Available online at www.sciencedirect.com
SCIENCE DIRECT
Molecular Phylogenetics and Evolution 40 (2006) 620–627
MOLECULAR PHYLOGENETICS AND EVOLUTION
www.elsevier.com/locate/ympev

Short communication

A nuclear DNA phylogeny of the woolly mammoth (*Mammuthus primigenius*)

Cristian Capelli [a,1], Ross D.E. MacPhee [b,1], Alfred L. Roca [c,1], Francesca Brisighelli [a], Nicholas Georgiadis [e], Stephen J. O'Brien [d], Alex D. Greenwood [b,f,g,*]

[a] *Istituto di Medicina Legale, Universitá Cattolica del Sacro Cuore, Rome, Italy*
[b] *Division of Vertebrate Zoology, American Museum of Natural History, New York, New York, USA*
[c] *Laboratory of Genomic Diversity, Basic Research Program, SAIC-Frederick, Frederick, MD 21702, USA*
[d] *Laboratory of Genomic Diversity, National Cancer Institute, Frederick, MD 21702, USA*
[e] *Mpala Research Center, P.O. Box 555, Nanyuki, Kenya*
[f] *GSF-National Research Centre for Environment and Health, Institute of Molecular Virology, Neuherberg, Germany*
[g] *Technical University of Munich, Institute of Virology, Munich, Germany*

Received 2 November 2005; revised 28 February 2006; accepted 7 March 2006
Available online 2 May 2006

1. Introduction

Although the woolly mammoth (*Mammuthus primigenius*) is one of the most intensively studied extinct species at the DNA level, mitochondrial DNA (mtDNA) markers have failed to unambiguously resolve its phylogenetic affiliation within Elephantidae. Most mtDNA-based elephantid phylogenies associate mammoths with African elephants (*Loxodonta africana* and *Loxodonta cyclotis*) to the exclusion of the Asian elephant (*Elephas maximus*) (e.g., Debruyne et al., 2003; Noro et al., 1998). However, other mtDNA studies (Ozawa et al., 1997), including recent sequencing efforts that yielded the complete mitochondrial genomes of two woolly mammoths (Krause et al., 2006; Rogaev et al., 2006), suggested that the Asian elephant is the closest living affine of mammoths. However, relationships inferred from mtDNA may be misleading due to the absence of a closely related outgroup species, or to the radiation of the three elephantid genera in rapid succession, which can produce discordance between a species tree and a gene (mtDNA) tree due to lineage sorting processes. Another difficulty is that in certain species—including elephants—the presence of nuclear insertions of mitochondrial sequences (Numts) can make identifying organellar mtDNA problematic (Greenwood and Pääbo, 1999; Thalmann et al., 2004). Moreover, Numt sequences are a routine, if unwanted, result of the procedures used in ancient DNA studies (Greenwood et al., 1999). Recently, cytonuclear genomic dissociation has been observed in African elephants, likely due to past hybridization between species (Roca et al., 2005). The existence of such dissociation phenomena could also confound mtDNA analysis within or among other elephantid species.

To date, the only extinct elephantid that has been amenable to confirmable molecular analysis by multiple research groups working with different specimens is the woolly mammoth (for a recent summary, see Greenwood, 2001). Yet, given the lack of consistent results across mtDNA phylogenetic studies, and given the possibility of discrepancies between the mtDNA tree and the species tree due to lineage sorting processes or to cytonuclear dissociation, nuclear DNA offers an alternative approach to studying woolly mammoth phylogeny. Nuclear DNA sequences from mammoths and other well-preserved extinct megafauna have been reported (Greenwood et al., 1999; Greenwood et al., 2001; Poinar et al., 2003; Poinar et al., 2006), and in principle it should be possible to characterize mammoth nuclear DNA sequences for the purpose of phylogenetic analysis. Of additional relevance, several nuclear genes have been investigated in a large number of individuals from different populations of *E. maximus*, *L. africana*, and *L. cyclotis* for the purpose of identifying fixed differences among groups and to establish their phylogenetic relationships (Roca et al., 2001). We have exploited and expanded this dataset to characterize the regions encompassing fixed differences among modern elephants in an

* Corresponding author. Fax: +49 (0) 89 31873329.
E-mail address: greenwood@gsf.de (A.D. Greenwood).
[1] These authors contributed equally to the study.

1055-7903/$ - see front matter Published by Elsevier Inc.
doi:10.1016/j.ympev.2006.03.015

effort to better ascertain the relationship of *M. primigenius* to extant elephantids.

2. Materials and methods

2.1. Samples

Two mammoth samples were included in this study. The first, from Engineer Creek, Alaska, has a radiocarbon date of 13,775 ± 145 years before present; nuclear and mitochondrial sequences for this specimen have been verified independently in different laboratories (radiocarbon dating described in Greenwood et al., 1999). Additional sequences have been reported for this mammoth (Binladen et al., 2006; Greenwood et al., 2001). The second sample, from Naskhok River in northeastern Wrangel Island (East Siberian Sea), has been dated to 4050 ± 40 years before present (Beta-195059; d13 corrected). For extant elephantids, our methods of sample collection, DNA extraction, PCR and sequencing have been previously described (Georgiadis et al., 1994; Roca et al., 2001, 2005).

2.2. Ancient DNA (aDNA)

Amplifications and re-amplifications were performed as described by Greenwood et al. (1999). To avoid contamination, processing was carried out in different research institutes: woolly mammoth samples were processed at the Istituto di Medicina Legale (Rome, Italy) while modern elephant samples were analyzed at the Laboratory of Genomic Diversity (Maryland, USA). Each aDNA amplification was performed in duplicate, cloned into a pGEM-T vector (Promega), transformed into electroporation competent bacteria, and five insert-positive clones per amplification sequenced to determine the consensus sequence of the clones (see Supplemental Figures 1–5 for all clone sequences used to derive the consensus sequences used in this study). PCR products ranged from 100 to 180 bp in length. Primer combinations used were, *CHRNA1*: L1 5′ GTTTAGTAGGTTGACTTCCA, R1 5′ GGACTCCATT ATGATCTTTA, L2 5′ GTGATGCACAGCATGAAC AT, R2 5′ AGCAGTTCGAATCCACCAGG, *GBA*: L 5′ GTAACCACTATGCTCCTCA, R 5′ CAGCCCTGAGG ACATCCAC, *BGN*: L1 5′ CTGAGCGCTAGGGCCAT CCA, R1 5′ ATGATGTTGCTGTGCAACA, L2 5′ TCAC ATCCACCAGTACAAAG, R2 5′ GTCTGTTTTAAAG CCTTTCC, *LEPR*: L 5′ TTATGGACTCTATATTGG AG, R 5′ TTGGTTGACCATCTGCAAGT. *VWF* sequences were taken from Greenwood et al. (1999).

2.3. Modern DNA

Genomic DNA (~50 ng) underwent amplification by PCR using 200 nM final concentration of each oligonucleotide primer in 1.5 mM $MgCl_2$, with AmpliTaq-GOLD DNA Polymerase (Applied Biosystems Inc. [ABI]). Primers were as previously reported for *BGN*, *CHRNA1*, and *GBA* (Lyons et al., 1997; Roca et al., 2001, 2005), but rock hyrax (*Procavia capensis*) *BGN* was amplified using new primers BGN-F1f (5′-AAGATCTCCAAGATCCAYGAGAARG) with BGN-R1f (5′-CCCARCCTGTACARCTTGGAGT A). *LEPR* used LEPR-F (5′-CCAAACCTCGAGGAAA GTTTACC) with LEPR-R (5′-AGGCTGCTCCTATGA TACCTCAA) for elephants and LEPR-F2 (5′-GCAGTG TACTGCTGCAATGA) with LEPR-R2 (5′-TGCAAAGT GCTTCCCACA) for hyrax. *VWF* was amplified using either vWF-F1a (5′-GATGGTGTCAACCTCACCTGT) or vWF-L1 (above) with vWF-R1a (5′-CAATGCCCACC GGGATCA); hyrax used vWF-F1a with vWF-R1a. For all primer pairs, PCR consisted of an initial 95 °C for 9:45 min; with cycles of 20 s at 94 °C, followed by 30 s at 60 °C (3 cycles); 58, 56, 54, or 52 °C (5 cycles each temperature); or 50 °C (last 22 cycles), followed by 75 s extension at 72 °C; with a final extension of 3 min at 72 °C. Sequences of several genes had been previously generated for multiple individuals of *E. maximus*, *L. africana*, and *L. cyclotis* (Roca et al., 2001, 2005), while novel elephant, mammoth and hyrax sequences generated for this study have been deposited in GenBank (*BGN*: DQ265804–DQ265820; *CHRNA1*: DQ265821–DQ265838; *GBA*: DQ265839–DQ265855; *LEPR*: DQ265856–DQ265888; *VWF*: DQ265889–DQ265919; Wrangel Island mammoth *BGN*: DQ267154, CHRNA1: DQ267155, DQ267156).

2.4. Phylogenetic analyses

Sequences were aligned using ClustalX (Thompson et al., 1997) and visually inspected. Two datasets were analyzed, each with concatenated DNA sequences from the genes *BGN*, *CHRNA1*, *GBA*, *LEPR* and *VWF*. The first dataset included sequences from elephantids and hyrax; 22 bp of the alignment in the 3′ fragment of *BGN* was excluded due to saturation of the region between hyrax and elephantids. The 3′ fragment of *CHRNA1* was an AfroSINE (Nikaido et al., 2003) present only in each of the elephantids; in hyrax it was coded as gaps and, to maximize resolution within elephantids, the maximum parsimony (MP) analysis treated gaps as a fifth state. The second dataset excluded the hyrax and used only elephantid sequences, including the complete 3′ sequence of *BGN*. In both datasets, a deletion (AAACC) was present in *CHRNA1* in one of the chromosomes (i.e., heterozygous) of elephant DS1534 and both chromosomes (homozygous) of LO3508; the deletion was part of the AfroSINE and removed from the alignment to avoid spurious affinity with hyrax. In a poly-T region of *LEPR* there was deletion of a thymine (in LO3505) or addition of a thymine (in LO3517); in each case the indel was present in only one of the chromosomes (heterozygous), and was not coded for analysis. These indels were present only in forest elephants and would not affect relationships inferred among elephantid genera. Modeltest 3.7 (Posada and Crandall, 1998) was used to determine the Akaike Information Criterion model of DNA sequence evolution that best fit the data; the model was implemented

for Neighbor Joining (NJ) and maximum likelihood (ML) methods in PAUP*4.0b10 (Swofford, 2002); MP was also run. Heuristic searches used 50 replicates of random taxon-addition and tree bisection-reconnection (TBR) branch swapping. Bootstrap resampling support was based on at least 100 replicates, with TBR branch swapping of starting trees obtained by stepwise addition. The model of evolution selected by Modeltest for each dataset was as follows. "Base" indicates the base frequencies for A, C, and G, with T inferred. "Nst" lists the number of substitution types listed in a rate matrix; the number of unique types may be inferred. "Rmat" is the rate matrix. "Rates" indicates the distribution of rates at variable sites. "Pinvar" indicates the proportion of invariant sites. For the elephantids + hyrax dataset: Lset Base = (0.2901 0.2364 0.2236) Nst = 6 Rmat = (1.0000 1.9849 0.3926 0.3926 3.7872) Rates = equal Pinvar = 0. For the elephantids-only dataset: Lset Base = equal Nst = 6 Rmat = (1.0000 1.3818 0.2787 0.2787 3.2657) Rates = equal Pinvar = 0. Tree scores are indicated on the Fig. 1 legend.

A Kishino Hasegawa (KH) test was run in PAUP* (Kishino and Hasegawa, 1989) using the following

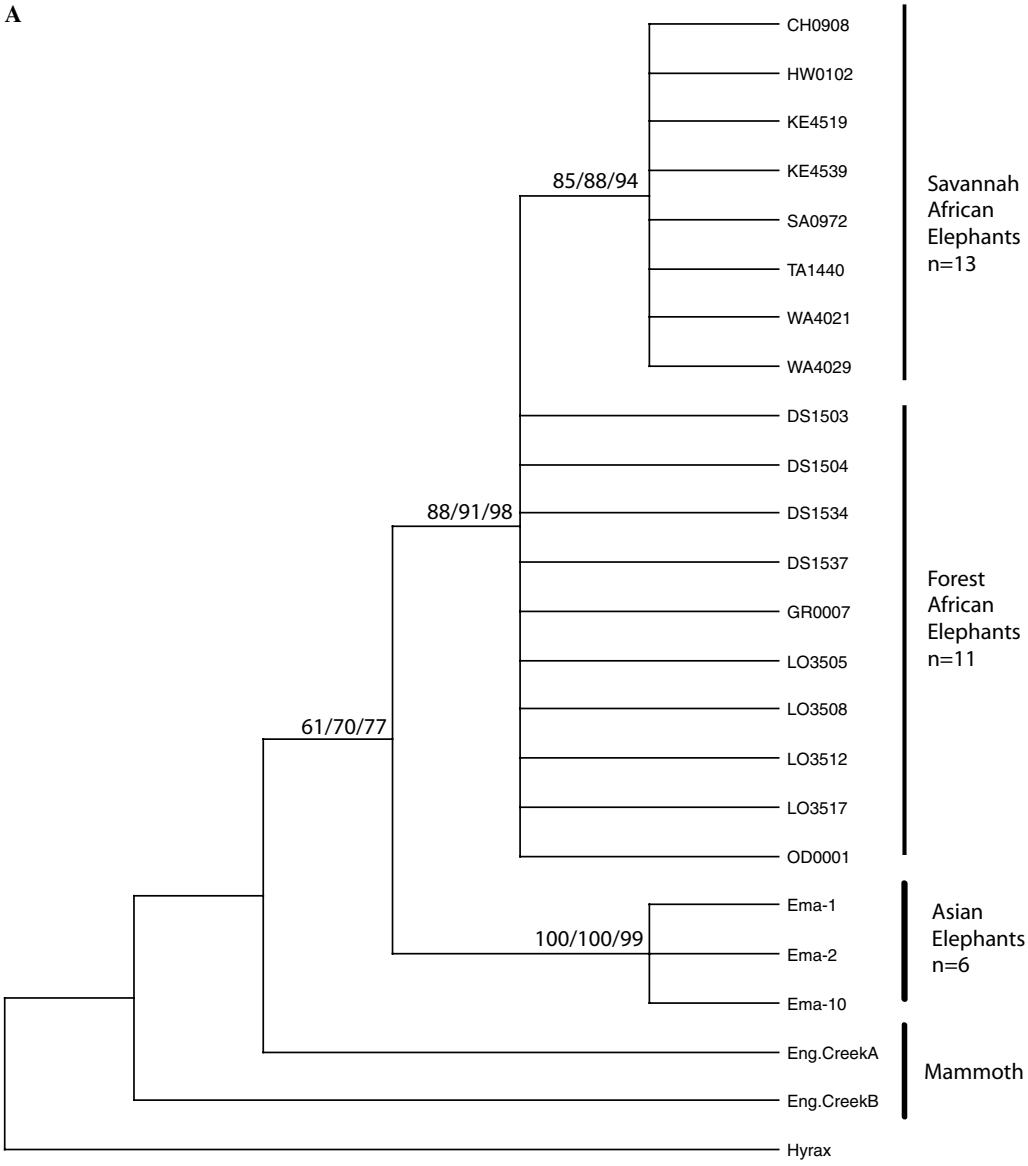

Fig. 1. Phylogenetic trees showing relationships among mammoths and living elephantids using DNA sequences from five nuclear genes (*BGN, CHRNA1, GBA, LEPR, VWF*). For both trees, bootstrap scores above 50% are shown for (left to right) maximum parsimony, Neighbor Joining, and maximum likelihood methods; "ns" indicates less than 50% bootstrap support for a given method. Modern elephant designations are taken from Table 1. Numbers indicated by species labels reflect the presence of additional individuals with duplicate sequences, not listed on the tree but shown in Table 1. (A) Strict consensus of 211,697 equally parsimonious trees produced by maximum parsimony analysis of 701 bp using hyrax as an outgroup, excluding a saturated portion of the 3′*BGN* sequence and treating gaps as a fifth state. The same interspecies relationships were suggested by MP (length 295; CI 0.990; RC 0.950), NJ (ME-score = 0.22282) and ML (-Ln likelihood = 1448.6734) methods. (B) The NJ tree, midpoint rooted for a 677 bp alignment excluding the hyrax sequence. The same interspecies relationships were suggested by MP (number of trees = 1000 [maxtrees], Length 62; CI 1.000; RC 1.000), NJ (ME-score = 0.03703) and ML (-Ln likelihood = 0.03758) methods.

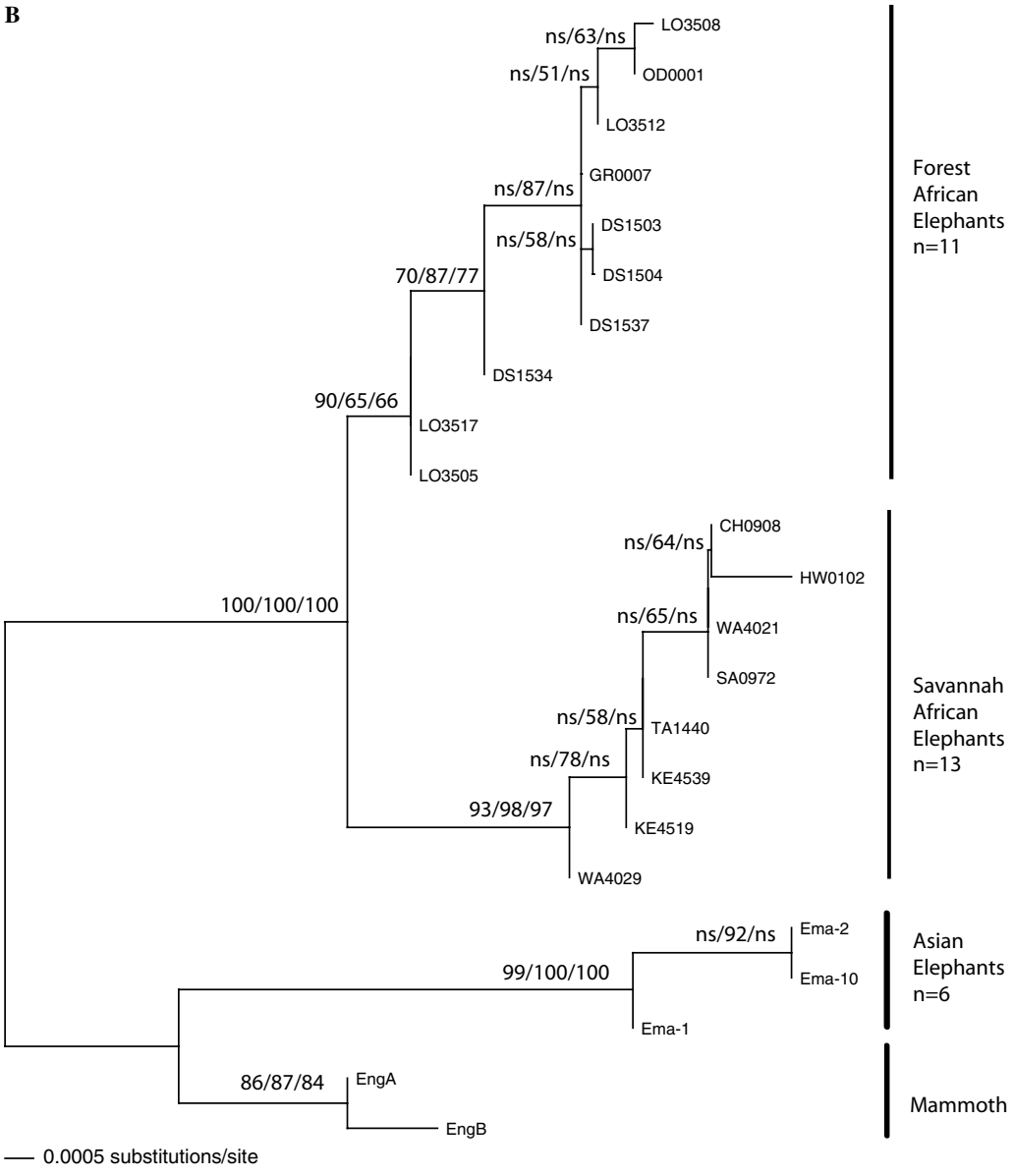

Fig. 1. (*continued*)

tree based on MP analysis of the dataset: (((((((CH0809, HW0102,SA0972,WA4021),KE4519,KE4539,TA1440, WA4029),(DS1503,DS1504,DS1537,GR0007,(LO3508, OD0001),LO3512),DS1534),LO3517),LO3505),(Ema-1, (Ema-2, Ema-10))),(Eng.CreekA,Eng.CreekB)),hyrax). This tree was compared to two other trees, both with the same intra-generic but different inter-generic relationships among the individuals. In the first tree used in both KH tests, *Loxodonta* and *Elephas* formed a clade excluding *Mammuthus*; in one comparison tree *Loxodonta* and *Mammuthus* were grouped to form a clade excluding *Elephas*, while in the other *Elephas* and *Mammuthus* formed a clade excluding *Loxodonta*.

Minimum spanning tree analysis was performed for the aligned elephantid sequences (without hyrax) using the TCS program (Clement et al., 2000).

3. Results

Two individual mammoths were genotyped at multiple nuclear DNA loci chosen for the potential presence of fixed nucleotide differences between *Elephas* and *Loxodonta*. A total of 681 bp of mammoth sequence was determined for loci *BGN* (175 bp), *CHRNA1* (193 bp), *GBA* (62 bp), *LEPR* (137 bp), and *VWF* (114 bp), with sequences for *BGN*, *CHRNA1*, and *VWF* amplified in two non-overlapping fragments. The mammoths were from different continents (Wrangel Island in northeastern Asia and Engineer Creek in Alaska) and chronologically separated by thousands of years. Thus, recovered sequences are likely to be minimally representative of geographic variation among mammoths for the loci characterized. In addition, little variation among mammoths has been

observed for cyt *b* (Debruyne et al., 2003). However, for nuclear loci the scale of variation would have to be determined by examining the sequences from additional mammoths.

The Wrangel Island mammoth specimen yielded lower quality DNA than the Alaskan sample and only produced replicable sequence for one *BGN* fragment and both *CHRNA1* fragments. It also yielded sequence in a single attempt to amplify *GBA*. For all fragments in common between the two mammoths, sequences were identical. For subsequent phylogenetic and network analyses, the Alaskan mammoth sequence was used.

In addition to the two mammoths, samples from six Asian elephants, 11 African forest elephants, 13 African savannah elephants, and one hyrax were characterized for all loci (Table 1 and Supplementary Table 1). Attempts to amplify the same genes in a manatee were unsuccessful. Both variable and fixed among-species elephantid differences were identified, with the mammoths exhibiting three unique polymorphisms, Asian elephants eight, African forest elephants ten, and African savannah elephants six (excluding indels). Transitions outnumbered transversions and indels were present in three of the genes including a homozygous 5 bp deletion in forest elephant LO3508 (Table 1). A single West African forest elephant, SL0001 from Sierra Leone, proved identical in combined gene sequences to LO3517 from Gabon in Central Africa (Table 1). This fails to confirm the suggestion of Eggert and colleagues (2002), based on mtDNA, that West African elephants comprise a separate species distinct from *L. africana* and *L. cyclotis*, although more specimens and nuclear markers would be required to confirm our result.

A 22 bp segment of the 3′ *BGN* elephantid alignment could not be aligned to hyrax sequence due to saturation and was removed. The remaining hyrax–elephantid alignment was 701 bp in length (21 parsimony informative sites), which included a 3′ *CHRNA1* fragment coded as gaps in the hyrax since it comprised part of an AfroSINE insertion present only in the elephantids (Nikaido et al., 2003). Phylogenetic analysis using hyrax as the outgroup suggested that *M. primigenius* is the most primitive elephantid with a subsequent branching of *Elephas* and the two *Loxodonta* species (Fig. 1A and Supplementary Figure 7). However, there are several

Table 1
Variable sites in nuclear genes among elephantids

		BGN 5′				BGN 3′							CHRNA1 5′					CHRNA1 3′					GBA		LEPR								VWF 5′				VWF 3′		
		19	20	49	83	22	24	30	31	32	86	5	26	56	63	82	16	67	70	103	10	60	3	19	58	72	107	128	131	6	12	24	47	9	13	27
Mammoth	Eng.CreekA	C	C	G	T	C	G	G	A	G	G	T	G	G	T	T	T	AAACC	A	A	C	C	C	C	G	T	A	A	A	G	A	C	G	C	A	A	G
	Eng.CreekB	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.
Asian Elephants	Ema-1	.	.	.	.	.	T	.	.	.	.	.	C/G	A	C	.	C/T	.	G	.	T	.	.	T	.	.	C	.	.	.	.	.	.	.	.	.	.	.
	Ema-2	.	.	.	.	.	T	.	.	.	.	.	C	A	C	.	C	.	G	.	T	.	.	T	.	.	C	.	.	.	.	.	.	.	.	.	.	.
	Ema-6	.	.	.	.	.	T	.	.	.	.	.	C/G	A	C	.	C/T	.	G	.	T	.	.	T	.	.	C	.	.	.	.	.	.	.	.	.	.	.
	Ema-7	.	.	.	.	.	T	.	.	.	.	.	C/G	A	C	.	C/T	.	G	.	T	.	.	T	.	.	C	.	.	.	.	.	.	.	.	.	.	.
	Ema-9	.	.	.	.	.	T	.	.	.	.	.	C	A	C	.	C	.	G	.	T	.	.	T	.	.	C	.	.	.	.	.	.	.	.	.	.	.
	Ema-10	.	.	.	.	.	T	.	.	.	.	.	C	A	C	.	C	.	G	.	T	.	.	T	.	.	C	.	.	.	A/G	.	.	.	.	.	.	.
Forest African Elephants	DS1503	A	A	A	C	.	.	T	G	-	-	.	.	.	C	.	.	.	G	.	.	.	.	.	.	.	.	C	.	T	.	.	.	.	.	T	.
	DS1504	A	A	A	C	.	.	T	G	-	-	.	.	.	C	.	.	.	G	.	.	.	C/T	.	.	.	.	C	.	T	.	.	.	.	.	T	.
	DS1534	A/C	A/C	A	C	.	.	T	G	-	-	.	.	.	C	.	.	.	G	.	.	.	C/T	.	.	C/T	.	C	.	T	.	.	.	.	A/G	T	A/G
	DS1537	A	A	A	C	.	.	T	G	-	-	.	.	.	C	.	.	.	G	.	.	.	.	.	.	C/T	.	A/C	.	T	.	.	.	.	.	T	.
	LO3505	.	.	A	C	.	.	T	G	-	-	.	.	.	C	.	.	.	G	.	.	.	.	.	.	C/T	.	A/C	.	T	.	.	.	.	.	T	.
	LO3508	A	A	A	C	.	.	T	G	-	-	.	.	.	C	.	.	-----	G	.	.	.	.	.	.	C	.	C	.	T	.	.	.	.	.	T	.
	LO3512	A	A	A	C	.	.	T	G	-	-	.	.	.	C	.	.	.	G	.	.	.	C/T	.	.	C/T	.	C	.	T	.	.	.	.	.	T	.
	LO3517	.	.	A	C	.	.	T	G	-	-	.	.	.	C	.	.	.	G	.	.	.	C/T	.	.	C/T	.	A/C	.	T	.	.	.	.	.	T	.
	OD0001	A	A	A	C	.	.	T	G	-	-	.	.	.	C	C/T	.	.	G	.	.	.	.	.	.	C	.	A/C	A/G	T	.	.	.	.	.	T	.
	SL0001	.	.	A	C	.	.	T	G	-	-	.	.	.	C	.	.	.	G	.	.	.	C/T	.	.	C/T	.	A/C	.	T	.	.	.	.	.	T	.
	GR0007	A	A	A	C	.	.	T	G	-	-	.	.	.	C	.	.	.	G	.	.	.	C/T	.	A/G	C/T	.	A/C	.	T	.	.	.	.	A/G	T	A/G
Savannah African Elephants	CH0908	.	.	A	C	T	.	.	G	.	-	C	.	.	C	.	.	.	G	A/G	.	T	.	.	.	.	.	C	.	T	.	.	.	T	.	T	.
	HW0102	.	.	A	C	T	.	.	G	.	-	C	.	.	C	.	.	.	G	G	.	T	.	.	.	.	.	C	.	T	.	.	.	T	.	T	.
	KE4519	.	.	A	C	T	.	.	G	.	-	C	.	.	C	.	.	.	G	.	.	T	.	.	.	.	.	C	.	T	.	.	.	.	.	T	.
	KE4539	.	.	A	C	T	.	.	G	.	-	C	.	.	C	.	.	.	G	.	.	T	.	.	.	.	.	C	.	T	.	.	.	C/T	.	T	.
	KR0014	.	.	A	C	T	.	.	G	.	-	C	.	.	C	.	.	.	G	.	.	T	.	.	.	.	.	C	.	T	.	.	.	C/T	.	T	.
	KR0138	.	.	A	C	T	.	.	G	.	-	C	.	.	C	.	.	.	G	G	.	T	.	.	.	.	.	C	.	T	.	.	.	T	.	T	.
	NA4721	.	.	A	C	T	.	.	G	.	-	C	.	.	C	.	.	.	G	.	.	T	.	.	.	.	.	C	.	T	.	.	.	C/T	.	T	.
	SA0972	.	.	A	C	T	.	.	G	.	-	C	.	.	C	.	.	.	G	.	.	T	.	.	.	.	.	C	.	T	.	.	.	T	.	T	.
	SE2100	.	.	A	C	T	.	.	G	.	-	C	.	.	C	.	.	.	G	.	.	T	.	.	.	.	.	C	.	T	.	.	.	C/T	.	T	.
	TA1440	.	.	A	C	T	.	.	G	.	-	C	.	.	C	.	.	.	G	.	.	T	.	.	.	.	.	C	.	T	.	.	A/G	C/T	.	T	.
	WA4020	.	.	A	C	T	.	.	G	.	-	C	.	.	C	.	.	.	G	.	.	T	.	.	.	.	.	C	.	T	.	.	.	.	.	T	.
	WA4021	.	.	A	C	T	.	.	G	.	-	C	.	.	C	.	.	.	G	.	.	T	.	.	.	.	.	C	.	T	.	.	A/G	T	.	T	.
	WA4029	.	.	A	C	T	.	.	G	.	-	C	.	.	C	.	.	.	G	.	.	C/T	.	.	.	.	.	C	.	T	.	.	.	.	.	T	.
Hyrax	Pca-1	.	.	A	.	.	?	?	?	?	?	.	.	.	.	.	.	-	-	-	-	T	.	.	A	.	.	G	.	.	.	T	A	T	.	.	.

Eng.CreekA and B refer to the two alleles found in this specimen for VWF (Greenwood et al., 1999). IUPAC designations for bases are shown,"." represents identity to the reference sequence and "-" represent deletions or gaps. Positions saturated when aligned to the hyrax are indicated by a question mark "?". Mammoth character states not present in any elephants are shaded grey. The numbering begins from the first base after the 5′ primer for each mammoth sequence; for BGN, CHRNA1, and VWF, the 5′ and 3′ sequences are separated by a line and numbered separately. Modern elephant sample designations are taken from Roca et al., 2001. Localities for the *Loxodonta* samples are: DS-Dzanga Sangha in Central African Republic; GR-Garamba in Congo (Kinshasa); LO-Lope in Gabon; OD-Odzala in Congo (Brazzaville); SL-Sierra Leone; CH-Chobe and SA-Savuti in Botswana; HW-Hwange in Zimbabwe; KE-Central Kenya; KR-Kruger in South Africa; NA-Namibia; SE-Serengeti and TA-Tarangire in Tanzania; WA-Waza in Cameroon. Asian elephants were from zoos; those of known geographic origin derived from India (Ema-1), Sri Lanka (Ema-2 & 9) and Thailand (Ema-10).

reasons to be cautious with this conclusion. First, the removal of part of *BGN* from the analysis removes several sites variable among the elephantids. Second, the branch length leading from the outgoup is extremely long; in the analysis by Springer et al. (2005), which notably does not support the traditional definition of Tethytheria, the base of the paeungulate divergence (elephants (sirenians, hyraxes)) is dated to approximately 63 Ma. Very long outgroup branch lengths have been a difficulty for mtDNA based phylogenies of elephantids and are clearly a problem for nuclear DNA based analysis. As the more recent lineages having a common ancestry with extant elephants are all extinct, it remains to be seen if an appropriate outgroup species will allow for nuclear DNA study. A Kishino Hasegawa (KH) test compared the tree with the relationships suggested by MP analysis of the dataset, in which *Elephas* and *Loxodonta* form a clade excluding *Mammuthus* (Fig. 1A), to alternative trees in which intra-generic relationships were maintained but inter-generic relationships were altered. The KH test found that support for a tree with the *Elephas–Loxodonta* clade (Fig. 1A) was not significantly different from support for a tree with a *Loxodonta–Mammuthus* clade ($p = 0.32$) or support for a tree with an *Elephas–Mammuthus* clade ($p = 1.00$).

The analysis with hyrax had excluded a region of the *BGN* sequence that was saturated between elephantids and hyrax. To include the full elephantid sequence for *BGN*, a second analysis was run that excluded the hyrax and used all of the elephantid *BGN* sequence, along with the four other gene sequences, in an alignment 677 bp long (25 parsimony informative sites). The tree was mid-point rooted (Fig. 1B). The results demonstrated the expected separation of *L. cyclotis* and *L. africana* (Roca et al., 2001). Although the tree appears to suggest a slightly closer relationship of *Elephas* and *Mammuthus*, this interpretation should be treated with caution given the paucity of informative sites, the lack of an appropriate outgroup, and the possibility that some lineages may be accelerated. Nonetheless, a *Mammuthus–Loxodonta* association was not suggested by either analysis, in contrast to several mtDNA studies.

Network analysis could not distinguish between a closer association of woolly mammoths and Asian elephants or African elephants, with nine steps separating *Mammuthus* from the nearest *Elephas* individual, versus ten steps to the nearest *Loxondonta*. In particular, it is of interest to note (cf. Roca et al., 2001) that the distance between *L. cyclotis* and *L. africana* is high relative to the difference between either of these species and Asian elephants or woolly mammoths (4–13 steps, including indels, between *L. cyclotis* and *L. africana* versus 10–19 steps between *L. cyclotis* and *M. primigenius*, Fig. 2). The same analysis with gaps

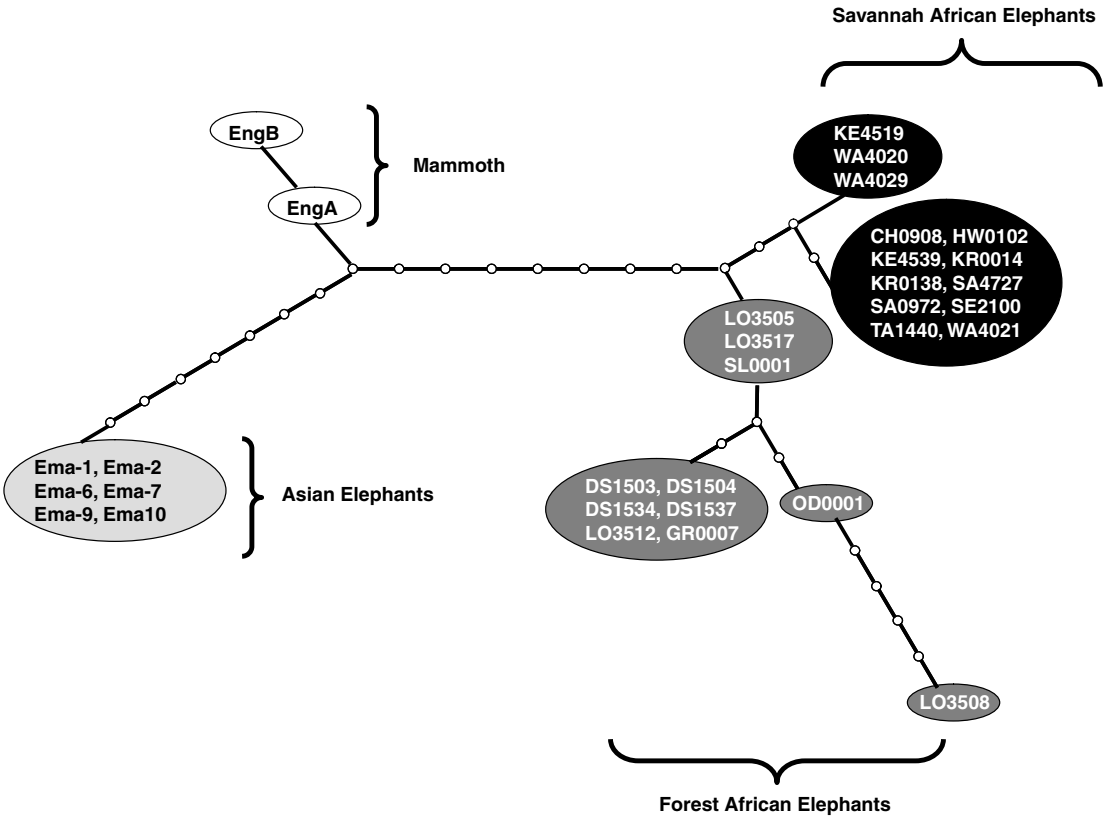

Fig. 2. Minimum spanning network of elephantid sequences. Open circles indicate the number of substitutions between nodes. Modern elephant designations are as in Table 1. The network depicted includes gaps as a fifth state. Heterozygous positions (Table 1) were scored as unknown. For example, position 107 in the *LEPR* gene was a C (cytosine) in all but six African elephants; these six individuals were heterozygous (C and A). Thus, some elephant individuals carried two different haplotypes with varying degrees of distance to sequences from other animals.

excluded yielded a similar network but reduced the number of steps along some branches (for example L03508 would be the same as OD0001; see also Table 1). The distance and diversity exhibited by *L. cyclotis* reflects a long history of reproductive isolation from *L. africana*.

4. Discussion

Although a small number of sites uniquely group woolly mammoths and Asian elephants, the phylogeny of the Elephantidae could not be resolved with the current dataset. However, the trend does not suggest a strong *Mammuthus–Loxodonta* association as has been reported in several mtDNA based studies (Greenwood et al., 1999). By contrast, the *VWF* gene suggests a mammoth–*Elephas* association, as does the *BGN* gene. While *GBA* is ambiguous, *CHRNA1* favors a mammoth–*Loxodonta* association and *LEPR* slightly favors *Loxodonta–Mammuthus* as there are heterozygous forest elephant individuals with only one difference compared to mammoth while Asian elephants uniformly display two fixed *LEPR* differences versus the mammoth sequence. Nonetheless, none of our analyses combining all the sequences produced a *Mammuthus–Loxodonta* grouping.

The three elephantid genera radiated in quick succession in the late Miocene/early Pliocene (Maglio, 1973; Vignaud et al., 2002). Their evolutionary patterns may be comparable to that produced by the contemporaneous rapid radiation of the gorilla, chimpanzee and human lineages, in which the correct (gorilla (human, chimpanzee)) relationship is supported by only 60% of nuclear loci and phylogenetically informative sites, due to random sorting, recombination, genetic drift or homoplasy (O'hUigin et al., 2002; Satta et al., 2000). An added difficulty for interpreting elephantid relationships is that one target group is extinct. Lack of an appropriate outgroup sequence is another difficulty. Hyracoids and sirenians are the groups most closely related to proboscideans, but since their divergences occurred at the beginning of the Cenozoic 63 Ma, they are poor candidates for determining among-species branching patterns. Although mtDNA sequences have been reported for the mastodon (*Mammut americanum*), the results have not been independently replicated and nuclear DNA has never been retrieved from a mastodon (Yang et al., 1996).

Nonetheless, the results of this study suggest that further sequencing of woolly mammoth nuclear genes should resolve their phylogeny conclusively, although it will require a substantial increase in the number of informative sites and independent loci examined. Recent developments in sequencing technology suggest that whole genome analysis of extinct animals, particularly mammoths will be feasible (Poinar et al., 2006). We also conclude that the application of nuclear markers is now practicable and indeed preferable for systematic study of a wide variety of extinct animals represented by well-preserved remains in museum collections.

Acknowledgments

The authors thank Lars Giesen and Uwe Kohler (both of Medigenomix GmbH, Martinsried, Germany) for technical support. We are grateful to Claudia Englbrecht for critically reading the manuscript. We thank A. Brandt, S. Rosendale, and S. Mordensky for assistance. For elephant samples, we thank A. Turkalo, J.M. Fay, R. Weladji, W. Karesh, M. Lindeque, W. Versvelt, K. Hillman Smith, F. Smith, M. Tchamba, S. Gartlan, P. Aarhaug, A.M. Austmyr, Bakari, Jibrila, J. Pelleteret, L. White, M. Habibou, M.W. Beskreo, D. Pierre, C. Tutin, M. Fernandez, R. Barnes, B. Powell, G. Doungoubé, M. Storey, M. Phillips, B. Mwasaga, A. Mackanga-Missandzou, M. Keele, D. Olson, B. York, and A. Baker at the Burnet Park Zoo, M. Bush at the National Zoological Park, and A. Lécu at Zoo de Vincennes (Paris Zoo). We thank the governments of Botswana, Cameroon, the Central African Republic, Congo (Brazzaville), Congo (Kinshasa), Gabon, Kenya, Namibia, South Africa, Tanzania, and Zimbabwe for permission to collect samples. Samples were obtained in full compliance with specific Federal Fish and Wildlife Permits (endangered/threatened species and CITES Permits US 750138 and US 756611 to N.G.). For funding, we thank R. Ruggiero and the US Fish and Wildlife Service African Elephant Conservation Fund; and the National Geographic Society and European Union (through the Wildlife Conservation Society). This publication has been funded in part with Federal funds from the National Cancer Institute, National Institutes of Health, under Contract No. N01-CO-12400. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the US Government. This research was supported in part by the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research. The work was financially supported by grants from the National Science Foundation (OPP 0117400), the Niarchos Fund, and the Evelyn Stefansson Nef Foundation.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ympev.2006.03.015.

References

Binladen, J., Wiuf, C., Gilbert, M.T., Bunce, M., Barnett, R., Larson, G., Greenwood, A.D., Haile, J., Ho, S.Y., Hansen, A.J., Willerslev, E., 2006. Assessing the fidelity of ancient DNA sequences amplified from nuclear genes. Genetics 172, 733–741.

Clement, M., Posada, D., Crandall, K.A., 2000. TCS: a computer program to estimate gene genealogies. Mol. Ecol. 9, 1657–1659.

Debruyne, R., Barriel, V., Tassy, P., 2003. Mitochondrial cytochrome b of the Lyakhov mammoth (Proboscidea, Mammalia): new data and phylogenetic analyses of Elephantidae. Mol. Phylogenet. Evol. 26 (3), 421–434.

Eggert, L.S., Rasner, C.A., Woodruff, D.S., 2002. The evolution and phylogeography of the African elephant inferred from mitochondrial DNA sequence and nuclear microsatellite markers. Proc. R. Soc. Lond. B Biol. Sci. 269, 1993–2006.

Georgiadis, N., Bischof, L., et al., 1994. Structure and history of African elephant populations: I. Eastern and southern Africa. J. Hered. 85, 100–104.

Greenwood, A.D., 2001. Mammoth biology: biomolecules, phylogeny, Numts, nuclear DNA, and the biology of an extinct species. Anc. Biomol. 3, 255–266.

Greenwood, A.D., Pääbo, S., 1999. Nuclear insertion sequences of mitochondrial DNA predominate in hair but not in blood of elephants. Mol. Ecol. 8, 133–137.

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

Greenwood, A.D., Lee, F., Capelli, C., Desalle, R., Tikhonov, A.N., Marx, P.A., MacPhee, R.D.E., 2001. Evolution of endogenous retrovirus-like elements of the woolly mammoth (*Mammuthus primigenius*) and its relatives. Mol. Biol. Evol. 18, 840–847.

Kishino, H., Hasegawa, M., 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in hominoidea. J. Mol. Evol. 29, 170–179.

Krause, J., Dear, P.H., Pollack, J.L., Slatkin, M., Spriggs, H., Barnes, I., Lister, A.M., Ebersberger, I., Paabo, S., Hofreiter, M., 2006. Multiplex amplification of the mammoth mitochondrial genome and the evolution of Elephantidae. Nature 439, 724–727.

Lyons, L.A., Laughlin, T.F., Copeland, N.G., Jenkins, N.A., Womack, J.E., O'Brien, S.J., 1997. Comparative anchor tagged sequences (CATS) for integrative mapping of mammalian genomes. Nat. Genet. 15, 47–56.

Maglio, V.J., 1973. Origin and evolution of the Elephantidae. Trans. Am. Phil. Soc. Philad., New Series 63, 1–149.

Nikaido, M., Nishihara, H., Hukumoto, Y., Okada, N., 2003. Ancient SINEs from African endemic mammals. Mol. Biol. Evol. 20 (4), 522–527.

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 (3), 314–326.

O'hUigin, C., Satta, Y., Takahata, N., Klein, J., 2002. Contribution of homoplasy and of ancestral polymorphism to the evolution of genes in anthropoid primates. Mol. Biol. Evol. 19, 1501–1513.

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 (4), 406–413.

Poinar, H., Kuch, M., McDonald, G., Martin, P., Pääbo, S., 2003. Nuclear gene sequences from a late Pleistocene sloth coprolite. Curr. Biol. 13, 1150–1152.

Poinar, H.N., Schwarz, C., Qi, J., Shapiro, B., Macphee, R.D., Buigues, B., Tikhonov, A., Huson, D.H., Tomsho, L.P., Auch, A., Rampp, M., Miller, W., Schuster, S.C., 2006. Metagenomics to paleogenomics: large-scale sequencing of mammoth DNA. Science 311, 392–394.

Posada, D., Crandall, K.A., 1998. Modeltest: testing the model of DNA substitution. Bioinformatics 14 (9), 817–818.

Roca, A.L., Georgiadis, N., Pecon-Slattery, J., O'Brien, S.J., 2001. Genetic evidence for two species of elephant in Africa. Science 293, 1473–1477.

Roca, A.L., Georgiadis, N., O'Brien, S.J., 2005. Cytonuclear genomic dissociation in African elephant species. Nat. Genet. 37 (1), 96–100.

Rogaev, E.I., Moliaka, Y.K., Malyarchuk, B.A., Kondrashov, F.A., Derenko, M.V., Chumakov, I., Grigorenko, A.P., 2006. Complete mitochondrial genome and phylogeny of Pleistocene mammoth *Mammuthus primigenius*. PLoS Biol. 4 (3), e73.

Satta, Y., Klein, J., Takahata, N., 2000. DNA archives and our nearest relative: the trichotomy problem revisited. Mol. Phylogenet. Evol. 14, 259–275.

Springer, M., Murphy, W., Eizirik, E., O'Brien, S.J., 2005. In: Rose, K., Archibald, J.D. (Eds.), The Rise of Placental Mammals: Origins and Relationships of the Major Extant Clades. Johns Hopkins Univ. Press, Baltimore, pp. 37–49.

Swofford, D.L., 2002. PAUP*: Phylogenetic Analysis Using Parsimony (* and Other Methods). Version 4.0b10. Sunderland, Massachusetts, Sinauer.

Thalmann, O., Hebler, J., Poinar, H.N., Pääbo, S., Vigilant, L., 2004. Unreliable mtDNA data due to nuclear insertions: a cautionary tale from analysis of humans and other great apes. Mol. Ecol. 13, 321–335.

Thompson, J.D., Gibson, T.J., et al., 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25, 4876–4882.

Vignaud, P., Duringer, P., Mackaye, H.T., Likius, A., Blondel, C., Boisserie, J.R., De Bonis, L., Eisenmann, V., Etienne, M.E., Geraads, D., Guy, F., Lehmann, T., Lihoreau, F., Lopez-Martinez, N., Mourer-Chauvire, C., Otero, O., Rage, J.C., Schuster, M., Viriot, L., Zazzo, A., Brunet, M., 2002. Geology and palaeontology of the Upper Miocene Toros-Menalla hominid locality, Chad. Nature 418, 152–155.

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 (3), 1190–1194.