Genetic Insight into an Extinct Population of Asian Elephants (*Elephas maximus*) in the Near East

Linus Girdland-Flink*,†, Ebru Albayrak‡ and Adrian M. Lister*

The current range of the Asian elephant is fragmented and restricted to southern Asia. Its historical range was far wider and extended from Anatolia and the Levant to Central China. The fossil record from these peripheral populations is scant and we know little of their relationship to modern Asian elephants. To gain a first insight to the genetic affinity of an *E. maximus* population that once inhabited Turkey we sequenced ca. 570 bp mtDNA from four individuals dating to ~3500 cal. BP. We show that these elephants carried a rare haplotype previously only observed in one modern elephant from Thailand. These results clarify the taxonomic identity of specimens with indeterminate morphologies and show that this ancient population groups within extant genetic variation. By placing the age of the common ancestor of this haplotype in the interval 3.7–58.7 kya (mean = 23.5 kya) we show that range-wide connectivity occurred at some time or times since the start of MIS 3, ~57 kya, probably reflecting range and population expansion during a favourable climatic episode. The genetic data do not distinguish natural versus anthropogenic origin of the Near Eastern Bronze Age population, but together with archaeological and paleoclimatic data they allow the possibility of a natural westward expansion around that time.

**Keywords:** Asian elephant; *Elephas maximus*; ancient DNA

---

1. Introduction

The Asian elephant (*Elephas maximus*, L. 1758) is an endangered species whose current distribution extends from the Indian subcontinent to Borneo ([Figure 1A](#)). Its range was far wider during the Pleistocene and early Holocene, extending from Anatolia and the Levant in the west, along a narrow land corridor by the southern Asian coastline, to the Indian subcontinent and as far eastward as central China (Santiapillai and Jackson 1990, Sukumar 2011) ([Figure 1A](#)). Since the fossil record is scant, often comprising single fragmented teeth or long bones that can be confused with other elephantid species, their affinities and distributional change are poorly understood (Turvey et al. 2013). Another possibly complicating factor is human translocation of living elephants from their core distribution (e.g. India) as part of emerging trade networks and long-range war campaigns (Collon 1977, Albenda 2008).

To gain insight to the *E. maximus* population that once inhabited the Near East we sequenced mitochondrial DNA (mtDNA) extracted from four ancient teeth excavated at Kahramanmaraş (Gavur Lake Swamp) in south-central Anatolia ([Figure 1A, Table 1](#)). The elephant remains have been directly 14C dated to approximately 3500 BP (ca. 1500 BCE) and represent one of the largest assemblages of Holocene *E. maximus* in the region (Albayrak and Lister 2012). The dates place the remains within the Late Bronze Age, when the region around Kahramanmaraş was likely an area of influence of both the Mittani and Hittite Kingdoms (Alaura 2016). The collection is primarily made up of loose teeth but also includes partial skulls and postcranial elements ([Figure 2](#)). Although most specimens show strong morphological affinity to modern *E. maximus*, a number of molars could not be identified with certainty as *E. maximus* compared to the Pleistocene species *Palaeoloxodon antiquus* (straight-tusked elephant).

Here we demonstrate the phylogenetic position of the elephant teeth (determinate and indeterminate) in relation to the known mitochondrial phylogeography of modern Asian elephant populations (Vidya, Sukumar and Melnick 2009). Further, by estimating mtDNA coalescence times we discuss the position of the Turkish remains in the broader historical context of *E. maximus*.

2. Methods

We extracted DNA from nine elephantid teeth from Kahramanmaraş and sequenced 579 bp mitochondrial DNA, including 574 bp of the fragment analysed in Vidya, Sukumar and Melnick (2009), comprising the C-terminal of cyt-b, t-RNA<sub>Thr</sub>, t-RNA<sub>Pro</sub> and the hypervariable left
domain of the control region. The sampling, DNA extraction and all pre-PCR work were carried out in a dedicated ancient DNA laboratory at the Natural History Museum, London, following strict protocols and recommendations (Gilbert et al. 2005). No work on other elephantid DNA had taken place previously in the ancient DNA laboratory.

2.1. DNA extraction

Teeth chosen for DNA extraction were decontaminated first by applying a 1% dilute solution of bleach (sodium hypochlorite) on and around the surface area targeted for dentine excision, followed by UV irradiation at 254 nm for 20 min at a distance of <10 cm. The outer surface

Table 1: A table summarizing the specimens analysed in this study, from Albayrak and Lister (2012). M = upper molar, m = lower molar.

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>NHM ID</th>
<th>Species (morphology)</th>
<th>Element</th>
<th>DNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>46GB05</td>
<td>EL001</td>
<td>cf. Elephas maximus</td>
<td>m3</td>
<td>No</td>
</tr>
<tr>
<td>46GB04</td>
<td>EL002</td>
<td>Elephas maximus</td>
<td>M3</td>
<td>No</td>
</tr>
<tr>
<td>46GB0X</td>
<td>EL003</td>
<td>Elephas maximus</td>
<td>M3</td>
<td>No</td>
</tr>
<tr>
<td>46GB02</td>
<td>EL004</td>
<td>Elephas maximus</td>
<td>M3</td>
<td>Yes</td>
</tr>
<tr>
<td>46GB20</td>
<td>EL005</td>
<td>cf. Elephas maximus</td>
<td>M2</td>
<td>No</td>
</tr>
<tr>
<td>46GB01</td>
<td>EL006</td>
<td>cf. Elephas maximus</td>
<td>m3</td>
<td>Yes</td>
</tr>
<tr>
<td>46GB09</td>
<td>EL007</td>
<td>indet.</td>
<td>m2</td>
<td>Yes</td>
</tr>
<tr>
<td>46GB03</td>
<td>EL008</td>
<td>indet.</td>
<td>m3</td>
<td>Yes</td>
</tr>
<tr>
<td>46GB07</td>
<td>EL009</td>
<td>cf. Elephas maximus</td>
<td>m1</td>
<td>No</td>
</tr>
</tbody>
</table>
Figure 2: A partially reconstructed skull of Elephas maximus from Gavur Lake Swamp (MTA Natural History Museum, Ankara).
(~1 mm) was removed by abrasion and tooth powder was obtained using a Dremel drill at the lowest possible rpm setting. We obtained a total of 30–60 mg powder per specimen.

DNA was extracted and purified following a modified version of a silica-binding protocol (Yang et al. 1998, Ottoni et al. 2013). Tooth powder was suspended and digested overnight at constant rotation at 50°C in 2 mL of extraction buffer consisting of 0.425 M EDTA (pH8), 1 mM Tris–HCl (pH8), 0.05% w/v SDS, and 0.33 mg/ml Proteinase K. The digested solution was concentrated to ≤100 μL using Amicon Ultra-4 centrifugal filter units with a molecular weight cut-off of 30 kDa (Amicon® Ultra, Millipore). We isolated and purified the DNA on silica columns (QIAquick® PCR purification kit, Qiagen) following the manufacturer's protocol and eluted in 100 μL EB buffer. One negative extraction control was processed alongside the nine tooth samples.

### 2.2. PCR amplification

A series of novel PCR primers were designed in Primer3Plus (Untergasser et al. 2012) to amplify short (101–119 bp) and mostly overlapping mtDNA fragments (Table 2). Singleplex PCRs were carried out in 25 μL reactions using 1.0 U Smart-Taq Hot DNA Polymerase and 1X Smart-Taq Buffer with (NH₄)₂SO₄ (Naxo), 2.5 mM MgCl₂, 200 μM of each dNTP, 0.5 mg/ml final concentration of RSA, and 0.2 μM of each primer. The cycling conditions were 95°C for 15 min followed by 55 cycles of 94°C for 20 sec, 55–57°C for 20 sec, 72°C for 20 sec, and a final extension for 5 min at 72°C. We obtained two PCR amplified products/fragment/specimen (amplified at different times) and included one negative PCR control for every eight PCRs.

### 2.3. DNA sequencing

PCR products were purified using the QIAquick® PCR purification kit (Qiagen), pooled into equimolar ratios following quantification on a Qubit® (Invitrogen), and constructed into barcoded sequencing libraries using the TruSeq Nano DNA sample prep kit (v2, Illumina) following manufacturer’s recommendations (but omitting initial shearing) at the sequencing facility at the Natural History Museum, London. Sequencing was performed in-house on the Illumina MiSeq platform the using MiSeq Reagent Kits v2 at ~10% of a full sequencing run. Libraries were de-multiplexed on the MiSeq platform and exported in FastQ format. FastQ files were trimmed in Geneious V7 (Kearse et al. 2012) first by removing regions with more than a 5% chance of an error per base, and secondly by retaining regions with no more than 2 bases with a quality of ≤30. Sequences were aligned against a reference *E. maximus* sequence (Genbank accession AY1245813), and consensus sequences were called using majority consensus and highest cumulative quality scores. Bases with a quality of ≤30 were ignored and called as Ns. Base coverage ranged from ca. 1000–30,000.

### 2.4. Phylogenetic analysis

We downloaded reference data from two *Loxodonta cyclotis* sequences (JQ438507 and KJ557424), two *Loxodonta africana* sequences (JQ438674 and JQ438767) (Ishida et al. 2013, Finch et al. 2014), as well as two *Mammuthus primigenius* sequences (GU984769 and KC427898) (Nyström et al. 2010, Palkopoulou et al. 2013) and aligned them with the Kahramanmaras sequences using MAFFT v7.017 (Katoh et al. 2002). We then aligned and pruned the DNA sequences to previously published *Elephas maximus* data (Fernando et al. 2000, Fernando et al. 2003, Vidya, Sukumar and Melnick 2009). Of the 579 bp...
amplified from the Kahramanmaraş samples we used 570 bp for the phylogenetic analysis described below. However, the Kahramanmaraş sequence data includes 32 bp missing data resulting from the difficulty in designing PCR primers to amplify short and fully overlapping amplicons.

We first assessed the phylogenetic position of the ancient Turkish specimens using MrBayes 3.2.2 (Huelsenbeck and Ronquist 2001) and a non-partitioned data set of unique haplotypes (n = 38 including Loxodonta and Mammuthus; n = 32 Elephas). We ran three heated chains (temperature = 0.4) over 10,000,000 generations with a subsampling frequency of 10,000 and a burn-in length of 1,000,000 generations. The consensus tree was visualized in Geneious 7 (Kearse et al. 2012). The nucleotide substitution model for the phylogenetic analysis of the non-partitioned dataset in MrBayes was HKY+I+G, which was the best-fitting model for all partitions but one in the Partition Finder analysis described below.

We then estimated the age of the common ancestor of the node defining the β1-subclade in which we observe the Kahramanmaraş haplotype using BEAST 2.1.3 (Bouckaert et al. 2014). Nucleotides were first grouped into five different partitions: the 1st, 2nd, and 3rd codon positions respectively for cyt-b; the tRNA\(_{\text{th}}\) and tRNA\(_{\text{yg}}\); and the control region. Nucleotide substitution models were estimated for each partition using the Bayesian Information Criterion in Partition Finder (Lanfear et al. 2012). The best-fit model for 1st and 3rd codon positions of the cyt-b gene, the tRNA\(_{\text{th}}\)/tRNA\(_{\text{yg}}\), and the control region was HKY+I+G, and the K80 model for the 2nd codon position in cyt-b. The BEAST analysis included all DNA sequences compiled for this study: Loxodonta (n = 4), Mammuthus (n = 2), Kahramanmaraş (n = 4), and Elephas maximus (n = 534).

We used tip dating on ancient DNA sequences as follows:

- GU984769 (4.8 kyr) (Nyström et al. 2010).
- KC427898 (33.7 kyr) (Palkopoulou et al. 2013).
- Kahramanmaraş elephants (3.5 kyr) (Albayrak and Lister 2012).

To obtain new MRCA estimates we used the following node age calibrations based on mitogenome divergence times estimated previously using a ‘narrow range fossil calibration’ (Brandt et al. 2012):

- Root (normal distribution, mean = 6.81 My, σ = 0.5 My),
- Loxodonta (normal distribution, mean = 5.51 My, σ = 0.7 My),
- Elephas and Mammuthus (normal distribution, mean = 6.01 My, σ = 0.7 My),
- Elephas (normal distribution, mean = 1.35 My, σ = 0.25 My).

We then placed uniform priors on the following clades and estimated their MRCA:

- *Elephas* β clade (lower = 0.0 My, upper = Infinity)
- *Elephas* α clade (lower = 0.0 My, upper = Infinity)
- Kahramanmaraş ‘clade’ (lower = 0.0035 My, upper = Infinity).

We used a strict clock model under a constant population size prior. The clock rate prior was set to default (uniform distribution ranging from 0 – infinity). Three runs of 30,000,000 generations were completed, reaching effective sample sizes (ESS) of >100 (though the vast majority were >200) for each parameter. Trees and trace files were sampled/logged every 3000 generations. The trace files were also visually inspected using Tracer V1.6 (Rambaut, Suchard and Drummond 2013) to ensure proper mixing and that the Markov chain had reached stationary distribution.

### 3. Results and Discussion

We successfully amplified and sequenced DNA from four specimens, yielding a success rate of ca. 44% (4/9), consistent with previously published data on ca. 120 bp mtDNA fragments sequenced from ancient pigs from Near East dating to approximately 3500 cal. BP (Ottoni et al. 2013). We furthermore observed damaged bases (C > T or G > A transitions) across all sequenced DNA fragments, consistent with cytosine deamination to uracil which is a common type of damage in ancient DNA molecules (Gilbert et al. 2003). The negative extraction and PCR controls did not produce PCR amplified products. These observations support the authenticity of our data.

#### 3.1. Genetic affinity of the Kahramanmaraş elephants

The mtDNA phylogeny of extant *E. maximus* consists of two highly divergent clades (α and β). The four ancient elephants from Kahramanmaraş all carried an mtDNA haplotype that groups in the β1 clade of the major β clade of *E. maximus* (Figure 1B). A comparison with previously published data on NCBI GenBank shows that the Kahramanmaraş sequences are identical to only a single modern elephant from Thailand (Unpublished GenBank accession JQ287727). While this haplotype appears to be rare among extant populations, the β1 sub-clade is present across their range (except Peninsular Malaysia, Borneo and Sumatra), being particularly abundant in South and Central India – and varies in frequency from 3.2–100% (Vidy, Sukumar and Melnick 2009). The lack of a clear phylogeographic structure means that it would be problematic attempting to ‘link’ the Kahramanmaraş elephants to any specific modern population. We therefore conclude only that they group within present-day variation.

The α and β clades are estimated to share a most recent common ancestor (MRCA) some 1.35 [0.9–1.84] Ma (Brandt et al. 2012). The internal MRCA for these clades have been estimated to 0.86 Ma (0.52–1.21 Ma) and 1.58 Ma (1.28–1.86 Ma) respectively (Vidy, Sukumar and Melnick 2009). Our analysis in BEAST provides somewhat younger dates and places the MRCA of the α-clade at 0.49 Ma (0.18–0.87 Ma) and the β-clade at 0.99 Ma (0.62–1.41 Ma). However, the MRCA for all *E. maximus* as well as the MRCA of the α and β clades are still older than the earliest morphologically distinct *E. maximus* fossils, which date to ca. 0.2–0.1 Ma (Maglio 1973, Lister et al. 2013), implying that the present-day mtDNA gene pool consists of lineages that originated in earlier, ancestral *Elephas* species
(Vidya, Sukumar and Melnick 2009). Morphology as well as geographic and temporal distribution suggests that *E. hysudricus* is a likely ancestor of *E. maximus* (Maglio et al. 2013). However, the deep divergences of the α and β clades might in theory reflect genetic contribution from more than one species, such as *E. hysudrinicus* or even *P. namadicus*, in addition to *E. hysudricus* (Lister et al. 2013). This possibility is congruent with a recent paper (Meyer et al. 2017) that analysed mitogenomes and partial autosomal genomes of *P. antiquus* and suggested possible historical gene flow between that species and *E. maximus*. Similarly, historical gene flow between African savanna and forest elephants (*L. africana* and *L. cyclotis*) explains the observation that some savanna elephants carry the deeply divergent ‘F-clade’ mitochondrial haplotypes and not ‘S-clade’ haplotypes, despite the former being characteristic of forest elephants while the latter is private to savanna elephants (Roca, Georgiadis and O'Brien 2004).

By estimating that the haplotype carried by the Kahramanmaraş elephants and a modern Thai elephant shared a common ancestor that lived 3.7–58.7 kyr ago (95% HPD; mean 23.5 kyr) we can conclude that it very likely belonged to a population of *E. maximus*. This implies that range-wide population connectivity, such as gene flow or migration, has taken place at some time or times since the start of MIS 3 some 57 kya, likely reflecting range and population expansion that led to the establishment of the Bronze Age southwest Asian population.

### 3.2. Molecular taxonomy and tooth morphology

Deraniyagala (1950) defined a subspecies, *E. maximus asurus*, for the Near Eastern elephants, but their supposed characteristics were based on doubtful interpretation of Bronze Age illustrations, aside from the suggestion, based on rather few bone measurements, that the animals were of unusually large size (Deraniyagala 1955, Pfälzner 2013). While the majority of molars from Kahramanmaraş are of typical *E. maximus* morphology, a few are more ambiguous (Albayrak and Lister 2012) (Table 1). The sequenc

### 3.3. Status of the Near-Eastern Asian elephant population

The sharing of a haplotype between the Turkish fossils and a modern individual from the Far East supports the hypothesis that the Near-Eastern elephants were the western-most expansion of the core population of India and Southeast Asia. It leaves open the possibility that the Turkish population was established only shortly before its Bronze Age date. However, the estimated age of the common ancestor of the Kahramanmaraş haplotype, extending to 58.7 kyr at its 95% lower bound, means that the Near Eastern populations could alternatively have been established, or at least separated from other populations in southern Asia, as long ago as MIS 3, assuming that the Turkish population and their direct ancestors were isolated without gene flow. The split time between the population ancestral to the Kahramanmaraş elephants and the eastern core populations could in theory have been even longer in the presence of gene flow. The genetic data do not distinguish between these possibilities, nor therefore between natural range expansion and human introduction of elephants into the Near East. The latter has been posited mainly due to the lack of well-dated skeletal remains of elephants before 3500 BP, or ivory (which could have been humanly transported) before 4000 BP (Cakır and Ikram 2016). The use of captive elephants likely did not begin in India until ca. 4600 BP (Sukumar 2011), and deliberate long-distance movement of elephants has been thought impracticable in the Bronze Age (Pfälzner 2013). Although there are historical/pictorial sources referring to the movement of live elephants in the Bronze and Iron Ages in southwest Asia (Collon 1977), the westward movement of elephants from India for war or other purposes is first clearly recorded around ca. 2600 BP (Sukumar 2011). Of considerable significance is that the Kahramanmaraş assemblage, although not systematically excavated, appears very likely to represent a wild-living population. It comprises various elements of the skeleton of multiple individuals, some of them probably associated, with no cut marks, artefacts, or other sign of human activity (Alaura 2016, Yar et al. 2016) in stark contrast to Near-Eastern archaeological sites with occasional elephant remains (Cakır and Ikram 2016). This suggests a naturally established population, or conceivably a wild-living population founded by humanly-introduced animals.

Combining these data suggests that the most likely scenario is the westward expansion of the elephants' range, presumably during favourable climatic episode(s), i.e. with sufficient warmth and moisture to provide the required vegetation and drinking water to support the animals along the route. These requirements would have been necessary whether the population arrived by natural expansion or human agency. Rowe et al. (2012) show cyclicity of wet and dry episodes in Turkey extending back into the Late Pleistocene, as do Rajagopalan et al. (1997) for the Indian subcontinent. Corresponding cyclical variation in the range of mammal species, including elephants, is likely. More recent studies show that the climate of the Near and Middle East underwent a shift from wetter
conditions in the early Holocene to drier conditions from around 6500 BP, but that this was interrupted by an interval of moister conditions between around 5000–3500 BP (Finné et al. 2011). For example, Migowski et al. (2006) indicate a wet phase in the Levant ca. 5600–3500 BP, based on Dead Sea lake level changes. A variety of data from Lake Van in southeastern Turkey and Soreq Cave in Israel suggests that a moister climate prevailed in the region during the Chalcolithic and Early Bronze Age, although the region became significantly more arid from around 2000 cal. BC (ca. 4000 BP) (Bar-Matthews et al. 1999, Wick, Lemcke and Sturm 2003). Together, these data suggest that wetter climatic periods suitable for the immigration of elephants did prevail in the region at intervals during the Holocene, although the general climatic trend is one of increasing aridity.

Cyclical expansion and contraction of range is also supported by earlier fossil records of Elephas in the Near East, where the earliest dated finds are an Early Pleistocene molar tooth of Elephas sp. from Evron Quarry in Israel (Tchernov et al. 1994), and five Middle Pleistocene teeth of Elephas cf. hysudricus from Ma‘ayan Baruch in Israel and ‘Ain Soda in Jordan (Lister et al. 2013).

Data Accessibility Statement

The consensus DNA sequences are publically available on GenBank (accession numbers MF314177-MF314180).

Acknowledgements

Raman Sukumar and Liora Kolska Horwitz are thanked for very helpful discussion.

Funding Information

The study was supported by the Natural History Museum (SIF fund).

Competing Interests

The authors have no competing interests to declare.

Author contributions

LGF designed the study, acquired and analysed data, and co-wrote the manuscript. EA provided samples and contextual information, and critically revised the manuscript. AML conceived the study and co-wrote the manuscript.

References


