U. Joger [1], G. Garrido [1], J. Hauf [2], A. Tikhonov [3] & S. Vartanyan [4]
[1] Hessisches Landesmuseum, Darmstadt
[2] Dep. of Biology, Technical University, Darmstadt
[3] Zoological Institute, Russian Academy of Sciences, St.Petersburg
[4] State National Reserve 'Wrangel Island', Ushakovskoye

Genetic investigations on mammoth (*Mammuthus primigenius*)

Joger, U., Garrido, G., Hauf, J., Tikhonov, A. & Vartanyan, S., 2003 - Genetic investigations on mammoth (*Mammuthus primigenius*) - In: Reumer, J.W.F., de Vos, J. & Mol, D. (eds.) - ADVANCES IN MAMMOTH RESEARCH (Proceedings of the Second International Mammoth Conference, Rotterdam, May 16-20 1999) - DEINSEA 9: 205-219 [ISSN 0923-9308] Published 24 May 2003

Previous authors have disagreed about the interpretation of molecular data concerning the phylogenetic affiliation of *Mammuthus* with respect to *Elephas* and *Loxodonta*. We compare sequence data of the mitochondrial cytochrome b gene, the gene studied by most authors. Hagelberg et al. (1994) as well as Hauf *et al*. (1995) found a closer affinity of *Mammuthus* to *Loxodonta*, whereas Yang *et al*. (1996) and Ozawa *et al*. (1997) found a common branch of *Mammuthus* with *Elephas*. This was rejected by Noro et al. (1998) whose data supported again a *Loxodonta-Mammuthus* association, whereas Hauf *et al.*(1999) gave additional support to an *Elephas-Mammuthus* clade. However Derenko *et al.* (1997) as well as Hauf *et al.* (1999/2000) could not find any evidence for either of these alternatives, while analyses by Barriel et al. (1999) and Thomas *et al.* (2000) support an association of *Mammuthus* with *Loxodonta*, but not with a satisfying certainty. In this contribution, we present new partial cytochrome b sequences for seven Wrangel island mammoths, but do not claim to present the definite solution. When all data now available for cytochrome b are taken together, the possibly synapomorphic bases counted for *Elephas* + *Mammuthus* are only slightly more in number than those countable for *Loxodonta* + *Mammuthus*, whereas less support is found for an *Elephas-Loxodonta* clade. A remarkable observation is that different portions of the same gene may indicate different phyletic affinities. We discuss the possible reasons for differing results and point out a way to solve the dilemma. Shortcomings of previous papers include: (1) Only one or two mammoth individuals were used, which sometimes even differed in their base sequence, yet intraspecific variability was largely disregarded; (2) The outgroup chosen was too distant (Mastodon or Dugong should be taken); (3) Sequences were too short (even 1000 bases do not provide sufficient phylogenetic signal); (4) Tree reconstructions were based on single-base transitions, which are probably switching easily back and forth between apomorphic and plesiomorphic states, hiding phylogenetic signal; (5) Computer methods were used uncritically. More data are needed to clarify the phylogenetic affiliation of *Mammuthus*.

Correspondence: Dr Ulrich Joger, Hessisches Landesmuseum, Friedensplatz 1, D-64283 Darmstadt, Germany, e-mail: u.joger@hlmd.de

Key words: mammoth, cytochrome b, phylogenetic analysis

INTRODUCTION

Since the invent of PCR (polymerase chain reaction) in the mid-eighties, it has become possible to amplify single genes from fossil bones in order to gain sufficient quantities for sequencing (Thomas & Pääbo 1993). Woolly mammoth (*Mammuthus primigenius*) has been one of the prime targets for sequencing fossil genes, for obvious reasons: (1) Well preserved carcasses from permafrost soils enhance the probability of preservation of long DNA fragments with intact base sequences; (2) These carcasses are rather recent (less than 100.000 y.b.p.), therefore there is a good chance that natural decay of nucleic acids has not yet proceeded too far to read sequences, whereas after millions of years DNA is unlikely to have survived unchanged (Lindahl 1993a, b); (3) DNA sequence information from living relatives of the same family Elephantidae (genera *Elephas* and *Loxodonta*) is available in order to design specific PCR primers. Without such specific primers, so-called 'universal' primers have to be used. As these bind to DNA from many different organisms, there is a risk of obtaining sequences from contaminating sources such as bacteria, fungi or human cells instead of the few fossil gene fragments that are still long enough to be bound between two primers. After such kind of frustrating experiences, we designed an elephant-specific primer to detect mammoth DNA (Hauf *et al.* 1995) .

The question which of the two, *Elephas* or *Loxodonta*, is closer to *Mammuthus* phylogenetically, or whether *Mammuthus* branched off from the common stem first, is a simple three-taxon problem with four possible solutions (Fig. 1). However, despite a number of efforts, this problem has presented considerable resistance to being solved. All three genera originate in Africa from the Pliocene *Primelephas* (Maglio 1973), but only *Loxodonta* did not migrate to Eurasia. Morphologically, the molar folding structure has been the principal character to monitor evolutionary progress in Elephantids. As *Loxodonta* retains the more primitive molar state, *Elephas* and *Mammuthus* share a derived molar condition which is often used to

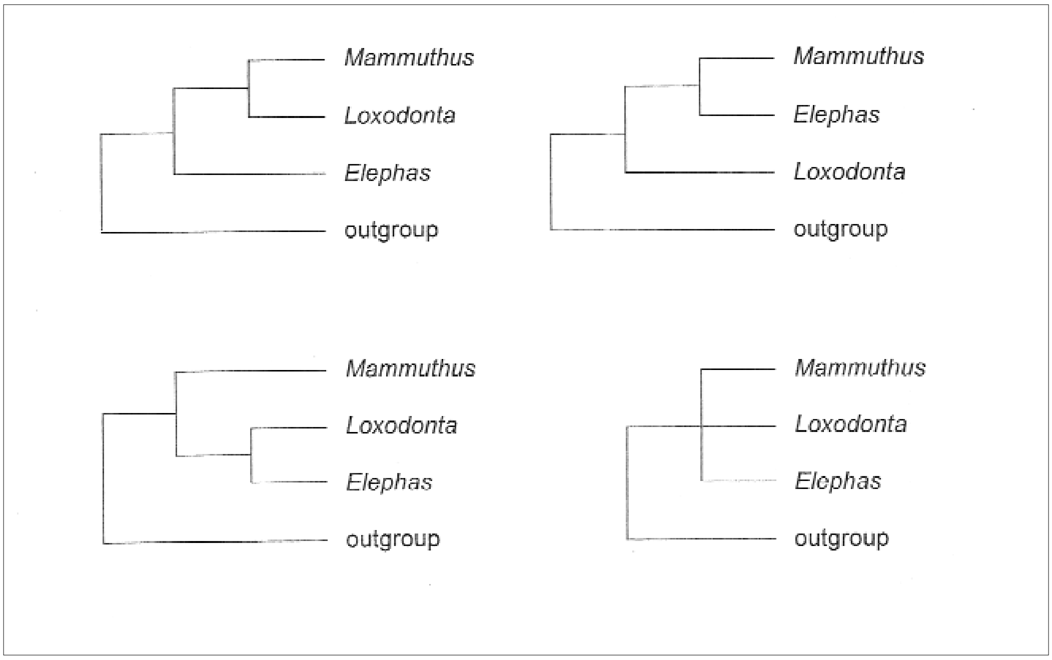


Figure 1 The four possible alternatives for a phylogenetic arrangement of the elephantid genera.

justify the assumption that they are closely related (Shoshani *et al.* 1985). However, a refined molar structure is a necessary adaptation to a higher proportion of grasses in the diet, and thus may occur independently in non-related species (as in the woolly rhinoceros *Coelodonta*). In other features, e.g. the shape of the tip of the trunk, *Mammuthus* is more similar to *Loxodonta* (Vereshchagin & Tikhonov 1990).

The first molecular studies on recent and fossil proboscidians, before the availability of PCR, were based on quantitative immunological assays of the protein serum albumin (Lowenstein 1985, 1988). However, pairwise immunological distances between the three genera were not significantly different from each other. Subsequent authors sequenced mitochondrial gene fragments. Höss *et al.* (1994) sequenced 93 base pairs (bp) of the 16s RNA gene of four mammoth carcasses and reported considerable variation between these. They did not attempt a phylogenetic placement of *Mammuthus*. Other authors sequenced fragments of the cytochrome b gene. Hagelberg *et al.* (1994) as well as Hauf *et al.* (1995) found a slightly closer affinity of *Mammuthus* to *Loxodonta*, whereas Yang *et al.* (1996) claimed to have detected a common branch of *Mammuthus* with *Elephas*. Derenko *et al.* (1997) added another partial mammoth cytochrome b sequence, largely overlapping with the four mammoths done by Hagelberg *et al.* (1994) and by Yang *et al.* (1996), but his attempt to resolve the phylogeny with the five mammoth individuals together failed. Extending the cytochrome b sequence to about 1,000 base pairs, Ozawa *et al.* (1997) reported a monophyletic *Elephas*-*Mammuthus* group. This was rejected by Noro *et al.* (1998) whose data supported a *Loxodonta*-*Mammuthus* clade, based on both cytochrome b and 12s RNA, whereas Hauf *et al.* (1999), again using cytochrome b, gave additional support to the *Elephas*-*Mammuthus* alternative. A re-analysis of eight previously published mammoth cytochrome b sequences of unequal length, in comparison with a larger number of *Elephas* and *Loxodonta* sequences (Barriel *et al.* 1999) led to a weakly supported association of *Mammuthus* with *Loxodonta*. Recently, we re-analyzed the 1,005 base pair sequence of Ozawa *et al.* (1997), but using our own new *Loxodonta* sequence for comparison (Hauf *et al.*, 1999/2000), and found *Elephas* and *Loxodonta* as possible sister groups, with *Mammuthus* occupying a more basal position. Finally, Thomas *et al.* (2000) compared cytochrome b (567 base pairs) of five mammoth individuals with 14 Asian and eight African elephants. Their preferred solution was again a *Mammuthus*-*Loxodonta* clade, though they could not reject any of the alternatives. Table 1 gives an overview of these studies and their differing outcomes.

In this contribution, we do not claim to present the definite solution, but intend to discuss the possible reasons for differing results and point out a way to solve the dilemma.

MATERIAL AND METHODS

In addition to two mammoth carcasses from Shandrin and Machsounotchka rivers, Yakutia (Republic of Sakha, Russian Federation) dated at 28,230 yBP and 27,330 yBP (see Hauf *et al.* 1999), bones of 14 mammoth individuals from Wrangel island, Chukotka (NE Siberia) were used as sources of DNA (see Joger 1996a). These bones belong to the subspecies *Mammuthus primigenius vrangeliensis* GARRUTT, AVERIANOV & VARTANYAN, 1993, which survived well into the Holocene (Vartanyan *et al.* 1993). Radiocarbon datings were done for three individuals from Wrangel (4,250 yBP, 5,281 yBP, and 7,710 yBP). We also sequenced an additional portion of the cytochrome b gene of the American Mastodon (*Mammut americanum*) already used by Yang *et al.* (1996) (10,200 yBP). We thank J. Shoshani for a gratious gift of bone powder of mastodon.

Total DNA was extracted from the mammoth bone samples. To avoid contamination by extraneous DNA, the bone surface was removed using a hand grinder. A 8 mm Ø perforation was made into the bone to obtain about half a gram of clean powdered sample. DNA was prepared by a silica-based purification method using the GENECLEAN® Kit (BIO 101, Inc., La Jolla, USA). For the DNA amplification five or six µl of the DNA solution were used and it was performed in a reaction volume of 50 µl containing 1x PCR buffer (Qiagen), 2.5 mM MgCl2 150 µM dNTPs, 1.6 mg/ml BSA (MBI Fermentas), 0.5 mM of each primer and 1.25 units of Taq polymerase (Qiagen) in a Perkin Elmer Thermocycler (GENEAMP® PCR System 9700). The amplification conditions were: 95°C for 2 minutes, for initial denaturation, followed by 40 cycles consisting of 94°C for 10 seconds, 54°C for 10 seconds 71°C for 40 seconds, followed by 72°C for 5 minutes for a final extension. No amplification was detected by electrophoresis in extraction and PCR blanks. The primers were chosen to cover positions between 370 and 1040 of the cytochrome b gene: The first forward primer was a modification of L15144 (Ozawa *et al.* 1997); the second forward primer was L500f (Hauf *et al.* 1995). The reverse primers were L800r (Hauf *et al.* 1999) and L15755 (Ozawa *et al.* 1997). The amplification of a cytochrome b fragment was successful from 8 Wrangel island samples. From three bone samples (samples DM1, DM5a, and DM6), it was possible to amplify up to 650 bp, maybe because of the good conservation state of these samples. From six mammoth DNA samples, it was not possible to obtain any amplification.

PCR fragments were sequenced in both directions according to the chain-termination method of Sanger *et al.* (1977), using the cycle sequencing technique (Murray 1989). The sequencing reactions contained approximately 300 ng of amplified DNA as sequencing template and 5 pmol of the respective primer. To this mixture the appropriate amount of BIG DYE® Terminator Cycle Sequencing Ready Reaction Sequencing Mix (PE Applied Biosystems, Weiterstadt, Germany) was added, following the manufacturer's instructions. The cycling conditions were: the denaturation step at 96°C for 10 seconds, followed by the annealing step at 50°C for 5 seconds and the extension/termination step at 60°C for 4 minutes, total of 25 cycles. The sequencing samples were electrophoresed on a ABI PRISM® 377 DNA sequencer and analysed using the ABI PRISM™ Sequencing Analysis software, version 3.2 (PE Applied Biosystems, Weiterstadt, Germany). Sequences were compared with the GeneDoc programme, version 2.5.000.

The following published sequence data were used for comparison: Lox-Irwin: *Loxodonta africana* (acc. no X56285; Irwin *et al.* 1991); Lox-Noro1: *Loxodonta africana* (acc. no D84150; Noro *et al.* 1998); Lox-Noro2: *Loxodonta africana* (acc. no D84151; Noro *et al.* 1998); Lox-Noro3: *Loxodonta africana* (acc. no D84152; Noro *et al.* 1998); Lox-Yang: *Loxodonta africana* (acc. no U23741; Yang *et al.* 1996); Ele-Noro1: *Elephas maximus* (acc. no D50844; Noro *et al.* 1998); Ele-Noro2: *Elephas maximus* (acc. no D50846; Noro *et al.* 1998); Ele-Noro3: *Elephas maximus* (acc. no AB002412; Noro *et al.* 1998); Ele-Ozawa: *Elephas maximus* (acc. no D83048; Ozawa *et al.* 1997); Ele-Yang: *Elephas maximus* (acc. no U23740; Yang *et al.* 1996); Mam-Noro: *Mammuthus primigenius* (acc. no D50842; Noro *et al.* 1998); Mam-Ozawa: *Mammuthus primigenius* (acc. no D83047; Ozawa *et al.* 1997); Mam-Dere.: *Mammuthus primigenius* (acc. no U79411; Derenko *et al.* 1997); Mam-Hag1: *Mammuthus primigenius* (acc. no U79411; Hagelberg *et al.* 1994); MamHag2: *Mammuthus primigenius* (Hagelberg *et al.* 1994); Mam-Yang1: *Mammuthus primigenius* (acc. no U23739; Yang *et al.* 1996); Mam-Yang2: *Mammuthus primigenius* (acc. no U23738; Yang *et al.* 1996); Mast-Yang: *Mammut americanum* (acc. no U23737; Yang *et al.* 1996); Dug-Irwin: *Dugong dugong* (acc. no U07564; Irwin & Arnason 1994); Pro-Ozawa: *Procavia capensis* (acc. no D86909; Ozawa *et al.* 1997).

Our own sequences are coded as follows: Lox-Hauf: *Loxodonta africana* (Hauf *et al.* 1999/2000); 3-Mam-Hauf: *Mammuthus primigenius*, 3 identical individuals, one of them from Wrangel island (Hauf *et al.* 1999); 3-Mamm: *Mammuthus primigenius*, 3 identical individuals from Wrangel island (collection numbers DM18, DM19, DM21); DM 1, DM 5a, DM 6, DM 13: *Mammuthus primigenus*, four slightly differing individuals from Wrangel island; Mast-Hauf: *Mammut americanum* (same individual as Mast-Yang, but different fragment of cytochrome b).

Figure 2 shows the approximate regions covered by the partial cytochrome b sequences of *Mammuthus primigenius* mentioned above.

RESULTS AND DISCUSSION

High intraspecific variation or amplification errors?

The mammoth carcasses used for the diverse studies stem from all over Siberia, ranging from the Taimyr peninsula in the Northwest to the Magadan area and Alaska in the extreme East. Their geological age spans a considerable time range from beyond the credibility limits of the radiocarbon method (>50,000 yBP) to Holocene times (4,000 yBP from the Wrangel island mammoth, *M. primigenius vrangeliensis*). Therefore, an amount of genetic variation that exceeds the one known from recent species is not unlikely. However, mitochondrial gene variation in the recent species has not been studied prior to the mammoth studies.

Moreover, inferring genetic variation from base differences obtained from single PCR assays can be erroneous because of artifacts due to the polymerase chain reaction itself. A polymerase error can result in wrong nucleotides or in the amplification of non-orthologous genes (‘jumping PCR’). Even a nuclear copy of a mitochondrial gene may be accidentally amplified. The thermostabile DNA polymerases used in PCR differ in their accuracy of copying a giving sequence. Error frequency can vary between 1/9,000 and 1/1,800,000 per nucleotide, depending on the type of polymerase used and various reaction parameters (Eckert & Kunkel 1991). In the worst case, this can amount to one error per 500 base pairs after 40 polymerase cycles. In assays involving fossil genes, conditions may even favour errors, due to chemical denaturation of nucleotid bases and low number of intact target fragments, making more cycles necessary. Even with recent material, errors may occur: the cytochrome b sequence of *Loxodonta africana* deposited in EMBL under accession no. X56285 (Irwin *et al.* 1991) and used for comparison with *Mammuthus* by Hagelberg *et al.* (1984), Hauf *et al.* (1995, 1999), Yang *et al.* (1996), and Ozawa *et al.* (1997) proved to differ from other *Loxodonta* sequences in several features, including a supposed insertion in positions 965, 970 and 977 (Apendix 2), which

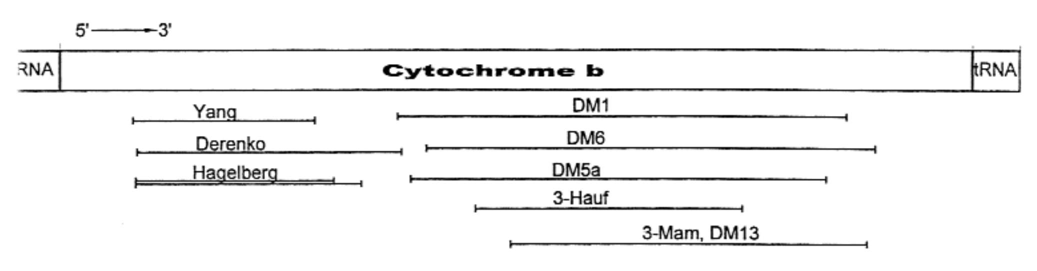


Figure 2 Overview of the mitochondrial cytochrome b molecule showing the partial sequences of *Mammuthus primigenius* done by different authors. Our own 10 sequences are shown on the right (3') part, whereas previous authors concentrated on the left (5') part. Ozawa *et al.* (1997) and Noro *et al.* (1998) sequenced more or less the whole cytochrome b (not shown).

were most likely erroneous (Hauf *et al.* 1999/ 2000). Of the 17 possible synapomorphic bases shared by *Mammuthus* and *Elephas* in our previous analysis (Hauf *et al.* 1999), no less than 10 have to be removed because we relied on the *Loxodonta* sequence of Irwin *et al.* (1991). New *Loxodonta* sequences show no difference from *Elephas* and *Mammuthus* at these 10 sites. Moreover, our own *Elephas* sequence differs from other *Elephas* in three of the supposed synapomorphic sites, thus reducing the number of sites supporting an *Elephas-Mammuthus* clade to 4 (hence not more than the number supporting a *Loxodonta-Mammuthus* clade).

For these reasons, reliable results cannot be based on single animals, but variation must be checked with several unrelated indivuals (at least three) of each elephantid species. This prerequisite was not met by many of the published studies (Table 1). Only Höss *et al.* (1994), Hauf *et al.* (1999) and Thomas *et al.* (2000) sequenced more than two mammoth individuals, but their results were contrary: the 16s RNA fragment sequenced by Höss *et al.* as well as the cytochrome b segment sequenced by Thomas *et al.* were highly variable whereas Hauf *et al.* found identical cytochrome b sequences in three mammoth individuals, one of which was from Wrangel island.

When comparing all the data on mammoth and elephant cytochrome b that are now available, areas of high heterogeneity are found in several regions of the gene, for example between bases no. 100-300 (Appendix 1). Derenko *et al.* (1997) compared five mammoth sequences that included this region and found that two of them clustered with *Elephas* whereas three clustered with *Loxodonta*. Thomas *et al.* (2000) found a diphyletic tree pattern, with the short mammoth sequences of Yang *et al.* (1997) clustering with *Elephas*, whereas longer sequences clustered with *Loxodonta*. This outcome is caused by variable positions. In position 117, the two mammoths of Yang *et al.* (1997) differ from those of other authors by a guanidine residue, as does their *Elephas* sequence. Again at site 165, their *Mammuthus* and *Elephas* sequences share the same base, cytosine. As their outgroup *Mammut americanum* is identical with *Loxodonta* at these sites, they were tempted to regard the shared states of

Table 1 Attempts to resolve mammoth phylogeny with mitochondrial sequences.

Year	**Authors**	**No of base Pairs**	**Gene**	**Individual sequences**			**Results: sister groups**
				Mam.	***Loxo.***	***Elephas***	
1994	Höss *et al.*	92	16s RNA	4	-	-	not resolved
1994	Hagelberg *et al.*	283	cytochrome b	2	1*	1	uncertain (*Mammuthus-Loxodonta?*)
1995	Hauf *et al.*	115	cytochrome b	1	1*	1	uncertain (*Mammuthus-Loxodonta?)*
1996	Yang *et al.*	228	cytochrome b	2	2	1	*Mammuthus-Elephas*
1997	Derenko *et al.*	331	cytochrome b	1(5)**	2**	2**	not resolved
1997	Ozawa *et al.*.	1005	cytochrome b	1	1*	1	*Mammuthus-Elephas*
1998	Noro *et al.*	1137	cytochrome b	1	3	3	*Mammuthus-Loxodonta*
		961	12s RNA	1	3	3	*Mammuthus Loxodonta*
1999	Hauf *et al.*	335	cytochrome b	3	1*	1	*Mammuthus-Elephas*
1999	Barriel *et al.*	varying	cytochrome b	8**	4(9)**	7(14)**	uncertain (*Mammuthus-Loxodonta?)*
2000	Hauf *et al.*	1005	cytochrome b	2**	1	1	uncertain (*Elephas-Loxodonta?*)
2000	Thomas *et al.*	545	cytochrome b	5	8	14	uncertain (*Mammuthus-Loxodonta?*)

* published sequence of Irwin *et al.* (1991) ** sequences published by previous authors reanalyzed

Mammuthus and *Elephas* as synapomorphic. However, other sequences show that *Mammuthus* (and at site 117, also *Elephas*) are variable at these particular positions. Sites 132, 144, 156, 207, 216, 219, 228 and 264 offer other examples of variability within *Mammuthus*. Site 171 and 273 vary between *Loxodonta* individuals, sites 195 and 270 vary within *Elephas*, and sites 246 and 267 vary in both *Elephas* and *Mammuthus*. If this high variability is not due to amplification or sequencing errors, these variable areas are certainly less suitable for phylogenetic purposes, and therefore the differing results of authors relying on those particular sequences (such as Hagelberg *et al.* 1994, Yang *et al.* 1996, and others) are easily explainable.

Contrary to this high variability, the cytochrome b fragment on which we concentrated our own efforts shows very little intraspecific variation. Between bases no. 560-818, our sequences from eight Holocene mammoth individuals from Wrangel island are identical to the much older one of Noro *et al.* (1998) from Taimyr Peninsula and the two of Hauf *et al.* (1999) from Yakutia. Only Ozawa *et al.* (1997) reported two differences in positions 621 and 652 (Appendix 2). Their individual was the Magadan baby ‘Dima’ of roughly the same geological age as the Taimyr mammoth. Position 621 is a third-codon position also variable in *Loxodonta*, whereas position 652 is a first codon position which is unlikely to vary. Therefore it could be erroneous. Our discussion (Hauf *et al.* 1999) about possible errors on our side that could be responsible for identical sequence readings has turned out to be over-cautious, as additional sequence data from this region of cytochrome b also show no variation within *Mammuthus primigenius*. In positions no. 819 and 822, four of our sequences are identical with the one from Noro *et al.* (1998), but another six are identical with those of Ozawa *et al.* (1997) and with our mastodon sequence. Apparently these are third-codon-sites of high individual variability. This variation is neither geographical nor spatial, i.e. not any variant is confined to *M. p. primigenius* or *M. p. vrangeliensis*, but both occur in both of them. Multiple mutations and backmutations at these sites are probable. These two base positions, identical to their homologues in *Loxodonta* (Noro *et al.* 1998) or *Elephas* (Ozawa *et al.* 1997) may explain these authors’ differing results, at least partly. Moreover, in position 822, the outgroups *Dugong* and *Procavia* differ.

Base sequences or amino acid sequences as characters?

Most researchers treat single bases in DNA sequences as if they were independent characters. This is, of course, not entirely correct because the bases of a given gene are part of one and the same functional complex and are likely to show concerted evolution. Moreover, as single bases have only four possible states, the likeliness for homoplasies (parallel mutations or back mutations) is rather high and shared derived states of two species are increasingly difficult to determine with rising total number of fixed mutations that have occurred in each clade (Joger 1996b). It is also well known that the type of base substitutions called transitions is much more frequent than the one called transversions.

If a protein coding gene is concerned, a codon, composed of three amino acids, is a functional entity that determines a single amino acid. Some authors (e.g., Ozawa *et al.* 1997) therefore prefer amino acid trees to trees derived from single base differences. However, that approach disregards the information content of the individual codons many of which differ only by synonymous (silent) mutations that may indicate the direction of evolutionary change. Codon variants were used for determining the phylogenetic position of mammoth by Hauf *et al.* (1999).

Influence of outgroup choice

The outgroup chosen for a phylogenetic analysis of a given group of organisms determines the inferred direction of evolution within the group and therefore its choice is crucial for the outcome of any phylogenetic analysis.

The ideal outgroup must be situated outside the group to be analysed, but as closely related to it as possible. If the common ancestor of outgroup and ingroups is located too far back in the past, the outgroup is likely to have lost most of its plesiomorphic character states that are needed for determining the respective apomorphic states within the ingroup. Most researchers on mammoth genes relied on gene bank sequences as outgroups. Initially the only appropriate mitochondrial sequences available were ungulates or whales. More recent authors (Ozawa *et al.* 1997, Noro *et al.* 1998) were the first that had sequences of Hyracoidea and Sirenia at their disposition, thus improving the outgroup comparison. Until now, only Yang *et al.* (1996) and Derenko *et al.* (1997) included another fossil proboscidean, the American mastodon *Mammut americanum*, as an outgroup, which makes their results more reliable. However, their mastodon sequence was much shorter (228 bp) than the sequences of more distant outgroups of other authors (Table 1). Hauf *et al.* (1999) re-analyzed the data of Hagelberg *et al.* (1994) by using Yang *et al.*'s mastodon sequence as an outgroup and found that, if data from both groups are taken together, their intraspecific variation within *Mammuthus primigenius* exceeds the possible number of shared derived bases with *Elephas* or *Loxodonta*. For this study, we sequenced an additional portion of the cytochrome b gene of *Mammut americanum* and found that several differences from outgroups that we used previously result in a different number of synapomorphic bases.

The use of computer programs for phylogeny reconstruction

Computer programs provide an impression of reliability and mathematical exactness, especially when high bootstrap probabilities are presented for a certain branching order. However, the fact that both Ozawa *et al.* (1997) and Noro *et al.* (1998) indicate high bootstrap support (90% or more) derived from a largely overlapping sequence of cytochrome b, but supporting opposing phyletic affiliations of *Mammuthus*, invalidates the bootstrap procedure as indicator for phyletic correctness and sheds some doubt on the uncritical use of computer methods. Of course, computer algorithms are necessary when a large data set containing many taxa is to be evaluated, but in a simple three-taxon comparison like the one under discussion here, computer-generated trees are not always needed. Thus we refrain from applying a computer program and use the traditional Hennigian phylogenetic method of determining shared derived states by outgroup comparison in aligned sequences.

CONCLUSIONS

From the shortcomings of previous attempts to reconstruct the phylogeny of the Elephantidae the following lessons can be learned.

(**1**) One individual is not enough to get correct sequences from fossils (and even from recent species, too). The sequences should be confirmed with at least three specimens. Individual variation, however, may require more than that minimum number of three specimens. Single bases are weak characters, especially transitions occur rather often and may switch back and forth. Phylogenetic conclusions should not be based on a small number of single base transitions but include transversions and amino acid changes.

(**2**) It is not always advisable to rely on computer methods for phylogeny reconstruction. Simple three taxon questions can be addressed by classical Hennigian phylogenetic analysis. The outgroups for phylogeny reconstruction should not be too distantly related. Of available outgroups for elephantids, mastodon would be the best choice, dugong the second best. Apparently, short sequences can be misleading. How long must a reliable sequence be?

(**3**) Some sequences are more variable than others. It is well known that evolutionary rate depends on functional constraints on the respective gene or part of gene. A sequence coding for the specific centre of an enzyme may not vary at all, whereas a non-functional sequence may show a high intraspecific variation that makes it unsuitable for comparisons between genera. Mutation saturation effects occur beyond a level of 15% base exchange. This level is reached by transitions much quicker than by transversions. While transitions may occur too often, transversions may be too rare to provide enough phylogenetic signal (as in the present data set). A reliable sequence shows intermediate variation, and the total number of phylogenetically informative sites (i.e. those which group two of three species together while the third remains in the assumed plesiomorphic state as inferred from the outgroup) must be statistically significant. The longer the sequenced portion of the genome, the more phylogenetically informative sites can be found, and the higher the probability that the information content is sufficient for a decision.

(**4**) Short sequences can be misleading: Table 2 shows that part of the cytochrome b gene indicates a *Mammuthus-Elephas* association, whereas a different part would favour a *Mammuthus-Loxodonta* clade. Dubrovo & Rautian (1999) showed that 200 to 300 base pairs of cytochrome b did not suffice to distinguish between the two alternatives *Elephas-Mammuthus* and *Loxodonta-Mammuthus* (nor did 961 bases of 12s RNA) but that the three genera appeared statistically equidistant. This is in line with the data of Derenko *et al.* (1997) and of Hauf *et al.* (1999) who showed that the phylogenetic information in the same portion of the cytochrome b gene is minimal.

(**5**) If the data in Table 2 are critically evaluated, it must be said that the total evidence is insufficient for a reliable support of any of the possible phylogenetic groupings (Fig. 1), especially taking into account that the evidence consists mainly of transitions, the least reliable mutation type.

Table 2 Possible synapomorphic base positions in 1100 bp of cytochrome b. ts = transition; tv = transversion; variable positions in brackets ( ); outgroups: *Mammut americanum*, *Dugong dugong*, *Procavia capensis*. [1] Outgroups differ, making plesiomorphic state uncertain. [2] Noro *et al.* (1998) found identity with *Loxodonta* in these positions. [3] Alignment not reliable.

Mammuthus + Elephas	***Mammuthus + Loxodonta***	***Loxodonta + Elephas***
	(12 TS)	
	45 TS	(40 TV)[1]
(75 TS)		
(171 TS)		
(264 TS)	(315 TS)	
	330 TS	
	402 TS	
	438 TS	
459 TS	(492 TS)	
507 TS	528 TS	
	552 TS	(597 TS)
627 TS		
(666 TS)		
(678 TS)		(709 TS)
724 TS	(741 TV)[1]	777 TS
780 TS		
(819 TS)[2]		
(822 TS)[2]		
942 TS		933 TS
		(985 TS)[3]
		(989 TS)[3]
		(994 TS)[3]
(1020 TS)	(1033 TS)	
(1092 TS)		

We are currently sequencing additional portions of the mitochondrial genome from several mammoth individuals of different geological age, both from the Siberian mainland and from Wrangel island, and thus hope to enlarge the data base. A first tree based on the mitochondrial d-loop region showed again a well supported *Mammuthus-Elephas* clade (Joger & Garrido 2001).

REFERENCES

Barriel, V., Thuet, E. & Tassy, P. 1999 - Molecular phylogeny of Elephantidae. Extreme divergence of the extant forest African elephant - C.R. Acad. Sci. Paris, Life Sci., 322: 447-454

Derenko, M., Malyarchuk, B. & Shields, G.F., 1997 - Mitochondrial cytochrome b sequence from a 33,000 year-old Woolly Mammoth (*Mammuthus primigenius*) - Ancient Biomolecules 1: 149-153

Dubrovo, I.A. & Rautian, G.S., 1999 - Comparative analysis of genetic diversity in mammoth - in: 2nd International Mammoth Conference, May 16-20, 1999, Natuurmuseum Rotterdam, Abstracts: 15-17

Eckert, K.A. & Kunkel, T.A., 1991 - The fidelity of DNA polymerases used in PCR - in: McPherson, M.J., Quirke, P. & Taylor, G.R. (eds.) - PCR. A practical approach. Vol. I - Oxford Univ. Press, New York: 225-244

Hagelberg, E., Thomas, M.G., Cook, C.E.Jr., Sher, A.V., Baryshnikov, G.F. & Lister, A.M., 1994 - DNA ancient mammoth bones - Nature 370, 333-334

Hauf, J., Baur, A., Chalwatzis, N., Joger, U., Zimmermann, F.K. & Lazarev, P., 1995 - Selective amplification of a mammoth mitochondrial cytochrome b fragment using an elephant-specific primer - Current Genetics 27: 486-487

Hauf, J., Joger, U., Zimmermann, F.K., Lazarev, P. & Vartanyan, S.L., 1999 - Protein and Nucleic acid sequences of Woolly Mammoth cytochrome b and the phylogenetic position of *Mammuthus* within the Elephantidae - In: Haynes, G., Klimowicz, J. & Reumer, J.W.F. (eds.) - Mammoths and the Mammoth Fauna: Proceedings of the First International Mammoth Conference, St. Petersburg, Russia - Deinsea 6: 211-217

Hauf, J., Chalwatzis, N., Joger, U. & Zimmermann, F.K., 1999/2000 - The complete mitochondrial genome sequence of the African elephant (*Loxodonta africana*) and its implication for the assessment of the systematic position of the Proboscidea - Zoology 102: 184-195

Höss, M., Pääbo, S. & Vereshchagin, N.K., 1994 - Mammoth DNA sequences - Nature 370: 333

Irwin, D.M., Kocher, T.D. & Wilson, A.C., 1991 - Evolution of the cytochrome b gene of mammals - Journal of Molecular Evolution 32: 128-144

Irwin, D.M. & Arnason, U., 1994 - Cytochrome b gene of marine mammals: phylogeny and evolution - Journal of Mammalian Evolution 2: 37-55

Joger, U., 1996a - Auf Mammutsuche auf der Wrangel-Insel - Informationen aus dem Hessischen Landesmuseum Darmstadt, 2/1996: 27-30

Joger, U., 1996b - Molekularbiologische Methoden in der phylogenetischen Rekonstruktion - Zoologische Beiträge, Berlin, N.F. 37: 77-131

Joger, U. & Garrido, G., 2001 - Phylogenetic position of *Elephas*, *Loxodonta* and *Mammuthus*, based on molecular evidence - in: The World of Elephants

Lindahl, T., 1993a - Instability and decay of the primary structure of DNA - Nature 362: 709-715

Lindahl, T., 1993b - Recovery of antediluvian DNA - Nature 365: 700

Lowenstein, J.M., 1985 - Radioimmune assay of mammoth tissue - Acta Zoologica Fennica 170: 233-235

Lowenstein, J.M., 1988 - Immunological methods for determining phylogenetic relationships - in: Runnegar, B. & Schopf, J.W. (eds.) - Molecular evolution and the fossil record - Short Courses in Paleontology 1, Paleontological Society, Knoxville, Tennessee: 12-19

Maglio, V. J., 1973 - Origin and evolution of the Elephantidae - Transactions American Philosophical Society, n.s. 63: 1-149

Murray, V., 1989 - Improved double stranded DNA sequencing using the linear polymerase chain reaction.- Nucleic Acids Research 17: 8889

Noro, N., Masuda, R., Dubrovo, I.A., Yoshida, M.C. & Kato, M., 1998 - Molecular phylogenetic inference of the woollly mammoth *Mammuthus primigenius*, based on complete sequences of mitochondrial cytochrome b and 12s ribosomal RNA – Journal of Molecular Evolution 46: 314-326

Ozawa, T., Hayashi, S. & Mikhelson, V.M., 1997 - Phylogenetic position of mammoth and Steller's sea cow within Tethytheria demonstrated by mitochondrial sequences - Journal of Molecular Evolution 44: 406-413

Sanger, F., Nicklen, S. & Coulsen, A.R., 1977 - DNA sequencing with chain terminating inhibitors - Proceedings National Academy Sciences USA 74: 5463-5467

Shoshani, J., Walz, D.A., Goodman, M., Lowenstein, J. M. & Prychodko, W., 1985 - Protein and anatomical evidence of the phyletic position of *Mammuthus primigenius* within the Elephantidae - Acta Zoological Fennica 170: 237-240

Thomas, M.G., Hagelberg, E., Jones, H.B., Yang, Z. & Lister, A.M., 2000 - Molecular and morphological evidence on the phylogeny of the Elephantidae - Proceedings Royal Society London B 267: 2493-2500

Thomas, W.K. & Pääbo, S., 1993 - DNA Sequences from old tissue remains - Methods in Enzymology 224: 406-419

Vartanyan, S.L., Garutt, V.E. & Sher, A.V., 1993 - Holocene dwarf mammoths from Wrangel Island in the Siberian Arctic - Nature 362: 337-340

Vereshchagin, N.K. & Tikhonov, A.N., 1990 - The exterior of mammoth - Irkutsk (in Russian)

Yang, H., Golenberg, E.M. & Shoshani, J., 1996 - Phylogenetic resolution within the Elephantidae using fossil DNA sequence from the American mastodon (*Mammut americanum*) as an outgroup - Proceedings National Academy Sciences USA 93: 1190-1194

Received 28 January 2000

NOTE ADDED IN PROOF:
This paper was submitted in the year 2000 and does not refer to more recent results. New and controversial mitochondrial data were published in the Proceedings of the 1st International Congress 'The World of Elephants' (G. Caravaretta, P. Gioia, M. Mussi, M.R. Palumba, eds., Rome 2001), in particular in articles by U. Joger & G. Garrido (p. 544-547), and by R. Debruyne (p. 630-632).

APPENDIX I

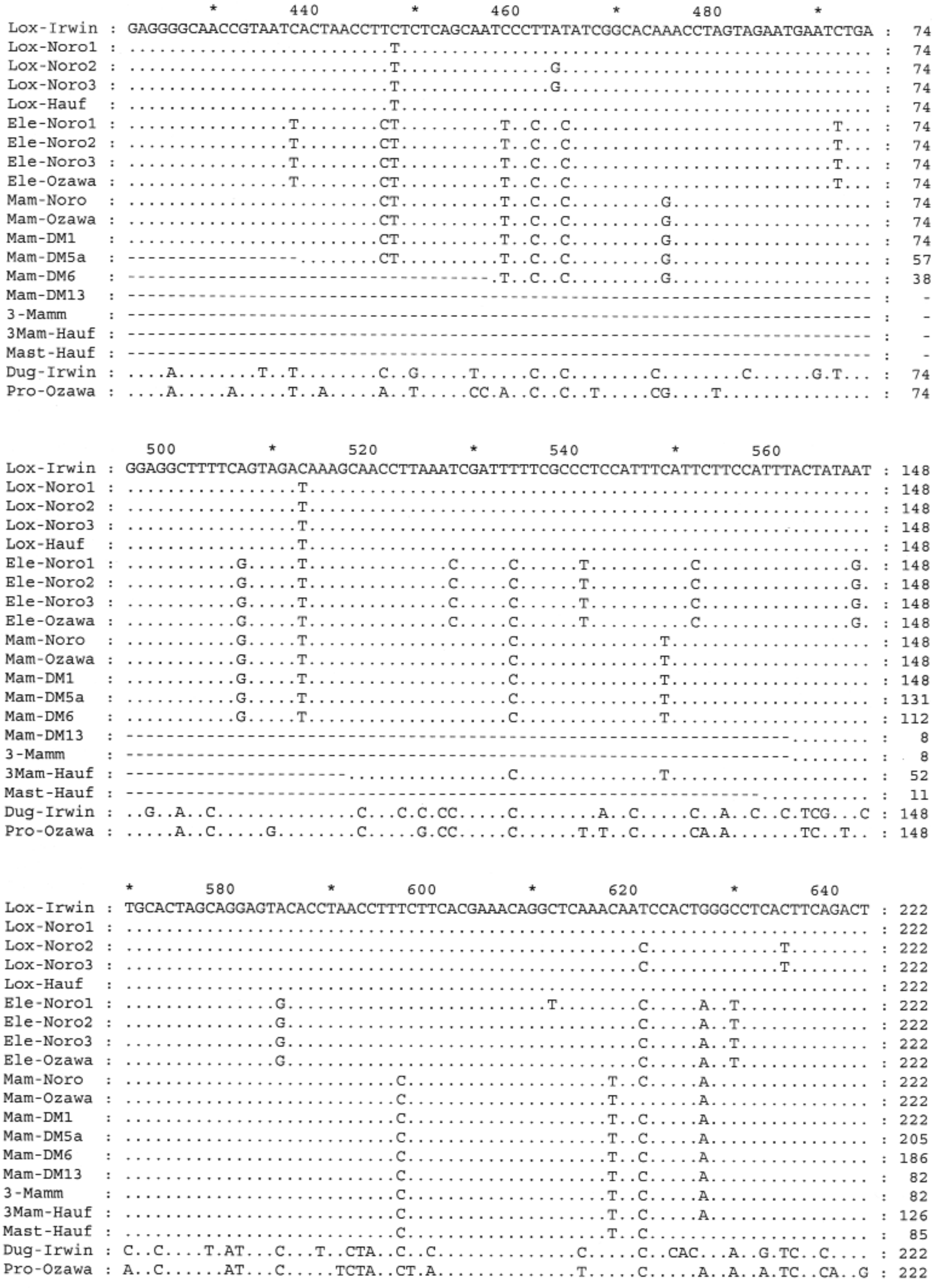
Alignment of several partial cytochrome b sequences published covering sites 90-320. See material and methods section for abbreviations.

```
            *       100         *       120         *       140         *       160         *       180         *       200
Lox-Irwin : ATGAAATTTCGGCTCACTACTAGGAGCATGCCTAATTACACAAATCCTAACAGGATTATTCCTAGCCATACATTATACACCCGACACAATAACTGCATTTTCATCTATATCCCATA : 116
Lox-Noro1 : .......................................C............................................................................ : 116
Lox-Noro2 : .......................................C.........................................T.................................. : 116
Lox-Noro3 : .......................................C..G......................................T.................................. : 116
Lox-Hauf  : .......................................C............................................................................ : 116
Lox-Yang  : --.....................................C............................................................................ : 114
Ele-Noro1 : .......................................C...................................C.....T.................................. : 116
Ele-Noro2 : ...........................G...........C...................................C.....T.................................. : 116
Ele-Noro3 : .......................................C...................................C.....T.................................. : 116
Ele-Ozawa : .......................................C...................................C.....T.................................. : 116
Ele-Yang  : --.........................G...........C...................................C.....T.......................C.......... : 114
Mam-Noro  : .......................................C..............G.....T....................T.................................. : 116
Mam-Ozawa : .......................................C..............G.....T....................T.................................. : 116
Mam-Der.  : --------...............................C..............G.....T....................T.................................. : 108
Mam-Hag.1 : ------.................................C..G.................T....................T.................................. : 110
Mam-Hag.2 : ------.................................C..............G.....T.....T..............T.................................. : 110
Mam Yang1 : --.........................G...........C...................................C.....T.................................. : 114
Mam-Yang2 : --.........................G...........C...................................C.....T.................................. : 114
Mas-Yang  : --.....................................C.....................................................C.....C................ : 114
Dug-Irwin : ......C..T.....C.....C..G........G....TT.....T..C..............G........C..C...T.A......C....C.....C..C..AG..A...... : 116
Pro-Ozawa : ......C.....A..C..C........C........C.TC...G.T........CC..........T........C...T.A..T........C.....CA....AG..A....C. : 116

                *       220         *       240         *       260         *       280         *       300         *       320
Lox-Irwin : TTTGCCGAGATGTAAACTACGGCTGAATTATTCGACAACTACACTCAAACGGAGCATCCATTTTCTTCCTCTGCCTATACACACACATTCGACGAAACATCTACTATGGCTCCTA : 231
Lox-Noro1 : .............G..................................................................................................... : 231
Lox-Noro2 : .............G.....................................................T............................................... : 231
Lox-Noro3 : .............G.....................................................T............................................... : 231
Lox-Hauf  : .............G..................................................................................................... : 231
Lox-Yang  : .............G....................................................................................................- : 228
Ele-Noro1 : .C........C..C..............................................T..................................................A..... : 231
Ele-Noro2 : .C........C..C..........................G.................T..C..T............................................A..... : 231
Ele-Noro3 : .C........C..C..............................................T..................................................A..... : 231
Ele-Ozawa : .C........C..C..............................................T..................................................A..... : 231
Ele-Yang  : .C........C..C..........................G.................T..C..T............................................A....- : 228
Mam-Noro  : .C...........C........T...................................T........................................................ : 231
Mam-Ozawa : .C...........C........T...................................T........................................................ : 231
Mam-Der.  : .C...........C........T...................................T........................................................ : 223
Mam-Hag.1 : .............C..............................................T..................------................................ : 219
Mam-Hag.2 : .C...........C........T...................................T..................------................................ : 219
Mam Yang1 : ..........C..G.....................................................C...............................................- : 228
Mam-Yang2 : .............G..........................G.................T..N...............................................A....- : 228
Mas-Yang  : .C...........C...A.T...............................................C.................T..........................C..A....- : 228
Dug-Irwin : .......G.................................T.T..T...G.T.....G.....A..A........A.....C...G.C...........CGGA.....T..C..C..A.. : 231
Pro-Ozawa : ....T.....C......C.T..A.....A..C...T.CT....TC.C..T.........C.A......C.A.....T...G.C.....C.....CGCT..T.....C..A..... : 231
```

APPENDIX 2

Alignment of our partial cytochrome b sequences with previously published sequences. See material and methods section for abbreviations.

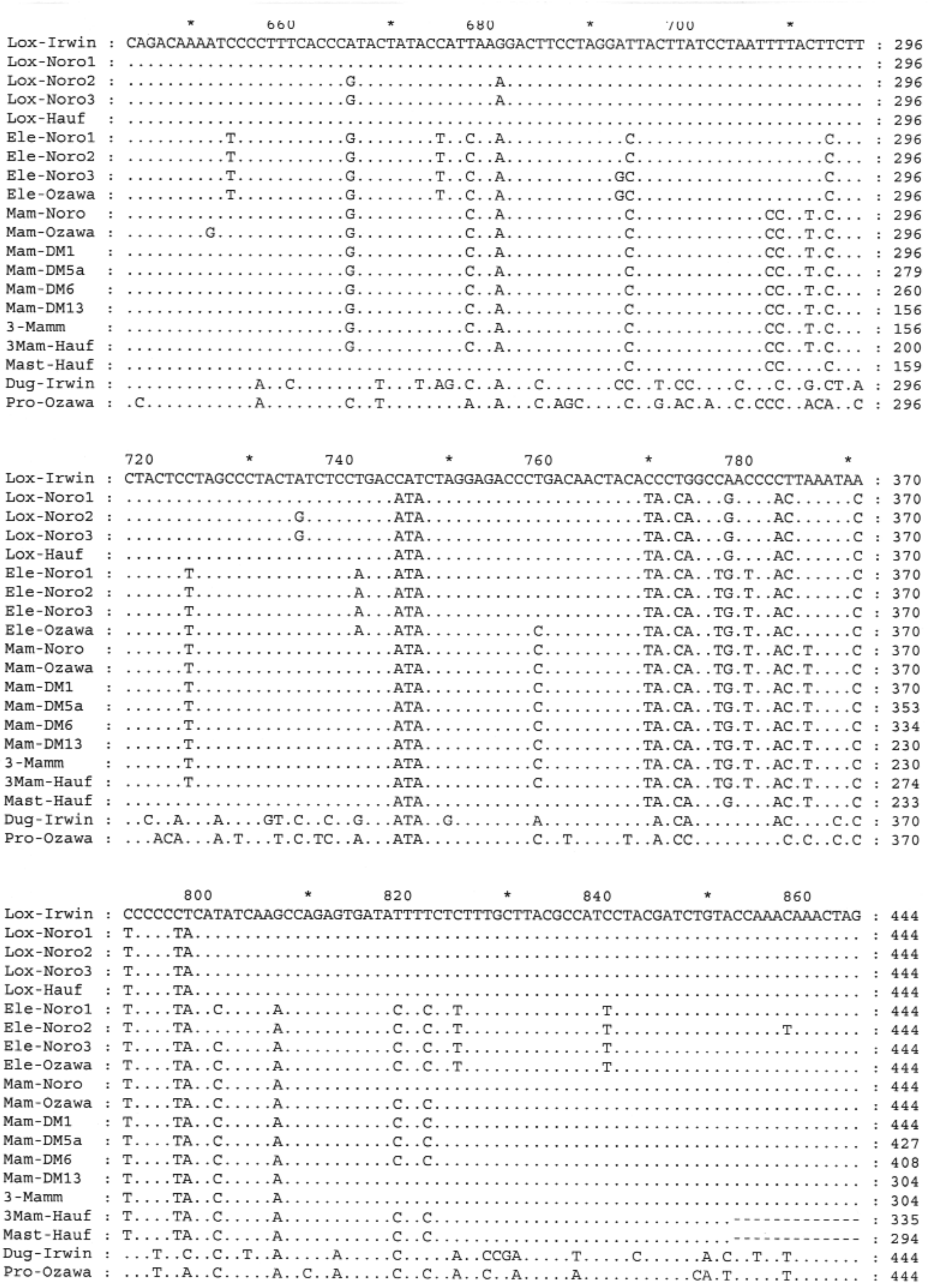


```
                    *       440         *       460         *       480         *
Lox-Irwin  : GAGGGGCAACCGTAATCACTAACCTTCTCTCAGCAATCCCTTATATCGGCACAAACCTAGTAGAATGAATCTGA :  74
Lox-Noro1  : ..........................T............................................... :  74
Lox-Noro2  : ..........................T...............G............................... :  74
Lox-Noro3  : ..........................T...............G............................... :  74
Lox-Hauf   : ..........................T............................................... :  74
Ele-Noro1  : ................T........CT..........T..C..C..........................T... :  74
Ele-Noro2  : ................T........CT..........T..C..C..........................T... :  74
Ele-Noro3  : ................T........CT..........T..C..C..........................T... :  74
Ele-Ozawa  : ................T........CT..........T..C..C..........................T... :  74
Mam-Noro   : .........................CT..........T..C..C.........G.................... :  74
Mam-Ozawa  : .........................CT..........T..C..C.........G.................... :  74
Mam-DM1    : .........................CT..........T..C..C.........G.................... :  74
Mam-DM5a   : -----------------........CT..........T..C..C.........G.................... :  57
Mam-DM6    : ------------------------------------.T..C..C.........G.................... :  38
Mam-DM13   : -------------------------------------------------------------------------- :   -
3-Mamm     : -------------------------------------------------------------------------- :   -
3Mam-Hauf  : -------------------------------------------------------------------------- :   -
Mast-Hauf  : -------------------------------------------------------------------------- :   -
Dug-Irwin  : ....A........T..T........C..G.....T.....C..C........C........C......G.T... :  74
Pro-Ozawa  : ....A.....A.....T..A.....A..T.....CC.A..C..C..T.....CG....T............... :  74

               500         *       520         *       540         *       560
Lox-Irwin  : GGAGGCTTTTCAGTAGACAAAGCAACCTTAAATCGATTTTTCGCCCTCCATTTCATTCTTCCATTTACTATAAT : 148
Lox-Noro1  : .................T........................................................ : 148
Lox-Noro2  : .................T........................................................ : 148
Lox-Noro3  : .................T........................................................ : 148
Lox-Hauf   : .................T........................................................ : 148
Ele-Noro1  : ...........G.....T..............C.....C......T..........C...............G. : 148
Ele-Noro2  : ...........G.....T..............C.....C......T..........C...............G. : 148
Ele-Noro3  : ...........G.....T..............C.....C......T..........C...............G. : 148
Ele-Ozawa  : ...........G.....T..............C.....C......T..........C...............G. : 148
Mam-Noro   : ...........G.....T....................C..............T.................... : 148
Mam-Ozawa  : ...........G.....T....................C..............T.................... : 148
Mam-DM1    : ...........G.....T....................C..............T.................... : 148
Mam-DM5a   : ...........G.....T....................C..............T.................... : 131
Mam-DM6    : ...........G.....T....................C..............T.................... : 112
Mam-DM13   : ------------------------------------------------------------------........ :   8
3-Mamm     : ------------------------------------------------------------------........ :   8
3Mam-Hauf  : ----------------------................C..............T.................... :  52
Mast-Hauf  : ---------------------------------------------------------------........... :  11
Dug-Irwin  : ..G..A..C..............C...C.C.CC.....C........A..C.....C..A..C..C.TCG...C : 148
Pro-Ozawa  : .....A..C.....G........C.....G.CC.....C......T.T..C.....CA.A.......TC..T.. : 148

             *       580         *       600         *       620         *       640
Lox-Irwin  : TGCACTAGCAGGAGTACACCTAACCTTTCTTCACGAAACAGGCTCAAACAATCCACTGGGCCTCACTTCAGACT : 222
Lox-Noro1  : .......................................................................... : 222
Lox-Noro2  : ...................................................C.............T........ : 222
Lox-Noro3  : ...................................................C.............T........ : 222
Lox-Hauf   : .......................................................................... : 222
Ele-Noro1  : ...............G..........................T........C.....A..T............. : 222
Ele-Noro2  : ...............G...................................C.....A..T............. : 222
Ele-Noro3  : ...............G...................................C.....A..T............. : 222
Ele-Ozawa  : ...............G...................................C.....A..T............. : 222
Mam-Noro   : ...........................C....................T..C.....A................ : 222
Mam-Ozawa  : ...........................C....................T........A................ : 222
Mam-DM1    : ...........................C....................T..C.....A................ : 222
Mam-DM5a   : ...........................C....................T..C.....A................ : 205
Mam-DM6    : ...........................C....................T..C.....A................ : 186
Mam-DM13   : ...........................C....................T..C.....A................ :  82
3-Mamm     : ...........................C....................T..C.....A................ :  82
3Mam-Hauf  : ...........................C....................T..C.....A................ : 126
Mast-Hauf  : ...........................C....................T..C...................... :  85
Dug-Irwin  : C..C....T.AT...C...T..CTA..C..C..............C.....C..CAC...A..G.TC..C.... : 222
Pro-Ozawa  : A..C......AT...C.....TCTA..CT.A..............T.....C.....A..A..A.TC..CA..G : 222
```

APPENDIX 2 continued



```
                  *       660         *       680         *       700         *
Lox-Irwin : CAGACAAAATCCCCTTTCACCCATACTATACCATTAAGGACTTCCTAGGATTACTTATCCTAATTTTACTTCTT : 296
Lox-Noro1 : .......................................................................... : 296
Lox-Noro2 : ......................G..............A.................................... : 296
Lox-Noro3 : ......................G..............A.................................... : 296
Lox-Hauf  : .......................................................................... : 296
Ele-Noro1 : ..........T...........G........T..C..A............C...................C... : 296
Ele-Noro2 : ..........T...........G........T..C..A............C...................C... : 296
Ele-Noro3 : ..........T...........G........T..C..A...........GC...................C... : 296
Ele-Ozawa : ..........T...........G........T..C..A...........GC...................C... : 296
Mam-Noro  : ......................G...........C..A............C.............CC..T.C... : 296
Mam-Ozawa : ........G.............G...........C..A............C.............CC..T.C... : 296
Mam-DM1   : ......................G...........C..A............C.............CC..T.C... : 296
Mam-DM5a  : ......................G...........C..A............C.............CC..T.C... : 279
Mam-DM6   : ......................G...........C..A............C.............CC..T.C... : 260
Mam-DM13  : ......................G...........C..A............C.............CC..T.C... : 156
3-Mamm    : ......................G...........C..A............C.............CC..T.C... : 156
3Mam-Hauf : ......................G...........C..A............C.............CC..T.C... : 200
Mast-Hauf : ..................................................C.............CC....C... : 159
Dug-Irwin : .............A..C........T...T.AG.C..A...C.......CC..T.CC....C...C..G.CT.A : 296
Pro-Ozawa : .C...........A........C..T........A..A...C.AGC....C..G.AC.A..C.CCC..ACA..C : 296

            720         *       740         *       760         *       780         *
Lox-Irwin : CTACTCCTAGCCCTACTATCTCCTGACCATCTAGGAGACCCTGACAACTACACCCTGGCCAACCCCTTAAATAA : 370
Lox-Noro1 : ...........................ATA......................TA.CA...G....AC......C : 370
Lox-Noro2 : .................G.........ATA......................TA.CA...G....AC......C : 370
Lox-Noro3 : .................G.........ATA......................TA.CA...G....AC......C : 370
Lox-Hauf  : ...........................ATA......................TA.CA...G....AC......C : 370
Ele-Noro1 : ......T................A...ATA......................TA.CA..TG.T..AC......C : 370
Ele-Noro2 : ......T................A...ATA......................TA.CA..TG.T..AC......C : 370
Ele-Noro3 : ......T................A...ATA......................TA.CA..TG.T..AC......C : 370
Ele-Ozawa : ......T................A...ATA...........C..........TA.CA..TG.T..AC......C : 370
Mam-Noro  : ......T....................ATA...........C..........TA.CA..TG.T..AC.T....C : 370
Mam-Ozawa : ......T....................ATA...........C..........TA.CA..TG.T..AC.T....C : 370
Mam-DM1   : ......T....................ATA...........C..........TA.CA..TG.T..AC.T....C : 370
Mam-DM5a  : ......T....................ATA...........C..........TA.CA..TG.T..AC.T....C : 353
Mam-DM6   : ......T....................ATA...........C..........TA.CA..TG.T..AC.T....C : 334
Mam-DM13  : ......T....................ATA...........C..........TA.CA..TG.T..AC.T....C : 230
3-Mamm    : ......T....................ATA...........C..........TA.CA..TG.T..AC.T....C : 230
3Mam-Hauf : ......T....................ATA...........C..........TA.CA..TG.T..AC.T....C : 274
Mast-Hauf : ...........................ATA......................TA.CA...G....AC.T....C : 233
Dug-Irwin : ..C..A...A....GT.C..C..G...ATA..G........A...........A.CA........AC....C.C : 370
Pro-Ozawa : ...ACA...A.T...T.C.TC..A...ATA...........C..T.....T..A.CC.........C.C..C.C : 370

                  800         *       820         *       840         *       860
Lox-Irwin : CCCCCCTCATATCAAGCCAGAGTGATATTTTCTCTTTGCTTACGCCATCCTACGATCTGTACCAAACAAACTAG : 444
Lox-Noro1 : T....TA................................................................... : 444
Lox-Noro2 : T....TA................................................................... : 444
Lox-Noro3 : T....TA................................................................... : 444
Lox-Hauf  : T....TA................................................................... : 444
Ele-Noro1 : T....TA..C.....A...........C..C..T..............T......................... : 444
Ele-Noro2 : T....TA........A...........C..C..T..............T.................T....... : 444
Ele-Noro3 : T....TA..C.....A...........C..C..T..............T......................... : 444
Ele-Ozawa : T....TA..C.....A...........C..C..T..............T......................... : 444
Mam-Noro  : T....TA..C.....A.......................................................... : 444
Mam-Ozawa : T....TA..C.....A...........C..C........................................... : 444
Mam-DM1   : T....TA..C.....A...........C..C........................................... : 444
Mam-DM5a  : T....TA..C.....A...........C..C........................................... : 427
Mam-DM6   : T....TA..C.....A...........C..C........................................... : 408
Mam-DM13  : T....TA..C.....A.......................................................... : 304
3-Mamm    : T....TA..C.....A.......................................................... : 304
3Mam-Hauf : T....TA..C.....A...........C..C..............................------------- : 335
Mast-Hauf : T....TA..C.....A...........C..C..............................------------- : 294
Dug-Irwin : ...T..C..C..T..A.....A.....C.....A..CCGA.....T.....C......A.C..T..T....... : 444
Pro-Ozawa : ...T..A..C.....A..C..A.....C..C..A..C..A.....A...........CA.T.....T....... : 444
```

APPENDIX 2 continued

```
                *       880         *       900         *       920         *       94
Lox-Irwin : GAGGCGTCCTAGCCCTACTCCTATCAATTCTAATCCTAGGATTAATACCACTTCTCCATACATCCAAGCACCGA : 518
Lox-Noro1 : .......................................................................... : 518
Lox-Noro2 : .......................................................................... : 518
Lox-Noro3 : .......................................................................... : 518
Lox-Hauf  : .......................................................................... : 518
Ele-Noro1 : .................T.T........C..G..TT...................................... : 518
Ele-Noro2 : .................T.............G..TT..............T....................... : 518
Ele-Noro3 : .................T.T........C..G..TT..............T....................... : 518
Ele-Ozawa : .................T.T........C..G..TT..............T....................... : 518
Mam-Noro  : ............................C............A.T...........A........T..A...... : 518
Mam-Ozawa : ............................C............A.T...........A........T..A...... : 518
Mam-DM1   : ............................C............A.T...........A........T..A...... : 518
Mam-DM5a  : ............................C............A.T...........A........T..A...... : 501
Mam-DM6   : ............................C............A.T...........A........T..A...... : 482
Mam-DM13  : ............................C............A.T...........A........T..A...... : 378
3-Mamm    : ............................C............A.T...........A........T..A...... : 378
3Mam-Hauf : -------------------------------------------------------------------------- :   -
Mast-Hauf : -------------------------------------------------------------------------- :   -
Dug-Irwin : .C.....GT.......CG.A..C..C..C..........CGC.CC.C.....C.....C........A..A... : 518
Pro-Ozawa : ....A..AA.T......A..A..........C..T....C.C.CC.C..C..A.....C.....A..A..G... : 518

            0         *       960         *       980         *      1000
Lox-Irwin : AGCATAATACTCCGACCTCTTAGCCTATGTGCCTATTGCTGAACTCTAACAATAGATTTACTAACACTT : 587
Lox-Noro1 : .........................-.A..-......-............................... : 584
Lox-Noro2 : .........................-.A..-......-............................... : 584
Lox-Noro3 : .........................-.A..-......-............................... : 584
Lox-Hauf  : .........................-.A..-......-............................... : 584
Ele-Noro1 : ..T.................C....-.G..-......-............................... : 584
Ele-Noro2 : ..T.................C....-.G..-......-............................... : 584
Ele-Noro3 : ..T.................C....-.G..-......-............................... : 584
Ele-Ozawa : ..T.................C....-.G..-T.....-............................... : 584
Mam-Noro  : ..T..G.....T.............-.A..-......-..........G...C....C......T.... : 584
Mam-Ozawa : ..T..G.....T.............-.A..-......-..........G...C....C......T.... : 584
Mam-DM1   : ..T..G.....T.............-.A..-......-T....-------------------------- : 558
Mam-DM5a  : ..T..G.....T......--------------------------------------------------- : 519
Mam-DM6   : ..T..G.....T.............-.A..-......-..........G...C....C......T.... : 548
Mam-DM13  : ..T..G.....T.............-.A..-......-..........G...C....C...C..T...- : 443
3-Mamm    : ..T..G.....T.............-.A..-......-..........G...C....C......T...- : 443
3Mam-Hauf : --------------------------------------------------------------------- :   -
Mast-Hauf : --------------------------------------------------------------------- :   -
Dug-Irwin : ...C..TC.T..........A....-.ATG-...C..-T....T...GGT.GCC..CC.GA.C.....C : 584
Pro-Ozawa : ...C.....T.T..T..A..C....-.ATG-......-....CT....GT.GCT..CC.TA.C.....A : 584
```

DEINSEA - ANNUAL OF THE NATURAL HISTORY MUSEUM ROTTERDAM
P.O.Box 23452, NL-3001 KL Rotterdam The Netherlands