TA or T. As discussed by Wolstenholme (1992) some mammalian protein-coding genes are terminated with TA or T, rather than with a complete stop codon. In these cases the terminal nt of the stop codon is adjacent to the 5′ terminal nt of a tRNA gene. This is consistent with the notion that the transcripts of such genes are completed by the post-transcriptional polyadenylation (Ojala et al. 1981).

The control region of the mtDNA of the pig contains the repeat motif GTACACGTGC. The 5′ nt of the first repeat is at position 16,146. The motif has the purine/pyrimidine alternation characterizing many repetitive motifs occurring in the control region of mammalian mtDNAs. In an extensive study of the control region of the pig (Ghivizzani et al. 1993), the range of repeat numbers was 14–29. In the present study four PCR clones were sequenced. The repeat number in these clones was 20, 23, 25, and 25, respectively, i.e., within the previously described range. No variation occurred in the four clones with respect to the sequence of the repeated motif.

The conserved sequence blocks, CSB-1, CSB-2, and CSB-3, are located in positions 16,109–16,134, 16,378–16,396, and 16,430–16,448, respectively. The sequences of the CSBs are identical to those described by Ghivizzani et al. (1993). The L-strand origin of replication (L-ori) of the mtDNA of the pig is located at position 5166 to 5212. The loop of the probable secondary structure of the L-ori of the pig is 13 nt long. The stem is 12 base pairs long, with full complementarity between the two strands. The secondary structure of the L-ori of the pig conforms with that of other eutherians except the armadillo (Arnason et al. 1997).

Table 1. Location of features in the mitochondrial DNA of the pig, *Sus scrofa*[a]

	Position			Codon	
Feature	From	To		Start	Stop
tRNA-Phe	1	70			
12S rRNA	71	1,032			
tRNA-Val	1,032	1,099			
16S rRNA	1,100	2,668			
tRNA-Leu (UUR)	2,668	2,742			
NADH1	2,745	3,701		ATG	TAG
tRNA-Ile	3,700	3,768			
tRNA-Gln	3,766	3,838	(L)		
tRNA-Met	3,840	3,909			
NADH2	3,910	4,953		ATT	TAG
tRNA-Trp	4,952	5,019			
tRNA-Ala	5,093	5,026	(L)		
tRNA-Asn	5,169	5,095	(L)		
Or. L-stand repl.	5,166	5,212			
tRNA-Cys	5,267	5,202	(L)		
tRNA-Tyr	5,332	5,267	(L)		
COI	5,334	6,878		ATG	TAA
tRNA-Ser (UCN)	6,950	6,882	(L)		
tRNA-Asp	6,958	7,025			
COII	7,026	7,713		ATG	T--
tRNA-Lys	7,714	7,780			
ATPase8	7,782	7,985		ATG	TAA
ATPase6	7,943	8,623		ATG	TAA
COIII	8,623	9,407		ATG	TA-
tRNA-Gly	9,407	9,475			
NADH3	9,476	9,822		ATA	TA-
tRNA-Arg	9,823	9,891			
NADH4L	9,892	10,188		GTG	TAA
NADH4	10,182	11,559		ATG	T--
tRNA-His	11,560	11,628			
tRNA-Ser (AGY)	11,629	11,687			
tRNA-Leu (CUN)	11,688	11,757			
NADH5	11,758	13,577		ATA	TAA
NADH6	14,088	13,559	(L)	ATG	TAA
tRNA-Glu	14,157	14,089	(L)		
Cyt *b*	14,162	15,301		ATG	AGA
tRNA-Thr	15,302	15,369			
tRNA-Pro	15,433	15,369	(L)		
Control region	15,434	16,679			
CSB-1	16,109	16,134			
CSB-2	16,378	16,396			
CSB-3	16,430	16,448			

[a] (L), light-strand sense; NADH1–6 and NADH4L, subunits 1–6 and 4L of nicotinamid dinucleotide dehydrogenase; ATPase6 and 8, subunits 6 and 8 of adenosine triphosphatase; COI–COIII, cytochrome *c* oxidase subunits I–III; cyt *b*, cytochrome *b*.

The cytochrome *b* (cyt *b*) sequence of the pig was described by Irwin et al. (1991). That and the present sequences differ by one first codon position nonsynonymous transversion and one synonymous transition in the third codon position.

Phylogeny

Phylogenetic analyses were carried out on the concatenated sequences of all H stand-encoded protein-coding

genes of the mitochondrial genome. The alignment included 17 species. After excluding gaps and ambiguous sites adjacent to gaps, the length of the alignment was 10,365 nt, or 3455 amino acids (aa).

The nt analyses included all changes at the first codon position except synonymous leucine transitions, all changes at the second colon position, and third codon position transversions. The tree reconstruction analyses were performed at both the aa and the nt levels. At the nt level, support values were calculated separately for first, second, first plus second, and all codon positions.

Phylogenetic analyses were based on the three most commonly used methods of tree reconstruction, maximum parsimony (MP) (Fitch 1971); neighbor joining (NJ) (Saitou and Nei 1987), and maximum likelihood (ML) (Felsenstein 1991); implemented by the PHYLIP (Felsenstein 1991), the MOLPHY (Adachi and Hasegawa 1996b), or the PUZZLE, Version 3.1 (Strimmer and von Haeseler 1996), program packages. The support values for ML were calculated using quartet puzzling (QP).

Support values for NJ and MP were established by bootstrapping. At the nt level the HKY model for sequence evolution (Hasegawa et al. 1985) was applied with 1000 replicates; at the aa level the analyses were based on the Dayhoff (1978) matrix. Due to computational constraints, the aa analyses were limited to 100 replicates. The ML/QP analyses were performed with 1000 puzzling steps and the TN (Tamura and Nei 1993) model for nt sequence evolution. Nucleotide frequencies and transition/transversion ratios were estimated from the data set. The aa analyses were according to the mtREV-24 model (Adachi and Hasegawa 1996a) and aa frequencies established from the data set.

Figure 1 shows a tree based on phylogenetic analysis of the concatenated aa sequences of 12 mitochondrial protein-coding genes. The tree was rooted with the wallaroo, *Macropus robustus* (Janke et al. 1997). With the exception of the hedgehog, none of the sequences deviated (5%-level χ^2 test) in aa or nt composition from the frequency distribution underlying the ML model. The inclusion or exclusion of the hedgehog had no effect on the topology of the tree.

The phylogenetic analyses reconstructed the commonly recognized eutherian relationship with the Lipotyphla, as represented by the hedgehog, basal to myomorph rodents (Krettek et al. 1995). Consistent with recent analyses of complete mtDNAs, the armadillo was identified as the sister group of the ferungulate (Artiodactyla, Perissodactyla, Carnivora, and Cetacea) clade (Arnason et al. 1997), and as recognized by Xu et al. (1996b) the primary split within the ferungulates is between an Artiodactyla/Cetacea clade and a Perissodactyla/Carnivora clade. Consistent with analyses of complete cytochrome *b* sequences (Irwin and Arnason 1994; Graur and Higgins 1994; Arnason and Gullberg 1996) and nuclear sequences (Gatesy et al. 1996, Gatsey 1997), the position of the pig was outside the Ruminantia (cow)/Cetacea (fin whale, blue whale) clade.

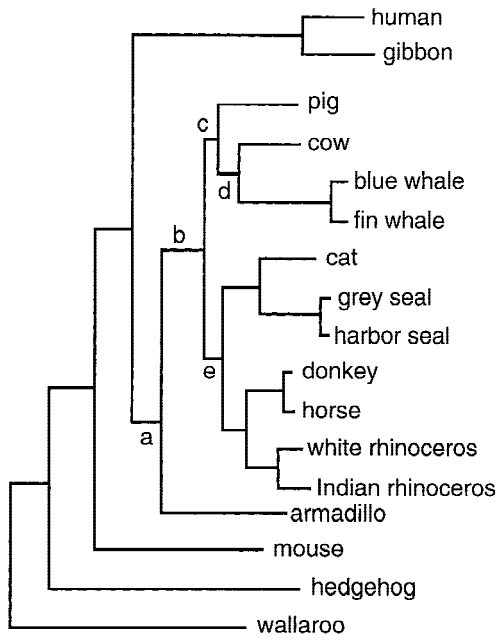

Fig. 1. Maximum-likelihood tree as reconstructed by the PUZZLE, Version 3.1, program showing the position of the pig relative to that of a number of other taxa represented by complete mitochondrial molecules. The tree is based on concatenated aa sequences of 12 mitochondrial protein coding genes. The branch lengths are proportional to the genetic distances among the taxa. The support for the branches labeled *a–e* is given in Table 2. The analysis included the following species: pig, *Sus scrofa* (present study); cow, *Bos taurus* (Anderson et al. 1982); blue whale, *Balaenoptera musculus* (Arnason and Gullberg 1993); fin whale, *Balaenoptera physalus* (Arnason et al. 1991); grey seal, *Halichoerus grypus* (Arnason et al. 1993); harbor seal, *Phoca vitulina* (Arnason and Johnsson 1992); cat, *Felis catus* (Lopez et al. 1996); horse, *Equus callabus* (Xu and Arnason 1994); donkey, *Equus asinus* (Xu et al. 1996a); Indian rhinoceros, *Rhinoceros unicornis* (Xu et al. 1996b); white rhinoceros, *Ceratotherium simum* (Xu and Arnason 1997); armadillo, *Dasypus novemcinctus* (Arnason et al. 1997); human ("Lund"), *Homo sapiens* (Arnason et al. 1996c); white-handed gibbon, *Hylobates lar* (Arnason et al. 1996b); mouse, *Mus musculus* (Bibb et al. 1981); hedgehog, *Erinaceus europaeus* (Krettek et al. 1995); and wallaroo, *Macropus robustus* (Janke et al. 1997).

The support for the topology shown in Fig. 1 in different data sets and according to different analytical methods is given in Table 2. There is general support for the depicted topology in most data sets, even though the results of some of the analyses of the more limited data sets deviate from the general pattern. These deviations are most pronounced in the MP analyses. The previously recognized armadillo/ferungulate clade (Arnason et al. 1997) was identified in all data sets and approaches; the ferungulate clade received almost-absolute support in all analyses. The Ruminantia/Cetacea clade was separated from the pig, with aa support and bootstrap values in the range 80–98.

The timing of the split between the pig and the cow/whale lineage was calculated according to the A/C-60

Table 2. Support Values for ML/QP and Bootstrap Values for MP and NJ

		Branch[a]				
Method	Position	a	b	c	d	e
ML/QP	aa	99	100	93	98	94
	1	98	99	89	76	—[b]
	2	98	99	85	75	—
	12	99	99	95	58	76
	123	100	99	96	58	76
MP	aa	87	100	76	80	62
	1	50	100	—	—	66
	2	70	92	—	83	—
	12	83	100	70	82	—
	123	64	100	60	95	63
NJ	aa	100	100	85	93	99
	1	84	100	76	62	93
	2	97	100	73	76	41
	12	99	100	92	82	95
	123	83	100	99	100	99

[a] Branches a–e are shown in Fig. 1.
[b] Not identified as a separate branch in these analyses.

reference (Arnason and Gullberg 1996; Arnason et al. 1996a). A/C-60 is based on the divergence between ruminant Artiodactyla (cow) and Cetacea (fin and blue whales), set at 60 MYBP. The dating was estimated on the basis of ML distances of aa and second codon positions. These data sets are the most conservative and are most likely to give robust results. Using the armadillo and the carnivores and perissodactyls included in Fig. 1 as the outgroup, the split between the pig and the cow/whale lineage was estimated as 65 MYBP. The evolutionary rate of branch d in Fig. 1 was taken as the mean of the branches leading to the cow and the whales.

Discussion

The present analyses have demonstrated artiodactyl paraphyly by joining a ruminant representative, the cow, and the two whales in a common clade to the exclusion of the pig. The support values for this relationship in the aa data set were 98 (QP), 93 (NJ), and 80 (MP), respectively. Artiodactyl monophyly was challenged in less comprehensive data sets by Irwin and Arnason (1994) and Graur and Higgins (1994), who showed that cetacean origin was within the Artiodactyla. The cytochrome *b* data set utilized by Arnason and Irwin (1994) included a hippopotamid sequence and the analysis suggested a sister-group relationship between Hippopotamidae and Cetacea. Cetacean/hippopotamid affinities have been supported in recent studies of nuclear data (Gatesy et al. 1996; Gatesy 1997). Molecular data tend not to support a suid/hippopotamid relationship (Beintema et al. 1986; Gelusniak et al. 1990; Stanhope et al. 1993), even though the hippopotamid/cetacean relationship was not specifically identified.

Cetacean origin and evolution were contentious issues for a long time, with most authorities advocating cetacean diphyly. Monophyly of extant cetaceans and a sister-group relationship between Artiodactyla and Cetacea was proposed on molecular grounds (Boyden and Gemeroy 1950; Goldstone and Smith 1966; Beintema et al. 1977) and on morphological grounds by Van Valen (1966), who argued that the two orders had originated separately from Mesonychidae. The present study is inconsistent with Van Valen's (1966) proposal of separate origins of Artiodactyla and Cetacea from Mesonychidae, since the analysis placed the cetacean origin within the Artiodactyla and thus more recent than early artiodactyl diversification.

The demonstration of cetacean origin within the Artiodactyla raises some taxonomic questions, and as argued by Graur and Higgins (1994), if taxa are required to be monophyletic, the molecular relationships between artiodactyls and cetaceans provide two options: either to degrade the Cetacea to a suborder within the Artiodactyla or to fragment the present Artiodactyla into three independent orders, Tylopoda, Suiformes, and Ruminantia (and probably also a fourth order, Hippopotamidea, according to more recent findings). The arrangement favored by Graur and Higgins (1994) was to include Cetacea as an artiodactyl suborder.

The present study supports the recent findings (Arnason et al. 1997) that the armadillo is a sister group to Cetferungulata. This position of the armadillo is inconsistent with previous molecular studies, which have placed the armadillo (Edentata) basal among the Eutheria (Goodman et al. 1985; Adkins and Honeycutt 1991). If the eutherian tree is rooted with the edenates, the artiodactyls are placed at a basal position.

The dating of the divergence between the pig and the Ruminantia/Cetacea clade, 65 MYBP, is consistent with the earliest postulated times of artiodactyl divergences (Miyamoto et al. 1993), yielding further support for the application of A/C-60 (Arnason and Gullberg 1996) as a molecular reference for dating various mammalian divergences. The present study permitted resolution of 60- to 65-MY-old divergences that took place within 5 MY. The findings underscore the importance of having large data sets for obtaining resolution in this interval.

Acknowledgments. We express our gratitude to Scan Slaughterhouse in Kävlinge, Sweden, for the sample, Suzette Mouchaty for help with language, and Axel Janke for valuable comments. This study was supported by the Swedish National Sciences Research Council, the Jörgen Lindström Foundation, and the Nilsson-Ehle Foundation.

References

Adachi J, Hasegawa M (1996a) Model of amino acid substitution in proteins encoded by mitochondrial DNA. J Mol Evol 42:459–468

Adachi J, Hasegawa M (1996b) MOLPHY version 2.3: programs for molecular phylogenetics based on maximum likelihood. Comput Sci Monogr 28:1–150

Adkins RM, Honeycutt RL (1991) Molecular phylogeny of the superorder Archonta. Proc Natl Acad Sci USA 88:10317–10321

Anderson S, de Brujin MHL, Coulson AR, Eperon IC, Sanger F, Young G (1982) Complete sequence of bovine mitochondrial DNA. Conserved features of the mammalian mitochondrial genome. Nature 37:312–322

Arnason U, 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, 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, Johnsson E (1992) The complete mitochondrial DNA sequence of the harbor seal, *Phoca vitulina.* J Mol Evol 34:493–505

Arnason U, Xu X, Gullberg A (1996c) Comparison beween the complete mitochondrial DNA sequences of *Homo* and the common chimpanzee based on nonchimeric sequences. J Mol Evol 42:145–152

Arnason U, Gullberg A, Janke A (1997) Phylogenetic analyses of mitochondrial DNA suggest a sister group relationship between Xenarthra (Edentata) and ferungulates. Mol Biol Evol 14:762–768

Arnason U, Gullberg A, 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, Gullberg A, Johnsson E, Ledje C (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, Gullberg A, Janke A, Xu X (1996a) Pattern and timing of evolutionary divergences among hominoids based on analyses of complete mtDNAs. J Mol Evol 42:650–661

Arnason U, Gullberg A, Xu X (1996b) 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

Beintema JJ, Gaastra W, Lenstra JA, Welling GW, Fitch WM (1977) The molecular evolution of pancreatic ribonuclease. J Mol Evol 10:49–71

Beintema JJ, Fitch WM, Carsana A (1986) Molecular evolution of pancreatic-type ribonucleases. Mol Biol Evol 3:262–275

Bibb MJ, Van Etten RA, Wright CT, Walberg MW, Clayton DA (1981) Sequence and organization of mouse mitochondrial DNA. Cell 26:167–180

Boyden A, Gemeroy D (1950) The relative position of the Cetacea among the orders of Mammalia as indicated by precipitin tests. Zoologica 35:145–151

Dayhoff MO (1978) Atlas of protein sequence and structure. Vol 5, Suppl 3. Natl Biomed Res Found, Washington, DC

Felsenstein J (1991) PHYLIP: phylogeny inference package, Version 3.5. University of Washington, Seattle

Fitch WM (1971) Towards defining the course of evolution, minimum change from a specific tree topology. Syst Zool 20:406–415

Gatesy J (1997) More DNA support for a Cetacea-Hippopotamidae clade: the blood-clotting protein gene gamma-fibrinogen. Mol Biol Evol 14:537–543

Gatesy J, Hayashi C, Cronin MA, Arctander P (1996) Evidence from milk casein genes that cetaceans are close relatives of hippopotamid artiodactyls. Mol Biol Evol 13:954–963

Ghivizzani SC, Mackay SLD, Madsen CS, Laipis PJ, Hauswirth WW (1993) Transcribed heteroplasmic repeated sequences in the porcine mitochondrial DNA d-loop region. J Mol Evol 37:36–47

Goldstone A, Smith EL (1996) Amino acid sequence of whale heart cytochrome *c*. J Biol Chem 241:4480–4486

Goodman M, Czelusniak J, Beeber JE (1985) Phylogeny of primates and other eutherian orders: a cladistic analysis using amino acid and nucleotide sequence data. Cladistics 1(2):171–185

Graur D, Higgins DG (1994) Molecular evidence for the inclusion of cetaceans within the order Artiodactyla. Mol Biol Evol 11:357–364

Hasegawa M, Kishino H, Yano TA (1985) Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J Mol Evol 22:160–174

Horai S, Hayasaka K, Kondo R, Tsugane K, Takahata N (1995) Recent African origin of modern humans revealed by complete sequences of hominoid mitochondrial DNAs. Proc Natl Acad Sci USA 92:532–536

Irwin DM, Arnason U (1994) Cytochrome *b* gene of marine mammals: phylogeny and evolution. J Mammal Evol 2:37–55

Irwin DM, Kocher TD, Wilson AC (1991) Evolution of the cytochrome *b* gene of mammals. J Mol Evol 32:128–144

Janke A, Xu X, Arnason U (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

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

Lopez JV, Cevario S, O'Brien SJ (1996) Complete nucleotide sequences of the domestic cat (*Felis catus*) mitochondrial genome and a transposed mtDNA tandem repeat (Numt) in the nuclear genome. Genomics 33:229–246

Miyamoto MM, Kraus F, Laipis PJ, Tanhauser SM, Webb SD (1993) Mitochondrial DNA phylogenies within Artiodactyla. In: Slalay FS, Novacek MJ, McKenna MC (eds) Mammal phylogeny. Springer-Verlag, New York, pp 268–281

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

Okimoto R, Macfarlane JL, Wolstenholme DR (1990) Evidence for the frequent use of TTG as the translation initiation codon of mitochondrial protein genes in the nematodes *Ascaris suum* and *Caenorhabditis elegans.* Nucleic Acids Res 18:6113–6118

Saitou N, Nei M (1987) The neighbor-joining method a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

Sanger F (1981) Determination of nucleotide sequences in DNA. Science 214:1205–1210

Strimmer K, von Haeseler A (1996) Quartet puzzling: a quartet maximum-likelihood method for reconstructing tree topologies. Mol Biol Evol 13:964–969

Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol Biol Evol 10:512–526

Van Valen L (1966) Detratheridia, a new order of Mammals. Bull Am Mus Nat Hist 132:1–126

Wolstenholme DR (1992) Animal mitochondrial DNA structure and evolution. Int Rev Cytol 141:173–216

Xu X, Arnason U (1994) The complete mitochondrial DNA sequence of the horse, *Equus caballus:* extensive heteroplasmy of the control region. Gene 148:357–362

Xu X, Arnason U (1997) The complete mitochondrial DNA sequence of the white rhinoceros, *Ceratotherium simum,* and comparison with the mtDNA sequence of the Indian rhinoceros, *Rhinoceros unicornis.* Mol Phylogenet Evol 7:189–194

Xu X, Gullberg A, 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, 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