THE ROYAL SOCIETY

Molecular and morphological evidence on the phylogeny of the Elephantidae

Mark G. Thomas[1], **Erika Hagelberg**[2], **Hywel B. Jones**[3], **Ziheng Yang**[1] **and Adrian M. Lister**[1*]

[1] *Department of Biology, University College London, London WC1E 6BT, UK*
[2] *Department of Biochemistry, University of Otago, PO Box 56, Dunedin, New Zealand*
[3] *Department of Genetics, School of Medicine, Stanford University, Stanford, CA 94305, USA*

The African and Asian elephants and the mammoth diverged *ca.* 4–6 million years ago and their phylogenetic relationship has been controversial. Morphological studies have suggested a mammoth–Asian elephant relationship, while molecular studies have produced conflicting results. We obtained cytochrome *b* sequences of up to 545 base pairs from five mammoths, 14 Asian and eight African elephants. A high degree of polymorphism is detected within species. With a dugong sequence used as the outgroup, parsimony and maximum-likelihood analyses support a mammoth–African elephant clade. As the dugong is a very distant outgroup, we employ likelihood analysis to root the tree with a molecular clock, and use bootstrap and Bayesian analyses to quantify the relative support for different topologies. The analyses support the mammoth–African elephant relationship, although other trees cannot be rejected. Ancestral polymorphisms may have resulted in gene trees differing from the species phylogeny. Examination of morphological data, especially from primitive fossil members, indicates that some supposed synapomorphies between the mammoth and Asian elephant are variable, others convergent or autapomorphous. A mammoth–African elephant relationship is not excluded. Our results highlight the need, in both morphological and molecular phylogenetics, for multiple markers and close attention to within-taxon variation and outgroup selection.

Keywords: elephant; mammoth; gene tree; phylogeny; molecular evidence; morphological evidence

1. INTRODUCTION

The elephant family, Elephantidae, originated in Africa in the Late Miocene. In the period 4–6 million years (Myr) ago, the earliest representatives of the three main lineages—*Loxodonta*, *Elephas* and *Mammuthus*—all make their appearance in deposits in eastern and southern Africa (Maglio 1973; Kalb & Mebrate 1993). Today only two species remain: the African elephant (*Loxodonta africana* L.) and the Asian elephant (*Elephas maximus* L.). The last representative of the third lineage, the woolly mammoth (*Mammuthus primigenius*), became extinct *ca.* 3.7 thousand years (Kyr) ago (Lister & Bahn 2000).

Dental and skeletal morphology have been taken to indicate that *Elephas* and *Mammuthus* are more closely related than either is to *Loxodonta* (Tassy & Shoshani 1988; Kalb & Mebrate 1993; Kalb & Froehlich 1995; Kalb *et al.* 1996; Tassy 1996; Shoshani *et al.* 1998). DNA sequence data have produced conflicting results. Yang *et al.* (1996) and Ozawa *et al.* (1997) suggest an *Elephas–Mammuthus* relationship, while Hagelberg *et al.* (1994), Noro *et al.* (1998) and Barriel *et al.* (1999) favoured a *Loxodonta–Mammuthus* clade. All of these studies have used the mitochondrial cytochrome *b* gene, with the addition of a 12S rRNA gene sequence by Noro *et al.* (1998).

In this study, we present new mtDNA sequence data from the cytochrome *b* gene of five *M. primigenius*, 14 *E. maximus* and eight *L. africana* individuals. Use of multiple individuals from the same species allowed us to take account of intraspecific variation. None of the previous studies has compared more than one or two individuals of each species. We also compare trees built with and without outgroups, quantify the relative support for different topologies, and assess the significance of lineage sorting for studies of this kind. Finally, we reappraise the morphological data and compare them with the molecular results.

2. MATERIAL AND METHODS

(a) *DNA samples, PCR, and sequencing*

DNA was obtained from compact bone of five *M. primigenius* individuals (Mam-2–Mam-6). All samples are from permafrost localities in north-east Siberia. Mam-2 is from Oyagossky Yar: scapula at the Zoological Institute, St Petersburg (ZIN), collected in 1990. Mam-3 and Mam-4 are from Berelekh: juvenile humerus, no. 30957/22, and juvenile scapula, 30957/12 respectively, collected in 1970. The specimens were radiocarbon dated to *ca.* 14–12 Kyr ago (Lister & Bahn 2000; Vereshchagin 1997). Mam-5 is from Khatanga: humerus, no. 31829, from partial carcass at ZIN, excavated in 1977 (Vereshchagin & Nikolaev 1982). Three ^{14}C dates on this carcass (one from the sampled bone) have given an 'infinite' age of > 47 Kyr old. Mam-6 is from Allaicha, mandible excavated in 1975 from beneath a horizon ^{14}C-dated to > 47 Kyr (Kaplina *et al.* 1980).

Blood of living elephants was obtained from the Institute of Zoology, London. Fourteen *E. maximus* (Ele-1–Ele-14) and eight *L. africana* (Lox 1–8) were sampled. The *Elephas* individuals were from various regions of India, Sri Lanka and Burma. The *Loxodonta* individuals were all savannah elephants from southern and eastern Africa. Each DNA sample was prepared for

*Author for correspondence (a.lister@ucl.ac.uk).

Proc. R. Soc. Lond. B (2000) **267**, 2493–2500
Received 4 July 2000   Accepted 3 October 2000

2493

```
                    111111111112222222222333333333334444444444444555555555
                    881234455667901124667701233444560013445666790123456  68
            N       157924506251576986470305106258562548789245527384128  95

Lox-ref     1       AAAAAATCCTTCTTTACACTCCAGCGTTCCTCGAACTCCTATACACTTCTATA
Lox-1       2        G.CG......T...G......G.............T..G....T...
Lox-2       4        G.CG......T...G.....T...............T..G....T...
Lox-3       1        G.CG.C..C...C....T....A..C......GG.CTTC.C...T.C...C.
Ele-1       7        G.C......CT.CCC..T....A.AC.A..TA..TCTTC.C.TGTCCTCG.G
Ele-3       1        GGC......CTCCCC.GTCT..A.ACCATC.A..TCTTC.C.TGTCCTCG.G
Ele-2       4        GGC......CT.CCC.GTCT..ATACCATC.A..TCTTC.C.TGTCCTCG.G
Mam-3       1        CG.C.G.T...T.C.CT.T.......C.T.C.....CTTC.CG.GT.C.
Mam-4       1        CG.C.G.T...T.C.CT.T.......C...C.....CTTC.CG.GT.C.
Mam-6       1        CG.C.G.TT..T.C.CT.T.......C...C.....CTTC.CG.GT.C.
Mam-5       1         .CG..T...T...C..T.......C.
```

Figure 1. The ten haplotype sequences obtained in this study. Only sites that differ from the *Loxodonta* reference sequence (Lox-ref; Irwin *et al*. 1991) are shown.

sequencing as described in Thomas *et al*. (1997), and DNA sequencing was carried out according to Thomas *et al*. (1998).

DNA was extracted from *M. primigenius* bone samples as described in Hagelberg & Clegg (1991). For modern elephants, DNA was extracted from 10 µl of blood using a modification of the silica method (Höss & Pääbo 1993). Following extraction, DNA was dissolved in 100 µl of water.

A 567 base pair (bp) fragment of the cytochrome *b* gene was amplified from extracted DNA using a biotinylated version of the conserved mammalian primer L14841 [5′AAA AAG CTT CCA TCC AAC ATC TCA GCA TGA TGA AA 3′] (Irwin *et al*. 1991) and the degenerate primer H15347d [5′TGT (G/T)GG GTT GTT (G/T)GA TCC TGT TTC GTG 3′]. This oligo-nucleotide was designed to prime most mammalian cytochrome *b* sequences but with a bias towards elephantid sequences.

First-round amplifications were carried out as described previously (Thomas *et al*. 1998), except that each reaction contained 0.08 mg ml^{-1} bovine serum albumin (BSA) (Boehringer, Lewes, Sussex, UK) and 1.3 units of *Taq* DNA polymerase premixed with TaqStart antibody (Clontech, Palo Alto, CA, USA) in a molar ratio of 1:14. PCR cycling parameters: pre-incubation at 95 °C for 1 min, followed by 55 cycles of denaturation at 93 °C for 1 min, annealing at 55 °C for 1 min and extension at 72 °C for 1 min. Following a final incubation at 72 °C for 10 min, PCR reaction products were stored at 4 °C.

When re-amplification was employed, first-round amplification product was run on a 2% agarose gel and DNA was purified from the appropriate bands using the Geneclean method (Bio101). Second-round amplifications were then carried out as above except that no BSA or TaqStart antibody was used and the annealing temperature was increased from 55 to 70 °C. Prior to preparation for sequencing, each PCR sample was extracted with chloroform and a 5 µl aliquot was subjected to electrophoresis on a 2% agarose gel.

(b) *Phylogenetic analysis*

Sequences were entered on a computer using the UWGCG program SEQED and aligned using the program PILEUP (Genetics Computer Group 1994). Phylogenetic analysis using both maximum-likelihood and maximum-parsimony methods was conducted using the PHYLIP (Felsenstein 1995) and PAML (Yang 1997) packages. Heuristic tree search was performed using the Jumble option of the PHYLIP package, with 20 replicate runs with random addition of species. Candidate trees collected this way were further compared using bootstrap and Bayes posterior probabilities calculated with PAML, using two

nucleotide substitution models: JC69 (Jukes & Cantor 1969) and HKY85 (Hasegawa *et al*. 1985). The three codon positions were either treated equally or assigned different substitution rate parameters, which were estimated from the data.

In an initial analysis, we used the dugong (*Dugong dugon*; Irwin & Arnason 1994) as the outgroup to root the tree. The dugong is a member of the Order Sirenia, generally recognized to be the closest living relatives of elephants (Order Proboscidea) (Fischer 1996). Nevertheless, the divergence time between the two Orders is great. The earliest true fossil proboscideans are of Late Palaeocene age (Gheerbrant *et al*. 1996), *ca*. 58–54 Myr ago. This suggests a separation from the Sirenia at least 60 Myr ago. Novacek (1992) implied a separation at *ca*. 70 Myr ago. We thus used 65 Myr ago. We also used the 228 bp sequence from a fossil bone of American mastodon (*Mammut americanum*; Yang *et al*. 1996). This is a member of the proboscidean family Mammutidae, which probably diverged from the lineage leading to the Elephantidae in the Early Miocene, *ca*. 24 Myr ago (Tassy 1996; Yang *et al*. 1996).

(c) *Morphological data*

Several studies have adduced evidence on elephantid relationships based on morphology (e.g. Shoshani *et al*. 1985). Tassy & Shoshani (1988), Tassy (1996) and Shoshani *et al*. (1998) have listed key characters apparently synapomorphous between late *Elephas* and *Mammuthus*. In assessing all such studies, the character states and their distributions must be carefully examined, particularly in relation to (i) the possibilities of convergence, (ii) character polarity, (iii) correlations between characters, and (iv) taking a single (usually advanced) species to represent each genus, which might have autapomorphies not characteristic of the genus as a whole. Evidence on these issues can be obtained from examining outgroups, in conjunction with primitive fossil representatives of the ingroup taxa. Much morphological work has understandably been based on the terminal taxa *L. africana*, *E. maximus* and *M. primigenius*, but Kalb, Froehlich and colleagues have gone some way toward including earlier material in a consideration of elephantid phylogeny (Kalb & Mebrate 1993; Kalb *et al*. 1996; Kalb & Froehlich 1995).

In the present study, morphological characters for *Mammuthus*, *Loxodonta* and *Elephas*, which have previously been cited as evidence of their relationships, were examined on skulls of *M. meridionalis* (Early Pleistocene), *M. primigenius* (Late Pleistocene), *L. africana* and *E. maximus* (Recent) in the collections of the Natural History Museum, London, the Muséum National d'Histoire Naturelle, Paris, and the Geological–Palaeontological

Museum of the University of Florence. Additional data were obtained from published sources.

3. MOLECULAR RESULTS

Amplification of the cytochrome *b* sequence was attempted directly from DNA extracted from *M. primigenius* bone on four occasions. All PCR negative controls yielded no product, on all occasions except one. On this occasion, all PCR products were rejected. All extraction blank samples yielded no product. All PCR product obtained from *M. primigenius* template DNA was re-amplified to obtain sufficient DNA for direct sequencing. One set of first-round PCRs had, however, yielded sufficient DNA for this purpose prior to re-amplification. All amplifications from modern elephant DNA yielded sufficient DNA for sequencing without the need for re-amplification.

Lengths of sequence obtained ranged from 543 to 592 bp in *Elephas*, 542–594 bp in *Loxodonta*, and 102–545 bp in *Mammuthus*. Comparison of the 27 sequences revealed ten haplotypes: four *Mammuthus*, three *Loxodonta* and three *Elephas* (figure 1). Sample Mam-2 yielded only a 102 bp region, identical to Mam-3, Mam-4 and Mam-6, and was omitted. Each compared sequence was at least 453 bases long with the exception of Mam-5, which was only 255 bases long. We also include the *Loxodonta* sequence of Irwin *et al.* (1991).

The DNA fragment amplified here overlaps that of our former study (Hagelberg *et al.* 1994), and two of the *M. primigenius* samples (Mam-5 and Mam-6) are the same specimens, respectively, as Mam-1 and Mam-2 reported there. This allows us to compare cytochrome *b* sequences obtained using different primers. Because the PCR fragment in this work is longer than that recovered previously, it is extremely unlikely that previous PCR products could have acted as a DNA template for the PCR reactions. Over the region of overlap, the sequences of Mam-5 and Mam-6 were identical to the reported Mam-1 and Mam-2 sequences, respectively. The A indicated at position 366 for Mam-2 (Hagelberg *et al.* 1994) was a typographical error for G.

We first conducted phylogenetic analysis using the dugong as the outgroup. Twelve equally parsimonious trees were found (each requiring 163 steps); all had the same basic topology. To gauge the robustness of this topology, 1000 bootstrap samples were analysed and a majority rule consensus tree constructed (figure 2*a*). This is identical to one of the 12 equally parsimonious trees. The bootstrap score of 84% suggests that *Loxodonta* and *Mammuthus* are sister taxa. However, the percentages on the nodes delineating each of these two species are low (35 and 39%, respectively). Indeed, in many of the replicates, Mam-5 is placed as the earliest sequence in the *Loxodonta* group. The *Elephas* group is linked with a bootstrap value of 74%. The sequence Ele-1 shows the greatest similarity to *Mammuthus* sequences and is frequently placed as the earliest branch in that group. However, the other two *Elephas* haplotypes are never placed with *Loxodonta* or *Mammuthus* and are only separated from Ele-1 in 3.7% of the bootstrap replicates. Within *Mammuthus*, Mam-5 diverged earlier than all the other *Mammuthus* sequences in 88% of replicates. Among *Loxodonta*, Lox-3 branched off first in 88% of the replicates.
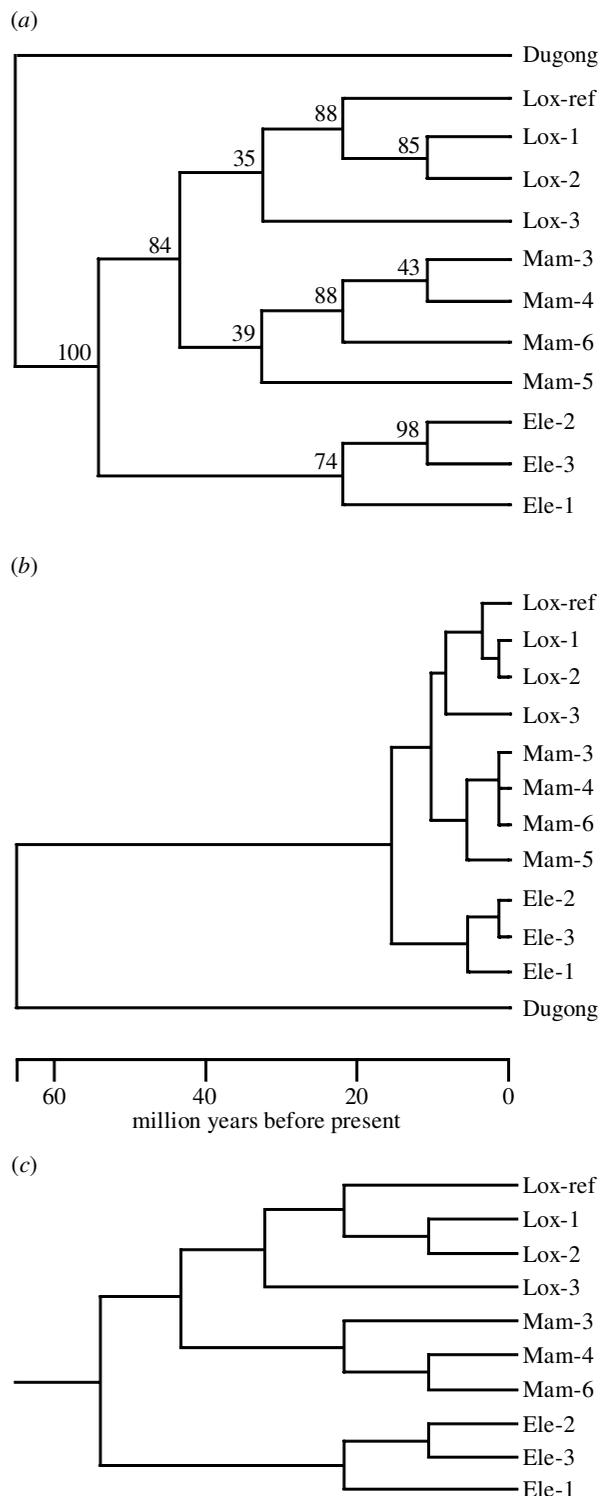
(*a*)

(*b*)

(*c*)

Figure 2. (*a*) Bootstrap consensus parsimony tree of elephantid cytochrome *b* sequences. The number at each node representing the percentage with which that group of sequences occurred in the 1000 bootstrap replicates. (*b*) Maximum-likelihood tree constructed under the assumption of a molecular clock and a dugong–elephant split of 65 Myr. (*c*) Maximum-likelihood tree obtained using the PAML package. The same topology was obtained with or without a molecular clock, and under two different substitution models.

Maximum-likelihood methods were used both for comparison with parsimony and to provide a rough estimate of the time of divergence of the species. Because the transition–transversion (Ts–Tv) ratio between the

Table 1. *Likelihood ratio test statistics of the molecular clock*
(d.f. = 8.)

model	with rates for codons	$2\Delta\ell$	P
JC69[a]	no	13.44	0.10
JC69	yes	13.38	0.12
HKY85[b]	yes	13.52	0.10

[a] Source: Jukes & Cantor (1969).
[b] Source: Hasegawa *et al*. (1985).

elephantid sequences and the dugong was considerably lower than that among the elephantids, we performed four separate analyses, using different Ts–Tv ratios (2, 10, 15 and 20). With the data set including the dugong sequence, a Ts–Tv ratio of 2 consistently gave the highest log-likelihood score. Figure 2*b* shows the maximum-likelihood tree constructed under the assumption of a molecular clock (i.e. all branch tips equidistant from the root). This tree shows *Elephas* separating first, followed by the *Loxodonta–Mammuthus* split. The topology of this maximum-likelihood tree is identical to one of the most parsimonious trees and differs from the consensus parsimony tree only in the placement of Mam-6. By using a Sirenia–Proboscidea split of 65 Myr for calibration, the common node of the cytochrome *b* tree for the three genera (the divergence of the *Elephas* branch) was estimated at 14.8 Myr ago, while the *Mammuthus–Loxodonta* node was at *ca*. 9.4 Myr ago. These date estimates are tentative and are based on the assumption that the Proboscidea and the Sirenia have evolved at the same rate. If the Proboscidea have had a faster rate of mtDNA evolution, as suggested by the results of Hauf *et al*. (1999) on the complete mitochondrial genome of *L. africana*, then our date estimates are likely to be too high.

We also note high genetic diversity reflected in ancient coalescence dates within each of the three species, *ca*. 5.0 Myr ago for *Elephas*, *ca*. 5.3 Myr ago for *Mammuthus*, and *ca*. 7.7 Myr ago for *Loxodonta* (figure 2*b*). This is consistent with the results of Georgiadis *et al*. (1994) for *Loxodonta*, while Fernando *et al*. (2000), using the faster mitochondrial control region, found a somewhat lower but still notably old coalescence date of 2.5–3.5 Myr ago for *Elephas*.

To test the assumption of a constant rate of evolution, a likelihood ratio test (Felsenstein 1995) was carried out, comparing the tree obtained under the assumption of a molecular clock (figure 2*b*) with an unrooted version of the same tree, assuming a Ts–Tv ratio of 2. Although the likelihood scores were not significantly different at the 95% level ($0.05 < p < 0.1$), the small *p*-value does indicate rate variation among the elephantid lineages. For this reason, the above divergence dates must be treated with caution.

Because of the large relative divergence time and dissimilarity in the Ts–Tv ratio between the elephantid sequences and the dugong sequence, this species appears problematic as an outgroup. As an alternative way of locating the root of the elephantid tree, maximum-likelihood analysis was performed on the ingroup species only, with and without the assumption of a molecular clock and using two substitution models in the PAML

package (Yang 1997). Because PAML ignores sites for which there are missing data for one or more taxa, the 257 bp Mam-5 sequence was removed to keep more sites in the data. Using this data set, the same tree was obtained with each substitution model, whether a molecular clock was enforced or not (figure 2*c*). Furthermore, the likelihood ratio test showed no significant difference between the clock and the no-clock models (table 1). So within the Elephantidae, the molecular clock assumption holds. The root can thus be located without an outgroup, and indicates an overall (Ele (Lox, Mam)) topology.

Bootstrap proportions (Felsenstein 1995) and posterior probabilities (Yang & Rannala 1997) were calculated to assess the support of this topology by the data. One hundred bootstrap samples were analysed by maximum likelihood, assuming a molecular clock, and a rooted consensus tree constructed. A monophyletic Lox–Mam clade was supported in 67% of replicates, whereas a monophyletic Ele–Mam clade was supported in 29%. Monophyletic groups for each of the three species, *Mammuthus*, *Elephas* and *Loxodonta*, were supported in 100, 99 and 77% of replicates, respectively.

Posterior probabilities for different tree topologies were calculated using the Bayesian method of Yang & Rannala (1997). Because this approach is computationally intensive, we removed one *Mammuthus* (Mam-6) and one *Elephas* (Ele-3) from the analysis. Probabilities for the six best trees summed to 0.998 and are shown in figure 3. Although the maximum-likelihood topology grouped *Mammuthus* and *Loxodonta*, the posterior probability score was only 0.43. Furthermore, the second and third best trees, with posterior probabilities of 0.28 and 0.20, respectively, grouped *Mammuthus* and *Elephas*.

To assess recent claims of resolution of elephantid phylogeny using American mastodon cytochrome *b* sequences (Yang *et al*. 1996), we constructed parsimony (with 1000 bootstrap re-samplings) and maximum-likelihood trees using Yang *et al*.'s mastodon sequence instead of dugong as the outgroup. The parsimony tree, as in Yang *et al*. (1996), did group *Mammuthus* and *Elephas* in a clade, although bootstrap support was low (35%) and the *Loxodonta* sequences did not form a single clade. Moreover, all maximum-likelihood trees (assuming four different Ts–Tv ratios and including those assuming a molecular clock) grouped *Mammuthus* with *Loxodonta* (results not shown). It should be borne in mind that only 200 bp of mastodon sequence were available for this analysis, whereas over 450 bp were available for most of the elephantid sequences in this study.

In a separate analysis, we constructed parsimony (with 1000 bootstrap re-samplings) and maximum-likelihood (with various Ts–Tv ratios) trees using all the sequences in the above analysis plus the *Mammuthus* and elephant sequences published by Yang *et al*. (1996), the *Mammuthus* and *Elephas* sequences of Ozawa *et al*. (1997), and human, pig, and black rhino sequences (Irwin *et al*. 1991; Anderson *et al*. 1981). The *Elephas* sequence of Yang *et al*. (1996) was not included because it is identical to our Ele-3 in the region of overlap.

Bootstrap support for most ingroup branches was poor. Despite a number of tree topologies with different methods, one consistent feature was that the *Mammuthus* sequences of Yang *et al*. (1996) always formed a clade with
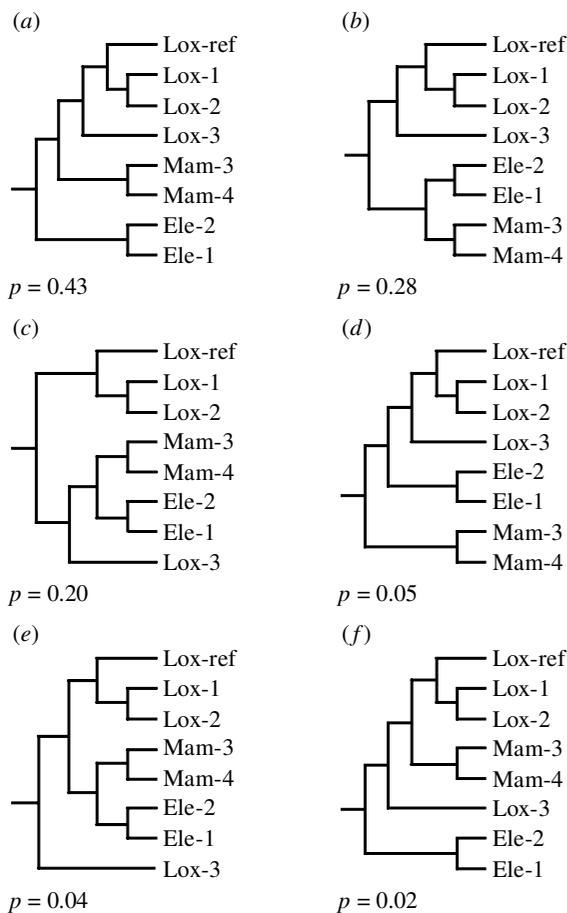
Figure 3. Posterior probabilities (p) for the six highest-scoring tree topologies, calculated using the Bayesian method of Yang & Rannala (1997).
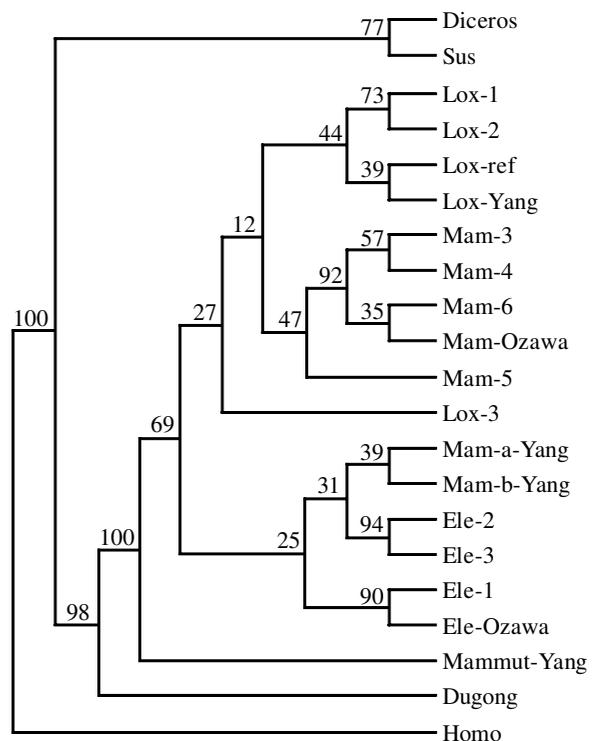

Figure 4. Consensus parsimony tree, including the *M. primigenius* sequences presented by Yang *et al.* (1996) and Ozawa *et al.* (1997). Bootstrap values (1000 re-samplings) are shown.

Elephas, whereas the *Mammuthus* sequences of this study and of Ozawa *et al.* (1997) fell outside the *Elephas* clade. This is typified by the consensus parsimony tree in figure 4. A single *Mammuthus* clade including all of our *M. primigenius* and that of Ozawa *et al.* (1997), together with those of Yang *et al.* (1996), is supported in only 0.018% of replicates, whereas a clade that grouped our sequences with only those of Ozawa *et al.* (1997), was supported in 47% of replicates, or 92% if Mam-5 was not included. Derenko *et al.* (1997) and Barriel *et al.* (1999) similarly observed an apparent biphyletic placement of *Mammuthus* cytochrome *b* sequences when combining their data with those of Yang *et al.* (1996). It seems that either polymorphism in *Mammuthus primigenius* is so great as to destroy inter-generic distinctions, or else Yang *et al.*'s (1996) reported mammoth sequences are erroneous.

4. DISCUSSION OF MOLECULAR STUDY

(a) *Phylogeny of Elephantidae*

The phylogenetic problem addressed in this paper is really one of placing the root in a three-taxon tree. We have shown that both outgroup rooting using the dugong, and maximum-likelihood ingroup rooting using a molecular clock, locate the root in the *Elephas* lineage and suggest an (Ele (Lox, Mam)) relationship. The statistical support for this relationship, evaluated using either the bootstrap or Bayesian method, is insufficient to confidently exclude the alternative *Mammuthus–Elephas* relationship. Of further relevance is the question of whether a single gene tree, even if well supported, represents the true species tree, since interspecific gene coalescence always precedes taxonomic divergence (Avise 1994). For mtDNA, the probability that the gene tree and species tree are different due to ancestral polymorphism is a function of the female effective population size at the time of the first speciation event (N_f) and the time in generations between the first and second speciation events (T), and can be calculated using the formula $P = \frac{2}{3}e^{-T/N_f}$ (Nei 1987). If we assume an upper limit for T in elephantids of 5 Myr, based on fossil evidence (see below), and a generation time of 30 years, then probabilities of less than 0.05 or 0.01 would require female effective population sizes in the ancestral species of $< 65\,000$ and $< 40\,000$, respectively. As N_f increases above these values, so does P. For example, if the female effective population size in the common ancestral species of living elephants was similar to that in modern elephant species, we can estimate N_f from the time to the most recent common ancestor (MRCA) in one of the species sampled in this study, using the formula $T_{\mathrm{MRCA}} = 2N_f$. Using a conservative estimate of $T_{\mathrm{MRCA}} = 5$ Myr ago for *Elephas* (see figure 2b) and assuming a generation time of 30 years, this corresponds to $N_f = 83\,000$ breeding females. This hypothetical but plausible combination of parameters gives a probability of 0.09—well above conventional significance levels—that the true gene tree does not represent the species tree.

If the mitochondrial gene tree is different from the species tree, more extensive sequencing of the mitochondrial genome will increase the chance of inferring

the correct gene tree, but will not improve the chance of obtaining the true species tree. It will then be important to sequence additional loci, specifically nuclear genes; this is plausible since the first nuclear sequences for *Mammuthus* have recently been obtained (Greenwood *et al.* 1999). However, it should be noted that an mtDNA gene tree is still more likely to reflect the species tree than a single autosomal gene tree. Moore (1995) has shown that for values of T and $\mathcal{N}_f$ that would give a 0.95 probability of the mtDNA tree being congruent with the species tree, the probability of a single autosomal gene tree being congruent is only 0.62. He has further shown that under the above conditions, 16 fully resolved, unlinked autosomal gene trees would be required to obtain the same probability of congruence as a fully resolved mtDNA tree.

Finally, the fossil record does not exclude the possibility that T is much less than 5 Myr. In this case, the two speciation events may have been so close in time that their relative order will be difficult to resolve using DNA sequence data, even with a large number of independent gene trees.

(b) *Speciation dates*

Coalescence dates calculated from the mtDNA data are older than the earliest known fossil occurrences of the taxa. The molecular clock approach, assuming the Proboscidea–Sirenia split occurred 65 Myr ago, gives 14.8 Myr ago for the common cytochrome *b* ancestor of the three elephantid genera, 9.4 Myr ago for that of the *Loxodonta–Mammuthus* clade, and 7.7, 5.3 and 5.0 Myr ago, respectively, for *Loxodonta*, *Mammuthus* and *Elephas* (figure 2*b*). In contrast, based on fossil material, Tassy (1995) placed the origin of the Elephantidae at *ca.* 9 Myr ago, while the earliest *Loxodonta* appears at *ca.* 5.5–6.0 Myr ago in Uganda (Tassy 1986, 1995). Kalb *et al.* (1996) give 4.0–4.5 Myr ago as the earliest date for *Mammuthus*, based on remains from South Africa and Ethiopia tentatively referred to as *M. subplanifrons*. Finally, the earliest known *Elephas*, *E. recki* and *E. ekorensis*, first occur in Ethiopian deposits some 4.0–4.5 Myr ago (Kalb & Mebrate 1993).

Earliest fossil occurrences are always subject to revision on the basis of new finds, but the African Miocene–Pliocene is very well sampled. One possible reason for significant differences between fossil and molecular dates is lineage sorting, in other words that mtDNA coalescence for the Family Elephantidae, reflected in the cytochrome *b* tree, predates the taxonomic origin of the family. Saturation of nucleotide substitutions along the outgroup (dugong) lineage, and possible rate differences between the proboscidean and sirenian lineages might also account for the high date estimates.

5. MORPHOLOGICAL RESULTS AND DISCUSSION

Our results indicate that many characters, previously cited as synapomorphies between two of the three elephantid genera (usually between *Elephas* and *Mammuthus*), may be convergent, primitive or autapomorphous.

Maglio (1973) stated that the crania of *Elephas* and *Mammuthus* differed in significant features, but were more similar to each other than to that of *Loxodonta*. However,

in several major features, the differences probably do not have phylogenetic significance. Thus, the 'globular skull' of *Loxodonta* appears from Kalb & Mebrate's (1993) analysis to be an autapomorphy, so it does not provide evidence on the relationship of this genus to *Elephas* and *Mammuthus* (those two taxa merely retaining the primitive, narrow-skulled condition). Similarly, the high cranium of late *Elephas* and *Mammuthus* is, at least in part, a convergent acquisition, early skulls of each being lower (Maglio 1973; Boeuf 1983; Lister 1996).

Correlation with the convergent cranial expansion in *Elephas* and *Mammuthus*, especially in the parietal region, may account for several other characters that have been described as synapomorphies of these genera. These include the 'large dorsal parietal bulges' (Tassy & Shoshani 1988; Shoshani *et al.* 1998); the 'recessed occipital condyles' (Kalb & Froehlich 1995; Kalb *et al.* 1996), which were described as due to overlapping by the posterior cranium; and the concave fronto-parietal region (Kalb & Froehlich 1995; Kalb *et al.* 1996; Shoshani *et al.* 1998), a feature which has been linked (Tassy & Shoshani 1988; Tassy 1996) to the large dorsal parietal bulges. Illustrations in Boeuf (1983) indicate that skulls of primitive *Mammuthus meridionalis* from Chilhac, France, *ca.* 1.9 Myr old, have modest parietals and prominent occipital condyles.

In the forelimb, Tassy & Shoshani (1988) listed a shared feature in the scaphoid of *Elephas* and *Mammuthus*, where the lateral bulging of the magnum articular facet is said to be reduced compared with *Loxodonta* and outgroup taxa. However, carpal and tarsal morphology shows considerable intraspecific variability in elephantids, and this feature is not evident in the scaphoids figured and described by Andrews & Forster (1928) where, if anything, *Mammuthus* appears closer to *Loxodonta*. In addition, the transition from 'serial' to 'aserial' carpus morphology between early and late *Mammuthus* entailed a reduction in the area of contact between scaphoid and magnum (Garutt 1954; Lister 1996), so this feature could be convergent and needs to be examined on more basal members of the group.

In dental morphology, the sharing of numerous enamel lamellae in the molars of late representatives of both *Elephas* and *Mammuthus*, mentioned as a synapomorphy by various authors (e.g. Tassy & Shoshani 1988), is clearly a convergent acquisition: the earliest representatives of these genera have no more lamellae than the earliest *Loxodonta* (Maglio 1973).

Kalb & Froehlich (1995) and Kalb *et al.* (1996) have described further dental characters. On the occlusal surface, the lamellae of *Elephas* and *Mammuthus*, and the primitive elephantid *Primelephas*, have the shape of a posteriorly convex crescent, compared with anteriorly convex crescents in more primitive Proboscidea. The significance for the *Elephas–Loxodonta–Mammuthus* trichotomy is not entirely clear, however, as some *Loxodonta* (*L. adaurora*) show the *Elephas–Mammuthus* condition, while others (e.g. *L. exoptata* and *L. africana*) show a unique condition of lamellae convex on both anterior and posterior sides. Treating the anterior and posterior sides of the lamellae as independent characters, Kalb & Froehlich (1995) regarded the situation in *L. exoptata* and *L. africana* as intermediate, with the

anterior side primitive and the posterior side derived. In this analysis, lamellae of *Elephas, Mammuthus* and even *Primelephas* are perceived as sharing a synapomorphous convex posterior side, with *L. adaurora* as convergently similar. However, it seems equally parsimonious to regard the lamellar shape of *Elephas, Mammuthus, Primelephas* and *Loxodonta adaurora* as primitive for the Elephantinae, and that of the other *Loxodonta* species as a single autapomorphous character, in which case the character would provide no evidence of a special relationship between any of these genera.

Three characters may be phylogenetically informative. Two of these, which have been listed as synapomorphies of *Mammuthus* and *Elephas*, are a gracile stylohyoid bone and prominent, close maxillary ridges (interalveolar cristae) (Tassy & Shoshani 1988; Tassy 1996; Shoshani *et al.* 1998). Based on our observation of specimens in London and Paris, these characters appear valid in the terminal species *M. primigenius* and *E. maximus*; and illustrations in Boeuf (1983) indicate that the maxillary ridge character occurs in the more primitive *M. meridionalis* too.

One character, conversely, appears to link *Loxodonta* and the earliest *Mammuthus*: the presence of flared premaxillary bones (Kalb & Froehlich 1995). This character is interesting because in later *Mammuthus*, the premaxillaries are subparallel and not at all flared (Lister 1996), apparently convergent to *Elephas*, whereas in some later *Elephas* (e.g. *E. antiquus*) they are very strongly flared (Maglio 1973; Osborn 1942), apparently convergent to *Loxodonta*. This illustrates the importance of scoring early representatives of each genus.

Further characters indicated as synapomorphies of *Elephas* and *Mammuthus* (Tassy & Shoshani 1988; Tassy 1996; Shoshani *et al.* 1998) were not confirmed by observations on skulls of the three genera examined during the present study. These include angular corners to the orbital rim, and the presence of a process on the anterior border of the mandibular foramen. These characters presumably show intraspecies variation so their phylogenetic significance (if any) is unclear.

Our results indicate that the morphological evidence does not, contrary to previous authors, contradict the *Mammuthus–Loxodonta* relationship suggested by our mtDNA data. Most of the similarities between *Mammuthus* and *Elephas* appear to be interesting examples of convergence, including dietary-based dental adaptations suggesting graminivorous ancestry (cf. Maglio 1973), and cranial features, which may be in part sexually selected (Lister & Blashford-Snell 1999).

6. CONCLUSIONS

A reassessment of morphological and molecular evidence indicates that, contrary to recent claims, elephantid phylogeny is not resolved in favour of an *Elephas–Mammuthus* relationship. Both our mtDNA data and those of Noro *et al.* (1998) now favour a *Mammuthus–Loxodonta* link, a suggestion consistent with our revised appraisal of the morphological evidence. Using our own data, we have quantified the support for different gene trees, allowing for the first time some objectivity in their relative appraisal. Our analysis also gives an estimate of the probability that lineage sorting has resulted in a gene tree not accurately reflecting the phylogenetic relationship among the species.

Limits on speciation dates, provided by the fossil record, have given a valuable perspective on both the molecular and morphological analyses. The fossil data set has given a framework of 0–5 Myr for the interval between the first and second nodes within the elephantid phylogeny. At the lower limit, the nodes may be so close together that their resolution with either molecular or morphological data is difficult. At the upper limit, with the nodes up to 5 Myr apart, sufficient change should have accumulated to allow their resolution. Morphological studies have the advantage that much closer fossil relatives of the elephants (early elephantids, and their immediate sister groups such as stegodontids) are available as outgroups, with divergence times only a few million years prior to the elephantids themselves. On the other hand, the fossil record is patchy and adaptive convergence a serious problem. We have shown that these problems can to some extent be overcome by examining more than one species in each genus, especially including more primitive members. On the molecular side, carefully chosen molecular markers provide the advantage of a neutral system, while the use of many independent loci should overcome problems of lineage sorting. For furthering our knowledge of elephantid phylogeny, therefore, molecular and morphological studies must go hand in hand.

REFERENCES

Anderson, S. (and 13 others) 1981 Sequence and organization of the human mitochondrial genome. *Nature* **290**, 457–465.

Andrews, C. W. & Forster, C. F. 1928 *On a specimen of* Elephas antiquus *from Upnor*. London: British Museum (Natural History).

Avise, J. C. 1994 *Molecular markers, natural history and evolution.* New York: Chapman & Hall.

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

Boeuf, O. 1983 Le site villafranchien de Chilhac (Haute-Loire), France. Etude paléontologique et biochronologique. Thèse d'État, Université de Paris VII, France.

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.

Felsenstein, J. 1995 *PHYLIP (Phylogeny Inference Package)*, v. 3.57c. University of Washington.

Fernando, P., Pfrender, M. E., Encalada, S. E. & Lande, R. 2000 Mitochondrial DNA variation, phylogeography and population structure of the Asian elephant. *Heredity* **84**, 362–372.

Fischer, M. S. 1996 On the position of Proboscidea in the phylogenetic system of Eutheria: a systematic review. In *The Proboscidea: evolution and palaeoecology of elephants and their relatives* (ed. J. Shoshani & P. Tassy), pp. 35–38. Oxford University Press.

Garutt, V. E. 1954 Southern elephant *Archidiskodon meridionalis* (Nesti) from the northern coast of the Azov Sea. *Trudy Komissi po Izucheniyu Chetverichnogo Perioda* **10**, 3–76.

Genetics Computer Group 1994 *Program manual for the Wisconsin package*, v. 8, August 1994. Madison, WI: Genetics Computer Group.

Georgiadis, N., Bischof, L., Templeton, A., Patton, J., Karesh, W. & Western, D. 1994 Structure and history of African elephant populations. I. Eastern and southern Africa. *J. Hered.* **85**, 100–104.

Gheerbrant, E., Sudre, J. & Cappetta, H. 1996 A Palaeocene proboscidean from Morocco. *Nature* **383**, 68–70.

Greenwood, A. D., Capelli, C., Possnert, G. & Pääbo, S. 1999 Nuclear DNA sequences from Late Pleistocene megafauna. *Mol. Biol. Evol.* **16**, 1466–1473.

Hagelberg, E. & Clegg, J. B. 1991 Isolation and characterisation of DNA from archaeological bone. *Proc. R. Soc. Lond.* B **244**, 45–50.

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

Hasegawa, M., Kishino, H. & Yano, T. 1985 Dating of the human–ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.* **22**, 160–174.

Hauf, J., Waddell, P. J., Chalwatzis, N., Joger, U. & Zimmermann, F. K. 1999 The complete mitochondrial genome sequence of the African elephant (*Loxodonta africana*), phylogenetic relationships of Proboscidea to other mammals, and D-loop heteroplasmy. *Zool. Anal. Complex Syst.* **102**, 184–195.

Höss, M. & Pääbo, S. 1993 DNA extraction from Pleistocene bones by a silica-based purification method. *Nucleic Acids Res.* **21**, 3913–3914.

Irwin, D. M. & Arnason, V. 1994 Cytochrome *b* gene of marine mammals: phylogeny and evolution. *J. Mamm. Evol.* **2**, 37–55.

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

Jukes, T. H. & Cantor, C. R. 1969 Evolution of protein molecules. In *Mammalian protein metabolism* (ed. H. N. Munro), pp. 21–132. New York: Academic Press.

Kalb, J. E. & Froelich, D. J. 1995 Interrelationships of late Neogene elephantoids: new evidence from the Middle Awash Valley, Afar, Ethiopia. *Geobios* **28**, 727–736.

Kalb, J. E. & Mebrate, A. 1993 Fossil elephantoids from the hominid-bearing Awash Group, Middle Awash Valley, Afar Depression, Ethiopia. *Trans. Am. Phil. Soc.* **83**, 1–114.

Kalb, J. E., Froehlich, D. J. & Bell, G. L. 1996 Phylogeny of African and Eurasian Elephantoidea of the late Neogene. In *The Proboscidea: evolution and palaeoecology of elephants and their relatives* (ed. J. Shoshani & P. Tassy), pp. 101–116. Oxford University Press.

Kaplina, T. N., Sher, A. V., Giterman, R. E., Zazhigin, V. S., Kiselyov, S. V., Lozhkin, A. V. & Nikitin, V. P. 1980 The key section of the Pleistocene deposits on the Allaikha River (the Indigirka lowstream). *Bull. Comm. Quat. Period Res.* **50**, 73–95.

Lister, A. M. 1996 Evolution and taxonomy of Eurasian mammoths. In *The Proboscidea: evolution and palaeoecology of elephants and their relatives* (ed. J. Shoshani & P. Tassy), pp. 203–213. Oxford University Press.

Lister, A. & Bahn, P. 2000 *Mammoths: giants of the ice age.* London: Marshall Editions.

Lister, A. M. & Blashford-Snell, J. 1999 Exceptional size and form of Asian elephants in western Nepal. *Elephant* **2**, 33–36.

Maglio, V. J. 1973 Origin and evolution of the Elephantidae. *Trans. Am. Phil. Soc.* **62**, 1–149.

Mahboubi, M., Ameur, R., Crochet, J. Y. & Jaeger, J. J. 1984 Earliest known proboscidean from early Eocene of north-west Africa. *Nature* **308**, 543–544.

Moore, W. S. 1995 Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-gene trees. *Evolution* **49**, 718–726.

Nei, M. 1987 *Molecular evolutionary genetics.* New York: Columbia University Press.

Noro, M., Masuda, R., Dubrovo, I. A., Yoshida, M. C. & Kato, M. 1998 Molecular phylogenetic inference of the woolly mammoth *Mammuthus primigenius*, based on complete sequences of mitochondrial cytochrome *b* and 12S ribosomal RNA genes. *J. Mol. Evol.* **46**, 314–326.

Novacek, M. J. 1992 Mammalian phylogeny: shaking the tree. *Nature* **356**, 121–125.

Osborn, H. F. 1942 *Proboscidea*, vol. 2. New York: American Museum of Natural History.

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

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

Shoshani, J., Golenberg, E. M. & Yang, H. 1998 Elephantidae phylogeny: morphological versus molecular results. *Acta Theriologica* (Suppl. 5) 89–122.

Tassy, P. 1986 Nouveaux Elephantoidea (Mammalia) dans le Miocène du Kenya. *Cahiers de Paléontologie Est-Africaine*. Paris: CNRS.

Tassy, P. 1995 Les proboscidiens (Mammalia) fossiles du Rift occidental, Ouganda. In *Geology and Palaeobiology of the Albertine Rift Valley, Uganda-Zaire*. II. *Palaeobiology*, pp. 217–257. CIFEG Occasional Publications 1994/29. Orléans, France: CIFEG.

Tassy, P. 1996 Who is who among the Proboscidea? In *The Proboscidea: evolution and palaeoecology of elephants and their relatives* (ed. J. Shoshani & P. Tassy), pp. 39–48. Oxford University Press.

Tassy, P. & Shoshani, J. 1988 The Tethytheria: elephants and their relatives. In *The phylogeny and classification of tetrapods*. II. *Mammals* (ed. M. J. Benton), pp. 283–315. Oxford, UK: Clarendon.

Thomas, M. G., Hesse, S. A., Foss, Y. J. & Farzaneh, F. 1997 The use of PCR for differential screening of cDNA libraries. In *PCR cloning protocols* (ed. B. A. White), pp. 405–418. Totowa, NJ: Humana Press.

Thomas, M. G., Cook Jr, C. E., Miller, K. W. P., Waring, M. J. & Hagelberg, E. 1998 Molecular instability in the COII/tRNALys intergenic region of the human mitochondrial genome: multiple origins of the 9-bp deletion and heteroplasmy for expanded repeats. *Phil. Trans. R. Soc. Lond.* B **353**, 955–965.

Vereshchagin, N. K. (ed.) 1997 Mammoth fauna of the Russian plain and Eastern Siberia. *Trudy Zool. Inst. AN SSSR* **72**, 1–114.

Vereshchagin, N. K. & Nikolaev, A. I. 1982 Excavation of the Khatanga mammoth. *Trudy Zool. Inst.* **111**, 3–17.

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. *Proc. Natl Acad. Sci. USA* **93**, 1190–1194.

Yang, Z. 1997 PAML: a program for package for phylogenetic analysis by maximum likelihood. *CABIOS* **13**, 555–556.

Yang, Z. & Rannala, B. 1997 Bayesian phylogenetic inference using DNA sequences: a Markov chain Monte Carlo method. *Mol. Biol. Evol.* **14**, 717–724.