J Mol Evol (1998) 46:314–326

JOURNAL OF MOLECULAR EVOLUTION

Molecular Phylogenetic Inference of the Woolly Mammoth *Mammuthus primigenius*, Based on Complete Sequences of Mitochondrial Cytochrome *b* and 12S Ribosomal RNA Genes

Miyuki Noro,[1] Ryuichi Masuda,[1,2,*] Irena A. Dubrovo,[3] Michihiro C. Yoshida,[1,2] Makoto Kato[4]

[1] Division of Biological Science, Graduate School of Environmental Earth Science, Hokkaido University, Sapporo 060, Japan
[2] Chromosome Research Unit, Faculty of Science, Hokkaido University, Sapporo 060, Japan
[3] Paleontological Institute, Russian Academy of Sciences, 117647 Moscow, Russia
[4] Department of Earth and Planetary Science, Faculty of Science, Hokkaido University, Sapporo 060, Japan

Received: 8 April 1997 / Accepted: 23 July 1997

Abstract. Complete sequences of cytochrome *b* (1,137 bases) and 12S ribosomal RNA (961 bases) genes in mitochondrial DNA were successfully determined from the woolly mammoth (*Mammuthus primigenius*), African elephant (*Loxodonta africana*), and Asian elephant (*Elephas maximus*). From these sequence data, phylogenetic relationships among three genera were examined. Molecular phylogenetic trees reconstructed by the neighbor-joining and the maximum parsimony methods provided an identical topology both for cytochrome *b* and 12S rRNA genes. These results support the ''*Mammuthus-Loxodonta*'' clade, which is contrary to some previous morphological reports that *Mammuthus* is more closely related to *Elephas* than to *Loxodonta*.

Key words: Mammoth — *Mammuthus primigenius* — Proboscidea — Mitochondrial DNA — Molecular phylogeny — Elephant evolution

* *Present address:* Chromosome Research Unit, Faculty of Science, Hokkaido University, Sapporo 060, Japan; e-mail: masudary@ees.hokudai.ac.jp
Correspondence to: R. Masuda

Introduction

Among the extinct genus *Mammuthus*, *M. primigenius* (woolly mammoth) is known as a well-preserved fossil from Siberian permafrost. *Mammuthus* is classified in the subfamily Elephantinae with two extant genera, *Loxodonta* represented by *L. africana* (African elephant) and *Elephas* represented by *E. maximus* (Asian elephant). These three genera are believed to have originated in Africa and to have diverged from *Primelephas gomphotheroides* before the early Pliocene, approximately 5 million years ago (Maglio 1973). Phylogenetic relationships among these three genera remain unclear because of lack of fossil evidence that clearly shows divergence process. This evolutionary problem has been argued repeatedly and is still a controversy, although various approaches have been carried out in order to resolve their trichotomy. Even with many discussions over this problem based on morphological data (e.g., Shoshani et al. 1985b; Valente 1983) and immunological analyses (Lowenstein 1985; Shoshani et al. 1985a), it is still in dispute.

Lowenstein (1985) showed immunological equidistance among the three genera. Valente (1983) compared hair structure among the genera and supported the equidistancy as well. A closer relationship between *Mammuthus* and *Elephas* than between *Mammuthus* and *Lox-*

odonta, based on dentition, is believed to be classical hypothesis. Shoshani et al. (1985b) examined nondental osteological and immunochemical characters and supported this hypothesis. Because dental morphologies of early elephant groups are quite similar, however, the dental differences among them are hardly distinguishable (Kamei 1991). Moreover, the evolutionary rate of morphology, especially in dentition, is variable (Kamei 1991; Shoshani and Tassy 1996). Therefore, it is difficult to resolve the trichotomy of these three elephants based on morphological analyses.

There have been some phylogenetical studies based on nucleotide sequence analyses. By contrast to morphological characters, diversity of molecular sequences evolves at an almost constant rate for each gene region (Zuckerkandl and Pauling 1962). The notion of a molecular clock can be applied especially to closely related species. For evolutionary study of closely related taxa, estimation of sequence differences of mitochondrial DNA is a useful tool, since this molecule evolves more rapidly than nuclear genome (Brown et al. 1979).

Höss et al. (1994) reported the highly diverged sequences of partial mitochondrial 16S rRNA gene region (92 bases) from four specimens of *M. primigenius.* Hagelberg et al. (1994) examined the partial cytochrome *b* gene region (242 bases) of two mammoths and suggested that *Mammuthus* is slightly closer to *Loxodonta* than to *Elephas*, though the relationship was not resolved conclusively.

After proboscideans had distributed all over the world except Australia and Antarctica, most of them have already been extinct. Therefore, *Loxodonta africana* and *Elephas maximus* are the only survivors in the order Proboscidea. Their closest outgroup among living mammals is sirenians (e.g., dugong and manatee), which comprise a distinct taxonomic order—Sirenia. Yang et al. (1996) reported the usefulness of fossil DNA sequences from the American mastodon *Mammut americanum* (an extinct proboscidean, classified to the different family Mammutidae) as an outgroup for resolving phylogeny of Elephantinae. Based on analysis of the partial cytochrome *b* sequences (228 bases), they showed a closer clustering of *Mammuthus* (two distinct individuals) and *Elephas* by using the American mastodon sequence as an outgroup. However, such sequence lengths previously investigated are too short to conclude their phylogenetic relationships.

Because these three genera have highly diverged and because divergence of three genera probably occurred within a short period, the greater parts of observed base substitutions were expected to be accumulated after final genus divergence. Therefore additional longer sequences of more genes are needed to raise the informative substitutions and to establish a reliable conclusion for this problem. Recently, Ozawa et al. (1997) analyzed a longer partial sequence (1,005 bases) of the mammoth cytochrome *b* gene and suggested that the mammoth is more closely related to the Asian elephant than to the African elephant. Though the sequence length they examined is much longer than those of previous studies, the African elephant cytochrome *b* sequence they used (obtained from Irwin et al. 1991) contains a three-base insertion which has not been detected in any other mammal. So it should be confirmed whether or not such an insertion is generally observed in African elephants.

The aim of this study is to delineate the evolutionary history of these three genera by means of DNA sequence analysis. We here report the longest (complete) cytochrome *b* sequences of *Mammuthus primigenius.* In addition, this is the first report of the complete 12S rRNA sequence of the mammoth. Based on analysis of these new sequence data and comparison with previous morphological data, we discuss phylogenetic relationships among these genera.

Materials and Methods

Sample Source and DNA Extraction. In the course of experiment, disposable and DNA-free tips and tubes were used. All glass implements were used after sterilization at 180°C for 1 h. STE buffer (100 mM NaCl, 10 mM Tris-HCl, 1 mM EDTA), TE buffer (10 mM Tris-HCl, 1 mM EDTA), and some of plastic implements were autoclaved at 120°C for 20 min.

Tissues of *Loxodonta africana* (LAF1–4) and *Elephas maximus* (EMA1–4) were offered from Asahikawa Asahiyama Zoo, Kobe Ohji Zoo, Gunma Safari Park, and Hiroshima Asa Zoo. LAF2 and EMA2 were muscle tissues, and the other specimens were body hairs. DNA extractions were performed by the proteinase K/sodium dodecyl sulfate (SDS) and the phenol/chloroform/isoamyl alcohol method (Sambrook et al. 1989). All samples were washed with STE buffer several times before chemical treatment. Then, each tissue was homogenized in 500 μl of STE buffer using a small glass homogenizer. Hair samples were excluded from above homogenization step. SDS and proteinase K were added so as to make a final concentration of 0.5% and 25 μg/ml, respectively. The mixtures were incubated at 37°C overnight and extracted with an equal volume of phenol/chloroform/isoamyl alcohol (25:24:1) and chloroform/isoamyl alcohol (24:1) until the insoluble matter was completely eliminated. All extracts, except the sample EMA2, were concentrated to 40–50 μl using Centricon30 microconcentrators (Amicon).

The specimen of *Mammuthus primigenius* (MPR1), supplied by Dr. I.A. Dubrovo, was the dried muscle tissue of the specimen excavated from the Pyasna River Valley in the Taimyr Peninsula, Russia, and dated at least 25,000 years before present by the radiocarbon method (Sample T-462 in Heintz 1966). It has been stored in a desiccater at room temperature for over 30 years without any chemical treatment. Approximately 0.5 × 0.5 × 0.5 cm of tissue was applied to DNA extraction performed by the phenol/chloroform/isoamyl alcohol method using the same procedure as for extant species except that the volume of STE buffer was increased from 500 μl to 10 ml. STE buffer without any tissue was extracted by the same procedure and used as a negative control of the following PCR amplification.

PCR Amplification. The mitochondrial cytochrome *b* and 12S rRNA gene regions were amplified due to polymerase chain reaction

Table 1. Conditions and results of PCR amplification for mammoth (MPR1) cytochrome *b* gene regions[a]

Primer pair	Annealing (°C, min)	Cycles	Expected length of PCR product	Result of PCR
E-BOL/Cb-EOH	55, 1	40	111 bases	+
Cb-EA/Cb-EB	55, 1	35	116 bases	+
Cb-EC/Cb-ED	55, 0.5	30	119 bases	+
Cb-EE/Cb-EF	60, 0.5	45	129 bases	+
Cb-EGL/Cb-EGH	53, 1	40	108 bases	+
Cb-EHL/Cb-EHH	53, 1	40	98 bases	+
Cb-EIL/Cb-EIH	55, 1	40	105 bases	+
Cb-EJL/Cb-EJH	52, 1	40*	116 bases	+*
Cb-EKL/Cb-EKH	52, 1	45*	96 bases	−*
Cb-EKL2/Cb-EKH	48, 1	45*	108 bases	+*
Cb-ELL/Cb-ELH	53, 1	40	93 bases	+
Cb-EML/Cb-EMH	54, 1	40	99 bases	−
Cb-EML2/Cb-EMH	50, 1	40*	107 bases	+*
Cb-ENL/T-BNH	47, 1	45*	94 bases	+*
E-BOL/T-BNH	47, 1	45*	1,175 bases	−*

[a] Denaturing step (94°C, 1 min) and extension step (72°C, 1 min) are common in each reaction cycle while annealing step is slightly modified as above for primer pairs. Reactions for E-BOL/T-BNH included relatively longer denaturing step (94°C, 1 min) and extension step (72°C, 2.5 min). All reactions were followed by block step (72°C, 10 min). *: 2nd PCR was performed. +: Expected size fragment was amplified. −: Not amplified

(PCR) of Kocher et al. (1989). Primer sequences used in this study are shown in Appendix 1 (cytochrome *b*) and Appendix 2 (12S rRNA). To avoid co-amplification of any contaminated DNA (especially of human), most primers (designated with ''−E'') were designed as specific for the elephant sequences reported by Irwin et al. (1991).

PCR reactions for extant species were performed using primers E-BOL and T-BNH to amplify the locus containing the entire cytochrome *b* gene region. Primers F-E1L, V-B8H, and V-E8H2 were used to amplify the entire 12S rRNA gene region. Each 0.5 μl extract was subjected to symmetric PCR. Amplifications were performed in a total volume of 25 μl containing 1.25 U *Taq* DNA polymerase (Takara), 10 mM Tris-HCl pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$, 0.001% gelatin, 200 μM dATP, 200 μM dGTP, 200 μM dCTP, 200 μM dTTP (TaKaRa), and 250 nM of each primer. Reaction conditions of symmetric PCR for each pair of primers were as follows: E-BOL/T-BNH, 35 cycles of [94°C for 1 min, 47°C for 1 min, 72°C for 2.5 min] and 20 cycles of connective second PCR with the identical temperature profile; F-E1L/V-B8H and F-E1L/V-E8H2, 35 cycles of [94°C for 1 min, 60°C for 1 min, 72°C for 2.5 min]. These runs were completed by the block step of 72°C for 10 min. Asymmetric PCR was performed in accordance with the method described by Gyllenstein and Erlich (1988) with the same conditions as symmetric PCR but 30 reaction cycles. The reaction mixture basically included the same contents as for symmetric PCR. However, it contained the primers at a 1:100 (2.5 nM:250 nM) ratio, and 1 μl of the symmetric PCR product was added as a template in a total volume of 100 μl.

Primers were used as to amplify the short loci less than 150 bases for the mammoth specimen because the probable DNA fragmentation was expected. Amplifications for some longer loci were also performed in the identical reaction solution. The reaction condition of symmetric PCR for each pair of primers is shown in Table 1 (cytochrome *b*) and Table 2 (12S rRNA). The second PCR was performed if necessary. Asymmetric PCR was performed with the same conditions as for symmetric PCR except for 30 reaction cycles. For check of DNA amplification, 10 μl of the PCR product was electrophoresed on a 3% agarose gel and stained with ethidium bromide.

Direct Sequencing of PCR Products. Asymmetric PCR products were concentrated to 40–50 μl using Centricon30 microconcentrators (Amicon). Each 7 μl of the concentrate was subjected to sequencing with a Sequenase version 2.0 kit (United States Biochemical) and [α-^{32}P]dCTP (Amersham). One picomole of an appropriate primer was used as a sequencing primer. Sequencing reaction was performed with the dideoxynucleotide chain reaction method (Sanger et al. 1977).

The resultant products were electrophoresed on 6% polyacrylamide gels with 7.0 M urea. Gels dried on a filter paper were exposed to an X-ray film (Fuji RX) for 1 day or more if necessary.

Sequence Analysis. Sequence alignment was made with GeneWorks computer software (IntelliGenetics). Cytochrome *b* sequences of the dugong (*Dugong dugong*: DDU) (Irwin and Arnason 1994), black rhinoceros (*Diceros bicornis*: DBI) (Irwin et al. 1991), Grevy's zebra (*Equus grevyi*: EGR) (Irwin et al. 1991), and domestic cow (*Bos taurus*: BTA) (Anderson et al. 1982) were used as outgroups. The alignment was analyzed with the Clustal W 1.5 program package (Thompson et al. 1994), PHYLIP Ver. 3.572 (Felsenstein 1996). Multiple alignment for 12S rRNA sequences was made using Clustal W, with minor modifications made by eyes; 12S rRNA sequences of the dugong (DDU) (Lavergne et al. 1996), manatee (*Trichechus manatus*: TMA) (Lavergne et al. 1996), cape hyrax (*Procavia capensis*: PCA) (Lavergne et al. 1996), tree hyrax (*Dendrohyrax dorsalis*: DDO) (Douzery and Catzeflis 1995), rhinoceros (*Ceratotherium simum*: CSI) (Douzery and Catzeflis 1995), Grevy's zebra (EGR) (Douzery and Catzeflis 1995), and domestic cow (BTA) (Anderson et al. 1982) were used as outgroups. The published sequences of the African elephant (LAF) (Lavergne et al. 1996) and the Asian elephant (EMA) (Lavergne et al. 1996) were analyzed together.

Phylogenetic tree construction according to the neighbor-joining method (Saitou and Nei 1987) and bootstrap analysis (Felsenstein 1985) were performed with Clustal W (Thompson et al. 1994). Genetic distances were corrected by Kimura's two parameter method (Kimura 1980). Bootstrap values were derived from 1,000 replications. Maximum parsimonious trees were constructed with the DNAPARS program, and bootstrapping was done with SEQBOOT, DNAPARS, and CONSENSE programs contained in PHYLIP (Felsenstein 1996). Most parsimonious trees were constructed via 10 times random input using the J (jumble) option on the DNAPARS program. Bootstrapped 100 data sets were generated with the SEQBOOT program, and the resultant sequences were randomly input 10 times using the J option on the

Table 2. Conditions and results of PCR amplification for mammoth (MPR1) 12S rRNA gene regions[a]

Primer pair	Annealing (°C, min)	Cycles	Expected length of PCR product	Result of PCR
F-E1L/12S-E1H	60, 1	40	128 bases	+
12S-E2L/12S-H	58, 1	40	126 bases	+
12S-E3L/12S-E3H	55, 1	40	130 bases	+
12S-E4L/12S-E4H	50, 1	45	125 bases	+
12S-E5L/12S-E5H	58, 1	40	128 bases	+
12S-E6L/12S-E6H	50, 1	40	114 bases	+
12S-E7L/12S-E7H	60, 1	40	119 bases	+
12S-E8L/V-B8H	55, 1	45*	157 bases	+*
12S-E8L/V-E8H2	55, 1	40	138 bases	+
F-E1L/V-B8H	55, 1	45*	1,001 bases	−*
F-E1L/V-E8H2	55, 1	45*	982 bases	−*

[a] Denaturing step (94°C, 1 min) and extension step (72°C, 1 min) are common in each reaction cycle while annealing step is slightly modified as above for primer pairs. Reactions for F-E1L/V-B8H and F-E1L/V-E8H2 included relatively longer denaturing step (94°C, 1 min) and extending step (72°C, 2.5 min). All reactions were followed by block step (72°C, 10 min). *: 2nd PCR was performed. +: Expected size fragment was amplified. −: Not amplified

DNAPARS program. Bootstrap values were calculated with the CONSENSE program.

Results

PCR Amplification and Direct Sequencing of PCR Product

DNA fragments with expected sizes were successfully PCR amplified from all specimens of extant elephants (Tables 1 and 2). No amplification was observed on the negative control.

Symmetric and asymmetric PCR amplifications of short loci from MPR1 extracts were successfully done (Tables 1 and 2). PCR from the mammoth DNA, with the pairs of primers E-BOL/T-BNH (containing the complete cytochrome *b* gene region), F-E1L/V-B8H, and F-E1L/V-E8L2 (containing the complete 12S rRNA gene region) for amplification of long fragments more than 1 kb, generated no visible product on the gel electrophoresis, indicating possible fragmentation or degradation of mammoth DNA.

By connecting partial sequences, complete sequences for cytochrome *b* (1,137 bases) and 12S rRNA (961 bases) gene regions were determined for all samples. These sequences were aligned with outgroup sequences (Appendix 3 for cytochrome *b*; Appendix 4 for 12S rRNA). Of four African and four Asian elephants, all animals were sequenced for cytochrome *b*, while two African and two Asian elephants were done for 12S rRNA. Cytochrome *b* nucleotide sequences of elephants encode predicted 378-amino-acid sequences that terminate with a TAA stop codon. The three-base insertion shown by Irwin et al. (1991) was not detected on any sample examined in this study. The nucleotide sequence data reported in this paper will appear in DDBJ, EMBL, and GenBank nucleotide sequence databases with the following accession numbers: D50841-D50847, D84150-D84152, AB002411, and AB002412.

Cytochrome b *Sequence Analysis*

The number of substitutions and percentage differences among sequences are shown in Table 3. Two African elephants (LAF1 and LAF3) and two Asian elephants (EMA3 and EMA4) shared identical sequences, respectively (Appendix 3). Most substitutions detected among the mammoth and extant elephants were transitional mutations. Sequence differences between *Mammuthus* and *Loxodonta* were 5.4–5.7% and those between *Mammuthus* and *Elephas* were 6.6–7.2%. Those between *Loxodonta* and *Elephas* showed larger values (7.1–7.7%) (Table 3).

Figure 1A shows the neighbor-joining tree. The ''*Loxodonta-Mammuthus*'' clade was supported by 92% bootstrap value. One most parsimonious tree was found regardless of input order, and its topology was identical to the neighbor-joining tree (Fig. 1B). Bootstrap values are shown above the nodes (some data sets generate multiple topologies). The ''*Loxodonta-Mammuthus*'' clade showed 73% bootstrap value (Fig. 1B). In order to check the possibility of generating other topologies aiming at the subfamily Elephantinae including the mammoth and extant elephants, the parsimony scores were compared among three distinct topologies which include the most parsimonious one. Even if the ''*Loxodonta-Mammuthus*'' clade is supported relatively strongly, the number of steps for most parsimonious topology of ((LAF, MPR), EMA) was three or five different from the

Table 3. Pairwise sequence divergence estimates (above diagonal) and numbers of substitution (below diagonal) based on cytochrome *b* gene sequences (1,137 bases)[a]

	MPR1	LAF1,3	LAF2	LAF4	EMA1	EMA2	EMA3,4	DDU	BTA
MPR1		5.55	5.45	5.65	6.63	7.13	6.73	28.29	31.09
LAF1,3	60 (54/6)		0.80	0.80	7.26	7.56	7.36	28.32	31.26
LAF2	59 (53/6)	9 (9/0)		0.18	7.16	7.46	7.26	28.46	31.55
LAF4	61 (55/6)	9 (9/0)	2 (2/0)		7.36	7.66	7.46	28.46	31.69
EMA1	71 (64/7)	77 (74/3)	76 (73/3)	78 (75/3)		1.70	0.26	28.31	30.90
EMA2	76 (69/7)	80 (77/3)	79 (76/3)	81 (78/3)	19 (19/0)		1.61	28.73	31.51
EMA3,4	72 (65/7)	78 (75/3)	77 (74/3)	79 (76/3)	3 (3/0)	18 (18/0)		28.31	30.80
DDU	264 (154/110)	264 (156/108)	265 (157/108)	265 (157/108)	264 (155/109)	267 (158/109)	264 (155/109)		27.33
BTA	287 (146/141)	288 (149/139)	290 (151/139)	291 (152/139)	286 (144/142)	290 (148/142)	285 (143/142)	259 (126/133)	

[a] Above diagonal includes pairwise sequence divergence estimates in percentage distances corrected by Kimura's two-parameter model. Below diagonal includes numbers of total substitutions and transition/transversion in parentheses

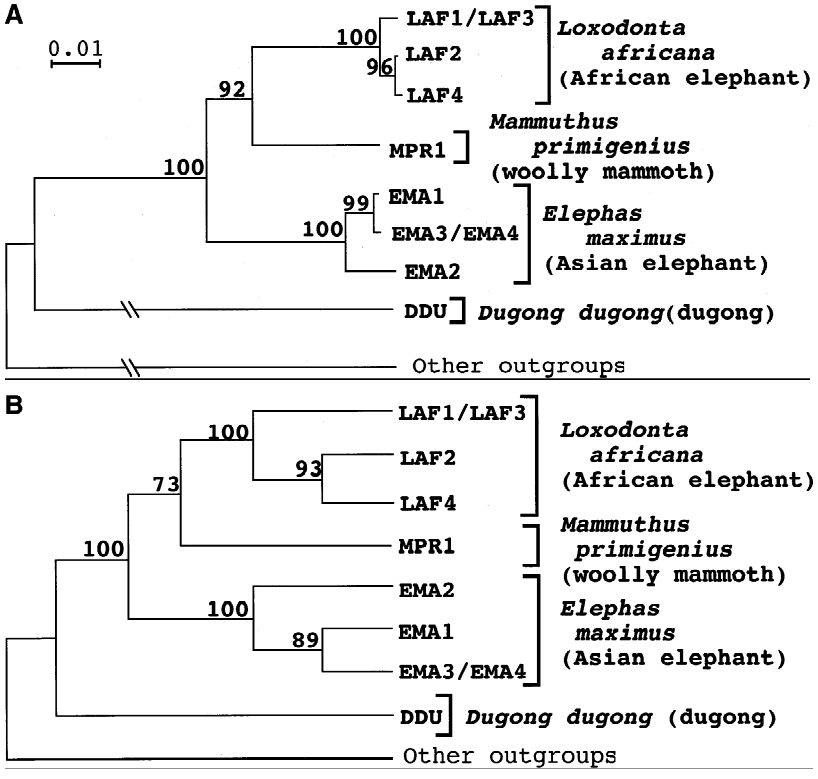


Fig. 1A. Phylogenetic tree of cytochrome *b* sequences (1,137 bases), constructed by the neighbor-joining method. *Numbers* on internal branches are bootstrap values derived from 1,000 replications. The *scale* indicates evolutionary distance of substitution per site, estimated by Kimura's two-parameter method. **B** Phylogenetic tree of cytochrome *b* sequences (1,137 bases), constructed by the maximum parsimony method. *Numbers* on internal branches are bootstrap values derived from 100 replications. The dugong and other species including the black rhinoceros, Grevy's zebra, and domestic cow were used as outgroups.

others. Two alternative trees tested were not significantly different than the most parsimonious one.

12S rRNA Sequence Analysis

Table 4 shows the number of substitutions and the percentage differences among sequences. Most substitutions detected among the mammoth and living elephants were transitions. Sequence differences between *Mammuthus* and *Loxodonta* were 1.0–1.4% and those between *Mammuthus* and *Elephas* were 1.1–1.3%. Those between *Loxodonta* and *Elephas* indicated larger values (1.5–2.3%).

The neighbor-joining tree (Fig. 2A) supported the ''*Loxodonta-Mammuthus*'' clade with 55% bootstrap

Table 4. Pairwise sequence divergence estimates (above diagonal) and numbers of substitution (below diagonal) based on complete 12S rRNA gene sequences[a]

	MPR1	LAF1	LAF2	LAF	EMA1	EMA2	EMA	DDU	BTA
MPR1		1.37	1.05	1.37	1.16	1.26	1.27	22.12	28.75
LAF1	13 (13/0)		0.31	0.00	1.91	2.23	2.24	20.99	28.93
LAF2	10 (10/0)	3 (3/0)		0.31	1.58	1.91	1.91	20.84	28.75
LAF	13 (13/0)	0 (0/0)	3 (3/0)		1.91	2.23	2.24	20.99	28.93
EMA1	11 (10/1)	18 (17/1)	15 (14/1)	18 (17/1)		0.31	0.31	20.70	28.93
EMA2	12 (10/2)	21 (19/2)	18 (16/2)	21 (19/2)	3 (2/1)		0.00	21.14	29.25
EMA	12 (10/2)	21 (19/2)	18 (16/2)	21 (19/2)	3 (2/1)	0 (0/0)		20.91	29.02
DDU	174 (125/49)	169 (120/49)	168 (119/49)	169 (120/49)	167 (119/48)	170 (121/49)	168 (119/49)		22.01
BTA	215 (137/78)	218 (138/80)	217 (137/80)	218 (138/80)	218 (138/80)	220 (139/81)	218 (137/81)	174 (110/64)	

[a] Above diagonal includes pairwise sequence divergence estimates in percentage distances corrected by Kimura's two-parameter model. Below diagonal includes numbers of total substitutions and transition/transversion in parentheses

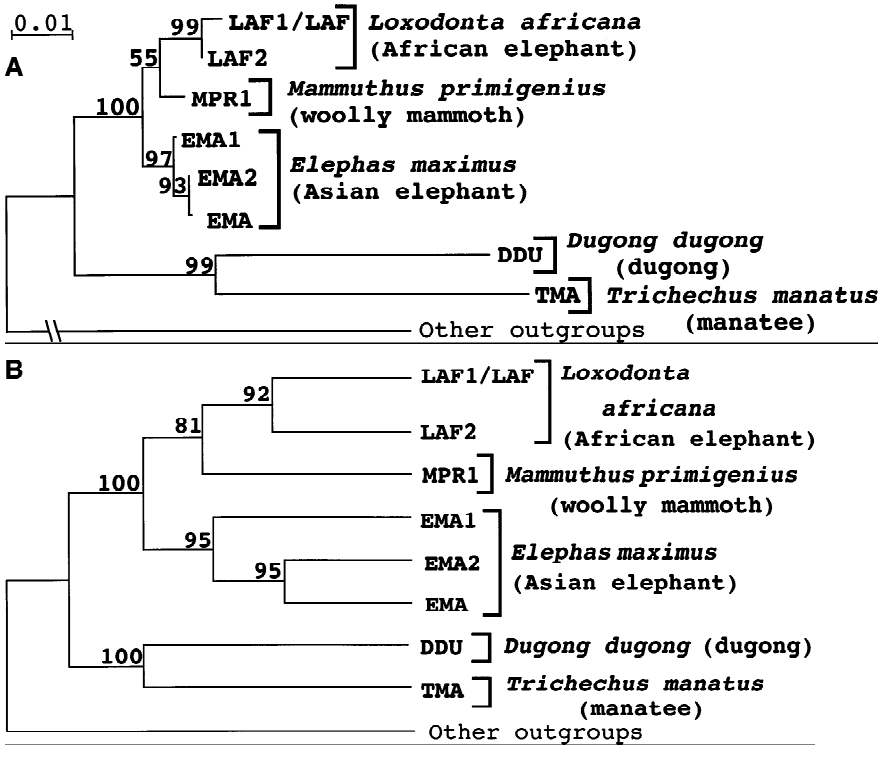


Fig. 2A. Phylogenetic tree of 12S rRNA sequences (961 bases for elephants), constructed by the neighbor-joining method. *Numbers* on internal branches are bootstrap values derived from 1,000 replications. The *scale* indicates evolutionary distance of substitution per site, estimated by Kimura's two-parameter method. **B** Phylogenetic tree of 12S rRNA sequences (961 bases for elephants), constructed by the maximum parsimony method. *Numbers* on internal branches are bootstrap values derived from 100 replications. The dugong, manatee, cape hyrax, tree hyrax, rhinoceros, Grevy's zebra, and domestic cow were used as outgroups.

value. One most parsimonious tree was found regardless of input order, and its topology is identical to the neighbor-joining tree (Fig. 2B). Bootstrap values are shown above the nodes. (Some data sets generate multiple topologies.) In the parsimonious tree, the ''*Loxodonta-Mammuthus*'' clade was supported with 81% bootstrap value (Fig. 2B). For a check of the possibility of generating other topologies in the subfamily Elephantinae, comparison of the parsimony scores was made on three distinct topologies including the most parsimonious one. The ''*Loxodonta-Mammuthus*'' clade was supported relatively strongly, and the number of steps for the most parsimonious topology of ((LAF, MPR), EMA) is five to 19 steps different from the others. The other two trees tested were significantly different than the best one.

Discussion

Phylogenetic Relationships Among the Subfamily Elephantinae

Phylogenetic analyses of complete sequences for both cytochrome *b* and 12S rRNA supported a closer relationship of *Mammuthus* to *Loxodonta* (Figs. 1 and 2) than to *Elephas.* An identical topology was generated using both the neighbor-joining method and the maximum parsimony method for both the cytochrome *b* and 12S rRNA sequences.

Based on cytochrome *b* sequences, the ''*Mammuthus-Loxodonta*'' clade was supported by high bootstrap values: 92% (neighbor-joining method, Fig. 1A) and 73% (maximum parsimony method, Fig. 1B). Parsimony test showed that alternative topologies need a few more steps than the most parsimonious one.

On analysis of 12S rRNA gene sequences, the ''*Mammuthus-Loxodonta*'' clade was also supported by 55% (neighbor-joining method, Fig. 2A) and 81% (maximum parsimony method, Fig. 2B) bootstrap values. Alternative topologies were significantly different than the most parsimonious one.

The rate of base substitution on 12S rRNA is much slower than that of cytochrome *b,* as observed in other mammals (Thomas et al. 1989; Masuda et al. 1996). Observed 12S rRNA variations of 1.0–2.3% among three genera are rather small, compared with 5.4–7.7% difference of cytochrome *b.* There is the same problem as for a short sequence. As mentioned in the Introduction, previous paleontological data showed that these three genera have highly diverged, and their divergences are expected to have occurred within a short period. Therefore, generally most of base substitutions in the genome could have occurred after the divergence. Actually, the sequence divergence among these three genera shows no large disparity, especially in 12S rRNA, suggesting the very phylogenetically close relationship among these three genera. Therefore, all nucleotides of cytochrome *b* sequences are presumably more informative for phylogenetic analysis of Elephantinae.

On the other hand, Ozawa et al. (1997) determined a partial sequence (1,005 bases) of cytochrome *b* from a different mammoth specimen. Between their data and our sequences, the nucleotide difference was only 0.5% (5/1,005 bases; all substitutions were transitions), indicating an intraspecific variation similar to that of living elephants (Table 3). From analysis of the first/second positions as well as amino acid sequences deduced from the nucleotides, Ozawa et al. (1997) suggested a closer relationship between the mammoth and the Asian elephant. However, because of such a small level of nucleotide difference among the three elephant species, it is reasonable for all positions in codons to be included for analysis. Moreover, the African elephant sequence data (Irwin et al. 1991) used by Ozawa et al. (1997) has a three-base insertion which has not been found in any other mammals. The sequence of Irwin et al. (1991) is also somewhat divergent from those of African elephants determined in our study. This is why we did not include their sequence data for the analysis.

In this study, the numbers of substitutions (transversion and transition) of cytochrome *b* between the outgroups and each elephant sequence are almost equivalent (Table 3). However, in comparison with living elephants, the *Mammuthus* sequence showed slightly more transversions, while fewer transitions were observed (Table 3). This causes the ''*Elephas-Loxodonta*'' clustering when only transversion substitutions are used for phylogenetic analyses (data not shown). On account of the effect of transversion frequency, predicted amino acid sequence also supports the ''*Elephas-Loxodonta*'' clustering (data not shown). The transversion ratios in all substitutions are 10% or less, and it is probable that transitional substitutions have not saturated yet. Therefore, there is no need to exclude transitions.

In this study, complete sequences of cytochrome *b* and 12S rRNA of the woolly mammoth *Mammuthus primigenius* were determined for the first time. The phylogenetic analysis of the sequences suggests that *Mammuthus* and *Loxodonta* are more closely related to each other than to *Elephas.* The larger sample size (both the number of specimens and genes, and the sequence length examined) is necessary to further understand the precise phylogenetic relationships in Elephantinae.

Acknowledgments. We thank K. Murata (Kobe Ohji Zoo), M. Kosuge (Asahikawa Asahiyama Zoo), Y. Fukumoto (Hiroshima Asa Zoo), and S. Kawakami (Gunma Safari Park) for supplying elephant samples. We are also grateful to Dr. N. Minoura (Hokkaido University) for critically reading the manuscript. This study was supported in part by a Grant-in-Aid for Scientific Research on Priority Areas (Molecular Evolution) from the Ministry of Education, Science, Sports, and Culture, Japan to R.M.

References

Anderson S, de Bruijn MHL, Coulson AR, Eperon IC, Sanger F, Young IG (1981) Complete sequence of bovine mitochondrial DNA: conserved features of the mammalian mitochondrial genome. J Mol Biol 156:683–717

Anderson S, Bankier AT, Barrell BG, de Bruijn MHL, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJH, Staden R, Young IG (1982) Sequence and organization of the human mitochondrial genome. Nature 290:457–465

Brown WM, George M Jr, Wilson AC (1979) Rapid evolution of animal mitochondrial DNA. Proc Natl Acad Sci USA 76:1967–1971

Douzery E, Catzeflis FM (1995) Molecular evolution of the mitochondrial 12S rRNA in Ungulata (Mammalia). J Mol Evol 41:622–636

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783–791

Felsenstein J (1996) PHYLIP (phylogeny inference package) Version 3.5c

Gyllenstein UB, Erlich HA (1988) Generation of single stranded DNA by the polymerase chain reaction and its application to direct sequencing of the HLA-DQA locus. Proc Natl Acad Sci USA 85: 7652–7656

Hagelberg E, Thomas MG, Cook CE Jr, Sher AV, Baryshnikov GF, Lister AM (1994) DNA from ancient mammoth bones. Nature 370: 333–334

Heintz A (1966) New radio carbon (C^{14}) age determinations of mammal remains from the permafrost in Siberia. Norsk Giologisk Tidsskrift 46:215–217

Höss M, Pääbo S, Vereshchagin NK (1994) Mammoth DNA sequences. Nature 370:333

Irwin DM, Arnason U (1994) Cytochrome *b* gene of marine mammals: phylogeny and evolution. J Mamm 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

Kamei T (ed) (1991) Japanese proboscidean fossils. Tsukiji-Shokan, Tokyo, 273 pp

Kimura M (1980) A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 16:111–120

Kocher TD, Thomas WK, Meyer A, Edwards SV, Pääbo S, Villablanca FX, Wilson AC (1989) Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. Proc Natl Acad Sci USA 86:6196–6200

Lavergne A, Douzery E, Stichler T, Catzeflis FM, Springer MS (1996) Interordinal mammalian relationships: evidence for paenungulate monophyly is provided by complete mitochondrial 12S rRNA sequences. Mol Phylogenet Evol 6:245–258

Lowenstein JM (1985) Radioimmune assay of mammoth tissue. Acta Zool Fennica 170:233–235

Maglio VJ (1973) Origin and evolution of the Elephantidae. Trans Am Phil Soc [new ser.] 63:1–149

Masuda R, Lopez JV, Pecon Slattery J, Yuhki N, O'Brien SJ (1996) Molecular phylogeny of mitochondrial cytochrome *b* and 12S rRNA sequences in the Felidae: ocelot and domestic cat lineages. Mol Phylogenet Evol 6:351–365

Ozawa T, Hayashi S, Mikhelson VM (1997) Phylogenetic position of mammoth and Steller's sea cow within Tethytheria demonstrated by mitochondrial DNA sequences. J Mol Evol 44:406–413

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

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. 2nd ed. Cold Spring Harbor Laboratory, NY, New York

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

Shoshani J, Tassy P (eds) (1996) The Proboscidea: evolution and palaeoecology of elephants and their relatives. Oxford University Press, New York, 472 pp

Shoshani J, Lowenstein JM, Walz DA, Goodman M (1985a) Proboscidean origins of mastodon and woolly mammoth demonstrated immunologically. Paleobiology 11:429–437

Shoshani J, Walz DA, Goodman M, Lowenstein JM, Prychodko (1985b) Protein and anatomical evidence of the phylogenetic position of *Mammuthus primigenius* within the Elephantinae. Acta Zool Fennica 170:237–240

Thomas RH, Schaffner W, Wilson AC, Pääbo S (1989) DNA phylogeny of the extinct marsupial wolf. Nature 340:465–467

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680

Valente A (1983) Hair structure of the woolly mammoth, *Mammuthus primigenius* and the modern elephants, *Elephas maximus* and *Loxodonta africana*. J Zool Lond 199:271–274

Yang H, Golenberg EM, Shoshani J (1996) Phylogenetic resolution within the Elephantidae using fossil DNA sequence from the American mastodon (*Mammut americanum*) as an outgroup. Proc Natl Acad Sci USA 93:1190–1194

Zuckerkandl E, Pauling L (1962) Molecular disease, evolution, and genic diversity. In: Kasha M, Pullman B (eds) Horizons in biochemistry. Academic Press, New York, pp 189–225

Appendixes 1–4 Follow

Appendix 1. Primer sequences for PCR amplification and direct sequencing for cytochrome *b* gene regions[a]

Primer name		Sequence
E-BOL	: L14735	5′-AAACCATCGTTGTCATTCAACTA-3′
Cb-EOH	: H15842	5′-CTCCTAGTAGTGAGCCGAAA-3′
Cb-EA	: L14840	5′-CCATCCAACATATCAACATGATGAA-3′
Cb-EB	: H14957	5′-TTCAGCCGTAGTTTACATCTCG-3′
Cb-EC	: L14950	5′-CTGCATTTTCATCTATATCCCAT-3′
Cb-ED	: H15070	5′-GGTATTTCAAGTTTCCGAGTAT-3′
Cb-EE	: L15064	5′-GACGAAACATCTACTATGGGTCC-3′
Cb-EF	: H15191	5′-GATATAAGGGATTGCTGAGAGAAG-3′
Cb-EGL	: L15178	5′-ATATCATTCTGAGGGGCAACC-3′
Cb-EGH	: H15287	5′-GTAAATGGAAGAATGAAATGGAG-3′
Cb-EHL	: L15274	5′-TCAGTAGACAAAGCAACCTTAAAT-3′
Cb-EHH	: H15373	5′-CTGAGTCTGAAGTGAGGCCC-3′
Cb-EIL	: L15358	5′-CCTTTCTTCACGAAACAGGC-3′
Cb-EIH	: H15464	5′-GATAGTAGGGCTAGGAGTAG-3′
Cb-EJL	: L15452	5′-TTCCTAGGATTACTTATCCTAA-3′
Cb-EJH	: H15569	5′-GTAAGCAAAGAGAAAATATCAC-3′
Cb-EKL	: L15554	5′-ACCCCTTAAATAACCCCCCT-3′
Cb-EKL2	: L15542	5′-GCTGATCCACTTAATACTCCCC-3′
Cb-EKH	: H15651	5′-GTATGGAGAAGTGGTATTAATC-3′
Cb-ELL	: L15640	5′-CTAGCCCTACTCCTATCAATT-3′
Cb-ELH	: H15734	5′-TCATGTAAGTGTTAGTAAATCTAT-3′
Cb-EML	: L15726	5′-CCTATGTGCCTATTGCTGAAC-3′
Cb-EML2	: L15718	5′-CGACCTCTTAGCCAAGTCCTA-3′
Cb-EMH	: H15826	5′-GCAGGAAGGCTAGGATAATG-3′
Cb-ENL	: L15811	5′-CATTATTGGTCAAATAGCCTCA-3′
T-BNH	: H15915	5′-TCTCCTTCTCTGGTTTACAAGAC-3′

[a] The letters L and H refer to the light and heavy strands, respectively, and the numbers refer to the positions of the 3′ end of the primers in the complete human mitochondrial DNA sequence (Anderson et al. 1981)

Appendix 2. Primer sequences for PCR amplification and direct sequencing for 12S rRNA gene regions[a]

Primer name		Sequence
F-E1L	: L627	5′-GCAAGGTACTGAAAATACCTAGACGA-3′
12S-E1H	: H753	5′-CACTTGGAGTGTGTGCTTGATGTCA-3′
12S-E2L	: L751	5′-CTAAATCATCGCTGATCAAAGAGAG-3′
12S-H	: H876	5′-GCGGTTGCTGGCACGAAATTGACCA-3′
12S-E3L	: L874	5′-GTTAGACTAAGTTATCCTAATAAAGGA-3′
12S-E3H	: H1009	5′-AGTCACTTTCGTAGGCTATTTTGTC-3′
12S-E4L	: L1005	5′-CTTATACTAGCTGTTTAAAGCTCAAGA-3′
12S-E4H	: H1130	5′-CTGGCTAGTAGTTCTCTGGCGGAT-3′
12S-E5L	: L1121	5′-GCCCTAAACTTTGATAGCTACCTTT-3′
12S-E5H	: H1250	5′-GAAGATGGTGGTATATGGACTGAATT-3′
12S-E6L	: L1243	5′-ATGAACCCCGATACACCTTACCGT-3′
12S-E6H	: H1352	5′-ATAATAGAAAATGTAGCCCATCTTTGA-3′
12S-E7L	: L1349	5′-GGCCGAGGTGTCGCCTACGTGAC-3′
12S-E7H	: H1467	5′-CGGGCGGTGTGTACGCGCTTCAT-3′
12S-E8L	: L1462	5′-CTAAGAATAGAGAGCTTAATTGAACAA-3′
V-B8H	: H1620	5′-TCTTCTGGGTGTAGGCCAGATGCTTT-3′
V-E8H2	: H1604	5′-CCAGATGCTTTTTGTTAAGCTACACTT-3′

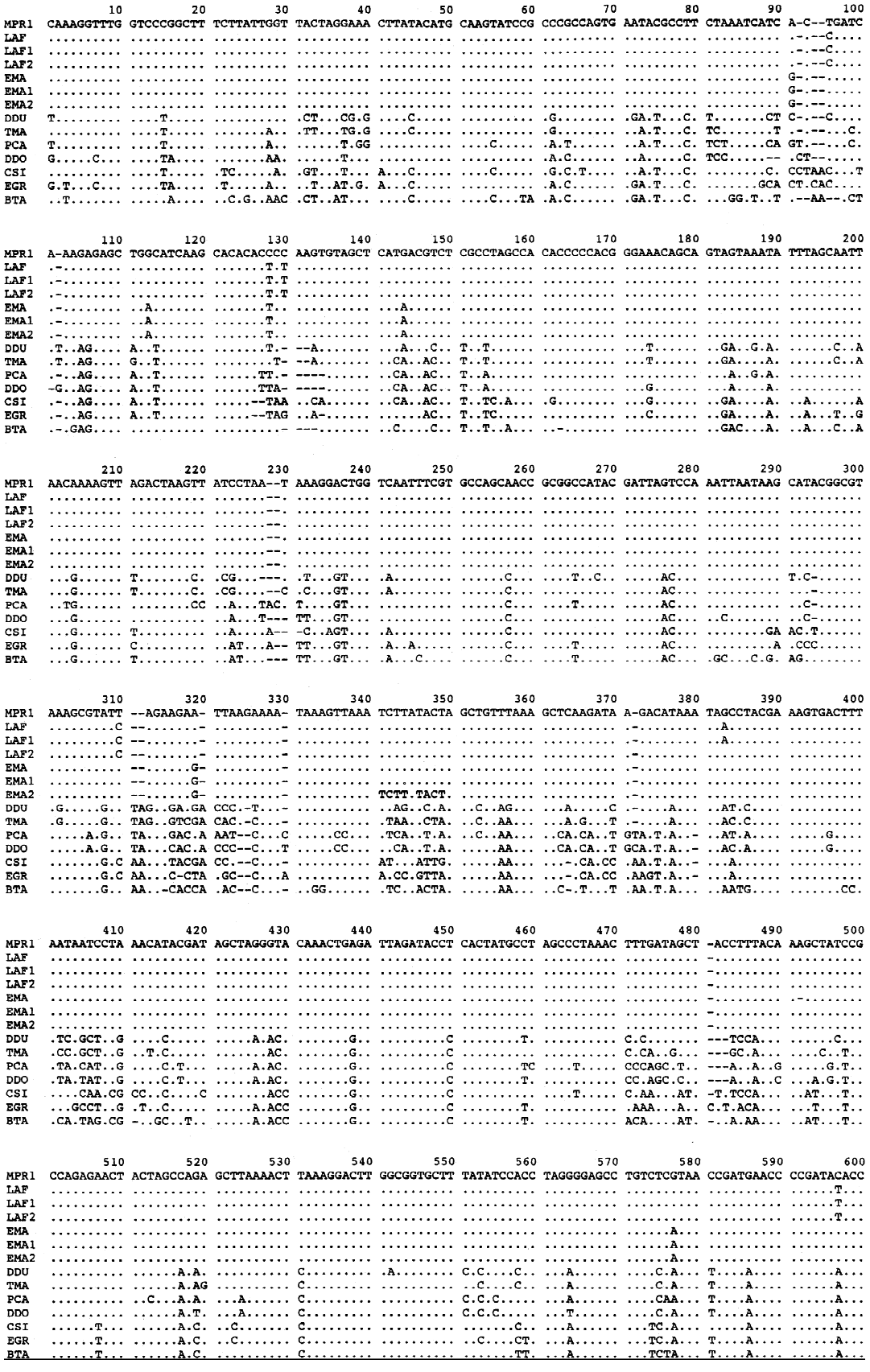
[a] The letters L and H refer to the light and heavy strands, respectively, and the numbers refer to the positions of the 3′ end of the primers in the complete human mitochondrial DNA sequence (Anderson et al. 1981)

```
          10         20         30         40         50         60         70         80         90        100
MPR1  ATGACCCACA TTCGAAAATC TCACCCCCTA CTTAAAATCC TTAATAAATC CTTCATTGAT CTACCTACCC CATCTAACAT CTCAACATGA TGAAATTTCG
LAF1  .......... .......... .......T.. .........A .C........ .......... .......... ....C..... .......... ..........
LAF2  .......... .......... .......T.. .........A .C........ .......... .......... ....C..... .......... ..........
LAF3  .......... .......... .......T.. .........A .C........ .......... .......... ....C..... .......... ..........
LAF4  .......... .......... .......T.. .........A .C........ .......... .......... ....C..... .......... ..........
EMA1  .......... CC........T .........G T........A .C..C..... .......... .....C.... .......... .......... ..........
EMA2  .......... CC........ .........G T........A .C..C..... .......... .....C.... .......... .......... ..........
EMA3  .......... CC........T .........G T........A .C..C..... .......... .....C.... .......... .......... ..........
EMA4  .......... CC........T .........G T........A .C..C..... .......... .....C.... .......... .......... ..........
DDU   ......A... .C........ A.....A... A.C....... .A..C..C.. .........C ..C....... .CGTA..T.. ....T..... .....C..T.
DBI   .....TA... .C..T..... C.....A... A.C.....TA .C...C.C.. A.....C..C .....C.... ....A..... T...G.C... .......T.
EGR   .....AA... .C..G..... ......A... A........A .C...C.C.. T..T.....C .....AG... .C..A..... T...T.A... .....C....
BTA   .....TA... .......G.. C.....A... A.A.....TG .A..C..TG. A.....C..C ..T..AG... ....A..... T...T.A... ..........

          110        120        130        140        150        160        170        180        190        200
MPR1  GCTCACTACT AGGAGCATGC CTAATTACCC AAATCCTAAC AGGGTTATTT CTAGCCATAC ATTATACACC TGACACAATA ACTGCATTTT CATCTATATC
LAF1  .......... .......... .......... .......... ...A.....C .......... .......... C......... .......... ..........
LAF2  .......... .......... .......... .......... ...A.....C .......... .......... .......... .......... ..........
LAF3  .......... .......... .......... .......... ...A.....C .......... .......... C......... .......... ..........
LAF4  .......... .......... .......... .G........ ...A.....C .......... .......... .......... .......... ..........
EMA1  .......... .......... .......... .......... ...A.....C .......... ....C..... .......... .......... ..........
EMA2  .......... ......G... .......... .......... ...A.....C .......... ....C..... .......... .......... ..........
EMA3  .......... .......... .......... .......... ...A.....C .......... ....C..... .......... .......... ..........
EMA4  .......... .......... .......... .......... ...A.....C .......... ....C..... .......... .......... ..........
DDU   ....C..... C..G...... ..G....TT. ....T..C.. ...A.....C ..G....... .C..C...T. A......C.. ..C.....C. .C..AG..A.
DBI   ....T..... ....ATC... .....CCTA. .......... C..AC..... ..T..T.... .......... A.......C. .....C..C. ....CG.TG.
EGR   ....C..C.. ....ATC... .....CCT.. .....T.... ...CC....C .......... .C..C...T. A.....G.C. ....C..C. ....CG.CA.
BTA   .T..C..C.. G...ATC... .....CCTA. .......C.. ...CC....C .....A.... .C..C...T. C.......C. ..A.....C. .C...G.TA.

          210        220        230        240        250        260        270        280        290        300
MPR1  CCATATCTGC CGAGATGTCA ACTACGGTTG AATTATTCGA CAACTACACT CAAACGGAGC ATCTATTTTC TTCCTCTGCC TATACACACA CATTGGACGA
LAF1  ......T... ........G. .......C.. .......... .......... .......... ...C...... .......... .......... ..........
LAF2  ......T... ........G. .......C.. .......... .......... .......... ...C...... ..T....... .......... ..........
LAF3  ......T... ........G. .......C.. .......... .......... .......... ...C...... .......... .......... ..........
LAF4  ......T... ........G. .......C.. .......... .......... .......... ...C...... ..T....... .......... ..........
EMA1  .......... .....C.... .......C.. .......... .......... .......... ........ .......... .......... ..........
EMA2  .......... .....C.... .......C.. .......... .....G.... .......... ......C..T .......... .......... ..........
EMA3  .......... .....C.... .......C.. .......... .......... .......... ........ .......... .......... ..........
EMA4  .......... .....C.... .......C.. .......... .......... .......... ........ .......... .......... ..........
DDU   ......T... ..G.....A. .......C.. .......... T.T..T...G .T.....G.. ...A..A... ..C..A.... .C...G.C.. .........C
DBI   ...C.....T .....G..A. .......C.. ......C..C T.C.....TG .......... ...C..A..T ...A...... ...T..TC.. ...A.....C
EGR   T..C..... .....C..T. .......A.. ........C T.C..C..TG .C.......... ...A..A..T ...A...... .C.T..TT.. .G.A.....C
BTA   .......... .....C..G. .......C.. ...C..C... T.CA.....G .......... T..A..G..T ...A.....T ....T.TG.. .G.A......

          310        320        330        340        350        360        370        380        390        400
MPR1  AACATCTACT ATGGGTCCTA CCTATACTCG GAAACCTGAA ATACCGGCAT TATACTACTA CTAATCACCA TAGCCACCGC CTTCATAGGA TATGTCCTTC
LAF1  .......... .......... .......... .....T.... .......... ....T..... .......... .......... .......... ..........
LAF2  .......... .......... .......... .....T.... .......... ....T..... .......... .......... .......... ..........
LAF3  .......... .......... .......... .....T.... .......... ....T..... .......... .......... .......... ..........
LAF4  .......... .......... .......... .....T.... .......... ....T..... .......... .......... .......... ..........
EMA1  .......... ....A..... .........A .......... ....A..... ....T..... .....T.... .......... .......... ..........
EMA2  .......... ....A..... T........A .......... .C..A..T.. .......... .......... .......... .......... ..........
EMA3  .......... ....A..... .........A .......... ....A..... ....T..... .....T.... .......... .......... ..........
EMA4  .......... ....A..... .........A .......... ....A..... ....T..... .....T.... .......... .......... ..........
DDU   GGA.....T. .C..C..A.. ......TC.A .......... .C.TT..T.. CG....G... ..C.CAGTT. ....T..T.. .........G ..C.....C.
DBI   GG.C....T. .C..A..... .ACC.T.CTA A......... .C.T...AG. ...T..... ..C.CAGTA. .......A.. A........C ........A.
EGR   GG.C..... .C..C..T.. .AC..T.CTA ..G..A.... .C.TT..A.. C..C.....T T.C.CAGTT. ....T..A.. A.....G..C .........A.
BTA   GG.T.A..T. .C.....T.. .ACT.TTCTA .....A.... ...TT..AG. A..C..T..G ..C.CAGTA. .......A.. A..T...... ..C.....A.

          410        420        430        440        450        460        470        480        490        500
MPR1  CGTGAGGACA AATATCATTC TGAGGGGCAA CCGTAATCAC TAACCTCTTC TCAGCAATTC CCTACATCGG CACAGACCTA GTAGAATGAA TCTGAGGAGG
LAF1  .......... .......... .......... .......... ......T... ........C. .T..T..... ....A..... .......... ..........
LAF2  .......... .......... .......... .......... ......T... ........C. .T.GT..... ....A..... .......... ..........
LAF3  .......... .......... .......... .......... ......T... ........C. .T..T..... ....A..... .......... ..........
LAF4  .......... .......... .......... .......... ......T... ........C. .T.GT..... ....A..... .......... ..........
EMA1  .A........ .......... .......... .......T.. .......... .......... .......... ....A..... .......... .T........
EMA2  .A........ .......... .......... .......T.. .......... .......... .......... ....A..... .......... .T........
EMA3  .A........ .......... .......... .......T.. .......... .......... .......... ....A..... .......... .T........
EMA4  .A........ .......... .......... .......T.. .......... .......... .......... ....A..... .......... .T........
DDU   .A........ .......... .....A.... ....T..T.. .......C.G .....T..C. .......... ...CA..... ..C..... .T.....G..
DBI   .A.....C.. ......C..T .....A.... .A..C..T.. A.....AC.. .....T..C. .G..T..... ....ACT..C .......... ..........
EGR   .A.....C.. ......C..T .....A.... .A..C..... G......C.A .......... .......... T..TAC...C ..C..G..... .......T..
BTA   .A........ .......... .....A.... .A..C..... C........A ........C. .A........ ....A.TT.. ..C..... .......C..

          510        520        530        540        550        560        570        580        590        600
MPR1  CTTTTCGGTA GATAAAGCAA CCTTAAATCG ATTCTTCGCC CTCCATTTTA TTCTTCCATT TACTATAATT GCACTAGCAG GAGTACACCT AACCTTCCTT
LAF1  ......A... .......... .......... ...T...... ........C. .......... .......... .......... .......... .....T...
LAF2  ......A... .......... .......... ...T...... ........C. .......... .......... .......... .......... .....T...
LAF3  ......A... .......... .......... ...T...... ........C. .......... .......... .......... .......... .....T...
LAF4  ......A... .......... .......... ...T...... ........C. .......... .......... .......... .......... .....T...
EMA1  .......... .......... .......C.. .......... T.......C. .C........ .......G.. .......... ....G..... .....T...
EMA2  .......... .......... .......C.. .......... T.......C. .C........ .......G.. .......... ....G..... .....T...
EMA3  .......... .......... .......C.. .......... T.......C. .C........ .......G.. .......... ....G..... .....T...
EMA4  .......... .......... .......C.. .......... T.......C. .C........ .......G.. .......... ....G..... .....T...
DDU   A..C..A... ..C.....C. ..C.C.CC.. .......... ..A..C..C. .C..A..C.. C.TCG...CC ..C....T.A T...C...T. .CTA.....C
DBI   G..C..C..C ..C.....C. .AC.T.CA.. .........T T....C.... ....C..C.. ..TC..CTCA ..C....C.A TCACC..... .TTA.....A
EGR   A..C..A... ..C.....C. ..C.T.CC.. ...T..T..T T.....C..C .C..A..C.. C.TC..C.CA ..C..G.T.. TC.....TT. .CTA..T...
BTA   A..C..A... ..C.....C. ..C.T.CC.. .........T T.....C... .C........ ..TC..C..A ...A.T.CCA T...C..... .CTA.....C
```

Appendix 3. Complete sequences of cytochrome *b* gene used in this study. Sequences of DDU (*Dugong dugong*), DBI (*Diceros bicornis*), EGR (*Equus grevyi*), and BTA (*Bos taurus*) were extracted from GenBank. The three-base insertion between positions 765 and 766 (corresponding to the nucleotide numbers 15,721 and 15,722 in human system) (Irwin et al. 1991) was not found in this study. The cytochrome *b* sequences of elephant encode 378 amino acids and terminate with the 379th TAA stop codon.

```
           610        620        630        640        650        660        670        680        690        700
MPR1  CACGAAACAG GCTCAAATAA CCCACTAGGC CTCACTTCAG ACTCAGACAA AATCCCCTTT CACCCGTACT ATACCATCAA AGACTTCCTA GGACTACTTA
LAF1  .......... .......C.. T.....G... .......... .......... .....A.... .......T.. G......... ...T......
LAF2  .......... .......C.. ......G... ....T..... .......... .......... .......T.. .......... ...T......
LAF3  .......... .......C.. T.....G... .......... .......... .....A.... .......T.. G......... ...T......
LAF4  .......... .......C.. ......G... ....T..... .......... .......... .......T.. .......... ...T......
EMA1  .......... .T.....C.. .........T .......... .......... ...T...... .......... ....T..... .......... ..........
EMA2  .......... .......C.. .........T .......... .......... ...T...... .......... ....T..... .......... ..........
EMA3  .......... .......C.. .........T .......... .......... ...T...... .......... ....T..... .......... ..G.......
EMA4  .......... .......C.. .........T .......... .......... ...T...... .......... ....T..... .......... ..G.......
DDU   .......... ....C..C.. ...CACG..A ..G.TC..C. .......... ......A..C .....A..T. ..T.AG.... ....C..... ..C...T.CC
DBI   .......... .A..C..C.. ....TC...A A..C.A..CA .TAT...... ...T..A..C .....A.... .C..A..... ...TA..... ...A.C..AC
EGR   .......... .A..T..C.. ...CTC...A A..C.A..C. .TATG..... ......A..C .....A..T. ....A..T.. ....A..... .....C..CC
BTA   .......... ....C..C.. ....AC...A A.TT.C.... ..GT...... ......A..C .....C.... .......T.. G...A..T.. ..GGCC..CT

           710        720        730        740        750        760        770        780        790        800
MPR1  TCCTAATCCT ATTCCTTCTA CTCTTAGCCC TACTATCTCC TGACATACTA GGAGACCCCG ACAACTACAT ACCAGCTGAT CCACTTAATA CTCCCCTACA
LAF1  .......TT. .C.T...... ...C...... .......... .......... ........T. .......... ......C..C .....A.... ..........
LAF2  .......TT. .C.T...... ...C...... ....G..... .......... .......... ........T. .......... ......C..C .....A.... ..........
LAF3  .......TT. .C.T...... ...C...... .......... .......... ........T. .......... ......C..C .....A.... ..........
LAF4  .......TT. .C.T...... ...C...... ....G..... .......... .......... ........T. .......... ......C..C .....A.... ..........
EMA1  .......TT. .C........ .......... .......... .......... ........T. .......... .......... .....A.... ..........
EMA2  .......TT. .C........ .......... .......... .......... ........T. .......... .......... .....A.... ..........
EMA3  .......TT. .C........ .......... .......... A......... ........T. .......... .......... .....A.... ..........
EMA4  .......TT. .C........ .......... .......... A......... ........T. .......... .......... .....A.... .T..C.TA..
DDU   ....C..T.. .G..T.A..C ..AC..A... .GT.C..C.. G........G ........A. .........C ......CA.C .....A..C. .C..T.CC..
DBI   .......... .ACA..A..C AC.C...T.. ..T.C..A.. CC..CAT... ........T. .........C C..C..CACC ..T..C.... .C..T.C...
EGR   ....G...T. GC....A... ACTC...TAT ..T.C..C.. C...C.C... ........A. .........C C......A.C ..T..C.GC. ......CT..
BTA   .A.....T.. .GCT..AA.. ..AC...TA. ..T.CG.A.. C...C.C..C ........A. .T.......C C.....CA.. .....C..C. .A....CT..

           810        820        830        840        850        860        870        880        890        900
MPR1  CATCAAACCA GAGTGATATT TTCTCTTTGC TTACGCCATC CTACGATCTG TACCAAACAA ACTAGGAGGC GTCCTAGCCC TACTCCTATC AATCCTAATC
LAF1  T.....G... .......... .......... .......... .......... .......... .......... .......... .......... ...T......
LAF2  T.....G... .......... .......... .......... .......... .......... .......... .......... .......... ...T......
LAF3  T.....G... .......... .......... .......... .......... .......... .......... .......... .......... ...T......
LAF4  T.....G... .......... .......... .......... .......... .......... .......... .......... .......... ...T......
EMA1  .......... ........C. .C..T..... .........T .......... .......... .......... .......... ..T.T..... ......G..T
EMA2  T......... ........C. .C..T..... .........T .......... .......... .......... .......... ..T..... ...T..G..T
EMA3  .......... ........C. .C..T..... .........T .......... .......... .......... .......... ..T.T..... ......G..T
EMA4  .......... ........C. .C..T..... .........T .......... .......... .......... .......... ..T.T..... ......G..T
DDU   ...T...... ..A.....C. ....A..CCG A.....T... ..C......A .C..T..T.. ......C.... ..GT...... .CG.A..C.. C.........
DBI   T......... .......... .CT.A..... C..T..A... ........C. .C..T..... ......C..A ..A..... ..GCA..T.. C.........
EGR   T..T.....G ..A..G..C. .C..G..... C.......... .....C..CA .T..C..... ......C... ..AT...... ..A....C.. C.....G...
BTA   .........C ........C. .CT.A..... A.....A... T.......AA .C..C..... .........A ..A..... ..GC.T.C.. T........T

           910        920        930        940        950        960        970        980        990       1000
MPR1  CTAGGAATTA TACCACTTCT ACATACATCT AAACACCGAA GTATGATACT TCGACCTCTT AGCCAAGTCC TATTCTGAAC TCTAGCAACA GATCTACTAA
LAF1  ......T.A. .......... C........C ..G..... .C..A..... C......... .......... .......... ....A...T. ...T......
LAF2  ......T.A. .......... C........C ..G..... .C..A..... C......... .......... .......... ....A...T. ...T......
LAF3  ......T.A. .......... C........C ..G..... .C..A..... C......... .......... .......... ....A...T. ...T......
LAF4  ......T.A. .......... C........C ..G..... .C..A..... C......... .......... .......... ....A...T. ...T......
EMA1  T.....T.A. .......... C........C ..G..... ....A..... C........C .....G.... .......... ....A...T. ...T......
EMA2  T.....T.A. .....T.... C........C ..G..... ....A..... C........C .....G.... .......... ....A...T. ...T......
EMA3  T.....T.A. .....T.... C........C ..G..... ....A..... C........C .....G.... .......... ....A...T. ...T......
EMA4  T.....T.A. .....T.... C........C ..G..... ....A..... C........C .....G.... .......... ....A...T. ...T......
DDU   ....CGC.CC .C.....C.. C..C.....C .....A.... .CC.ATC.T. C........A ......TG.. .C..T....T ...G.T.G.C ..C..GA.C.
DBI   ....CTC... .C..CA.... C..C.....C .....A.... .C..A.T.T. C.....C..A ......TGTA .G......CT A....T.G.C ..CT....C.
EGR   ....C.C.C. .C..CACC.. C..C.T...A .....A.... .C..A.T.T. C..G.....C ......TG.G ........CT CT...TGG.. ..CT....G.
BTA   ..T.CTC.A. .C..C..A.. ...C..C..C .....A.... .C..A.T.T. C.....A..C ......TG.. ........G. C....T.G.. ..C.....G.

          1010       1020       1030       1040       1050       1060       1070       1080       1090       1100
MPR1  TACTTACATG AATTGGCAGT CAACCAGTAG AATATCCCTA CATCATTATC GGCCAAAATAG CCTCAATTCT ATATTTCTCC ATTATCCTAG CTTTCCTGCC
MPR2  .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
MPR6  .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
LAF1  C......... .........C .......... .......T.. .......... ..T....... .......... .......... .......... .C........
LAF2  C......... .........C .......... .......T.. .......... ..T....... .......... .......... .......... ..........
LAF3  C......... .........C .......... .......T.. .......... ..T....... .......... .......... .......... .C........
LAF4  C......... .........C .......... .......T.. .........T ..T....... .......... .......... .......... .C........
EMA1  C......... .......... .......... ..C.C..... .......... .......... .......... ...C...... .....T.... .......A..
EMA2  C......... .......... .......... ....C..... ..C....... .......... .......... ...C...... .....T.... .......A..
EMA3  C......... .......... .......... ..C.C..... .......... .......... .......... ...C...... .....T.... .......A..
EMA4  C......... .......... .......... ..C.C..... .......... .......... .......... ...C...... .....T.... .......A..
DDU   C...C..... ...C...G.C .......... ..C..... .......... ......C... ......C... .......C.. G..C...... ..C.....CA TC..TA....
DBI   C......... ...C..AG.A ........C. .GC.C..A.T .........T ......C... .A.....C.. .......... C.A..... TAC.TA.A..
EGR   C...A..... ...C...G.A ..G.....G. ..C.C..A.. CG.A...... ......C.G. .......C.. C..C...... C.A..T..CA T...TA.A..
BTA   C...C..... ......AG.A ........C. ..C.C..A.. T....CC... ..A...C... .A..TG.C.. ...C..TCT. C.C....... TGC.AA.A..

          1110       1120       1130       1140
MPR1  AATTGCAGGA ATGATCGAAA ACTACCTCAT TAAGTAA---    (1137 bases)
MPR2  .......... .......... .......... .......---    (1137 bases)
MPR6  .......... .......... .......... .......---    (1137 bases)
LAF1  .......... G.A....... .......... .......---    (1137 bases)
LAF2  .......... G.A....... .......... .......---    (1137 bases)
LAF3  .......... G.A....... .......... .......---    (1137 bases)
LAF4  .......... G.A....... .......... .......---    (1137 bases)
EMA1  ...C...... ..A....... .......... C......---    (1137 bases)
EMA2  .......... ..A....... .......... C......---    (1137 bases)
EMA3  ...C...... ..A....... .......... C......---    (1137 bases)
EMA4  ...C...... ..A....... .......... C......---    (1137 bases)
DDU   C..C.....C C.A..T.... .TC....AC. ...A.G.AGG    (1140 bases)
DBI   CC.......C ..T....... ..A....TC. G..A.G.AGA    (1140 bases)
EGR   .C.C...A.C .CC....... ..A.T...C. A..A.G.AGA    (1140 bases)
BTA   ..CG..C..C .CA....... ..A.AT.AC. A..A.G.AGA    (1140 bases)
```

Appendix 3. Continued.



Appendix 4. Complete sequences of 12S rRNA gene used in this study. Sequences of LAF (*Loxodonta africana*), EMA (*Elephas maximus*), DDU (*Dugong dugong*), TMA (*Trichechus manatus*), PCA (*Procavia capensis*), DDO (*Dendrohyrax dorsalis*), CSI (*Ceratotherium simum*), EGR (*Equus grevyi*), and BTA (*Bos taurus*) were extracted from GenBank.

```
          610        620        630        640        650        660        670        680        690        700
MPR1  TTACCGTCAC TTGCTAATTC AGTCCATATA CCACCATCTT CAGCAAACCC C-TATAGGGC ACAAAAGTGA GCTTAATCAT AACCCATGAA AAAGTTAGGC
LAF   .......... .......... .......... .......... .......... .-........ .......... .......... .......... ..........
LAF1  .......... .......... .......... .......... .......... .-........ .......... .......... .......... ..........
LAF2  .......... .......... .......... .......... .......... .-........ .......... .......... .......... ..........
EMA   .......... .......... .......... .......... .......... --.......- .......... .......... .......... ..........
EMA1  .......... .......... .......... .......... .......... .-........ .......... .......... .......... ..........
EMA2  .......... .......... .......... .......... .......... .-........ .......... .......... .......... ..........
DDU   .C...CC... .......... ..C.T..... ..G..... ....C..... T-..A.A..T G.......A. ...C...G.. -.G....A.. .........T
TMA   .C...AC... .A..... ..C.T..... ..G..... ....C..... T-..G.A..T C.......... ...CG..... -.A....A.. .........T
PCA   .C...AC..T ....CTGCC. ....T..... ..G..... .......... T-..A.A..A .GT.....A. ..AC..CA.. -.T....A.. .........T
DDO   .C...AC..T ....CT.C.. ....T..... ..G..... .......... T-..A.A..A GGC.....A. ..AC...A.. -GCT...A.. .........T
CSI   CC...AA.C. .......... ..C.T..... ..G..... .......... T-A.A.A..A ..T.....A. ..AC..GT.. ..AA...A.. ..C......T
EGR   CC...A..C. ........C. ..C.T..... ..G..... .......... TAA.C.A... ..CG....A. ..AC..... CCAA..... ..C......T
BTA   .C...AATT. ........A. ....T..... ..G..... .......... TA-.A.A..A .A......A. ..G....T.. G.TA...A.. ..C......T

          710        720        730        740        750        760        770        780        790        800
MPR1  CGAGGTGTCG CCTACGTGAC GGTCAAAGAT GGGCTACATT TTCTATT--- ATAG--AATA G-----ACAA ACGGATACCA CTCTGAAATG GGTGGTTGAA
LAF   .......... .......... .......... .......... .......--- ......-.C. .-----.... .......T.. .......... ..........
LAF1  .......... .......... .......... .......... .......--- ......-.C. .-----.... .......T.. .......... ..........
LAF2  .......... .......... .......... .......... .......--- ......-.C. .-----.... .......T.. .......... ..........
EMA   .......... .......... .......... .......... .......--- ......-.C. .-----.T.. .......... .........A ..........
EMA1  .......... .......... .......... .......... .......--- ......-.C. .-----.... .......... .........A ..........
EMA2  .......... .......... .......... .......... .......--- ......-.C. .-----.T.. .......... .........A ..........
DDU   .A......A. ..C.T...GA ....T..... .......... ....C..--- .G..--..C. ATCA--..C. .....CGT.. ..A.....CA A...ACCA..
TMA   .A......A. ..C.T...GT ....C..... .......... ....C.C--- .G..--..C. ATCAA-..T. .....CGTT. T.A......A AA..AC.A..
PCA   .A......A. .TA.TA..GT ..A.TT.... .......... C.T..CAC-- ...--A.GC. ------.T-- ...A..GTT. .CA......T A..AAC....
DDO   .A......A. .TA.TA..GT ..C.C..... .......... C.T..CAC-- .A---A.GC. -------.T-- ...A..GTT. ..A.....CT AA.AAC....
CSI   .A......A. .T..T.G..T ..AG.G.A.. .......... .......TTA .G.ACA.CA. -TTAGCC... ...A.GGTTT T.A.....CC AAAAAC.A..
EGR   .A......A. .TC.T.G..T ..AG.G.A.. .......... .....C.CTA .G.ACA.GA. CTTAACC... ...A.AGT.T ..A......T ..A.ACC...
BTA   .A......AA ....T.AA.T ..GA.G.A.. .......... C....CACCA .G..-A.TC. ---AGC---- ...A.AGTT. T.A.....CC AA.AACCA..

          810        820        830        840        850        860        870        880        890        900
MPR1  GGCGGATTTA GTAGTAAACT AAGAATAGAG AGCTTAATTG AAC-AAGGCC ATGAAGCGCG TACACACCGC CCGTCACTCT CCTCAAGTAC CTCCA--CAT
LAF   .......... .......... .......... .......... ....-.....T .......... .......... .......... .........T .C...--...
LAF1  .......... .......... .......... .......... ....-.....T .......... .......... .......... .........T .C...--...
LAF2  .......... .......... .......... .......... ....-..... .......... .......... .......... .........T .C...--...
EMA   .......... .......... .......... .......... ....-..... .......... .......... .......... .......... T...--...
EMA1  .......... .......... .......... .......... ....-..... .......... .......... .......... .......... T...--...
EMA2  .......... .......... .......... .......... ....-..... .......... .......... .......... .......... T...--...
DDU   ..A....... .C......T. .......... .......... ....-T..... .......A.. C......... .......C.. .........T TCAT.AC...
TMA   ..A....... ........T. ....TC.... .......... ..T-T..... .......A.. C......... .......C.. ......A... .CAT.ATT..
PCA   ..A....... .C......T. .......... .......... ..T-C..... .......A.. C......... .......C.. .......... ..AA------.
DDO   ..A....... .C......T. .......... .......... ..TT...... .......A.. C......... .......C.. .......... .CAA-----C
CSI   ..A....... .C......T. .......... .......... ...CC-.... ..A....A.. C......... .......C.. ...T..A... .C.G.GC.TA
EGR   ..A....... ........T. .......... .......... ..T-..... .......... C......... .......C.. ...T..A..T .A.A.AT.-A
BTA   ..A....... .C......... .......... T.....G... ..TT-..... .......A.. C......... .......C.. ......A..G A.T..GT--G

          910        920        930        940        950        960        970        980        990        998
MPR1  CAAAC--AAT CAT-AT-TAC AGATTT---- AAACA-A-AT ACAAGAGGAG ACAAGTCGTA ACAAGGTAAG CGTACTGGAA AGTGTGCTTG GGTAA-CT (961 bases)
LAF   .....--... ...-..-... ......---- ......-.-. .......... .......... .......... .......... ....C..... .A...-.. (961 bases)
LAF1  .....--... ...-..-... ......---- ......-.-. .......... .......... .......... .......... ....C..... .A...-.. (961 bases)
LAF2  .....--... ...-..-... ......---- ......-.-. .......... .......... .......... .......... .......... .A...-.. (961 bases)
EMA   .T...--... ...-..-... ......---- ......-.-. .......... .......... .......... .......... .......... ..A..-.. (958 bases)
EMA1  .T...--... ...-..-... ......---- ......-.-. .......... .......... .......... .......... .......... .A...-.. (961 bases)
EMA2  .T...--... ...-..-... ......---- ......-.-. .......... .......... .......... .......... .......... ..A..-.. (961 bases)
DDU   TC-------- ..CC..A..T -...CC---- T..T.-GC.. .TG..... .......... .......... .A........ .......... .A.T.-.. (957 bases)
TMA   .T-------- T.CC.CAA.. -A....---- T..T.-.-G. .TG..... .......... .......... .A........ .......... .A.C.-.. (958 bases)
PCA   TCT..TC..- T.C-..AA.. .-CA------ T..-.-.-.. T..... .......... .......... .A........ .......... .AA..A.. (951 bases)
DDO   TT...C-..- T.C-C.AA.. .-CA------ C.CT.-.-.C .T..... .......... .......... .A........ .......... .AA..-T. (945 bases)
CSI   .-...TC..- ...-..TA.- --C.C---CG C.TT.C.CG. .T..A..... .......... .......... .......... ....C..... .AC..-TC (973 bases)
EGR   .C.GTAT..- ..C-.-AA.. --CG.----G .CC..A.C.. .TG.A..... .......... .......... TA...C.... G....A.... .A...-.C (975 bases)
BTA   .-=.TCT..- .CCT.----- --T..AAACG C.CT.GCT.C .TG..... .......... .......... .A........ .......... .A...A-. (955 bases)
```

Appendix 4. Continued.