Ancient DNA from Nubian and Somali wild ass provides insights into donkey ancestry and domestication

Birgitta Kimura1,†, Fiona B. Marshall2, Shanyuan Chen3, Sonia Rosenbom3, Patricia D. Moehlman1, Noreen Tuross5, Richard C. Sabin6, Joris Peters7, Barbara Barich8, Hagos Yohannes9, Fanuel Kebede10, Redae Teclai9, Albano Beja-Pereira3 and Connie J. Mulligan1,*

1Department of Anthropology, University of Florida, PO Box 103610, Gainesville, FL 32610, USA
2Department of Anthropology, Washington University, St Louis, MO 63130, USA
3Centro de Investigação em Biodiversidade e Recursos Genéticos (CIBIO-UP), Universidade do Porto, Campus Agrário de Vairão, 4485-661 Vairão, Portugal
4IUCN/SSC Equid Specialist Group, Box 2031, Arusha, Tanzania
5Department of Anthropology, Peabody Museum, Harvard University, 11 Divinity Avenue, Cambridge, MA 02138, USA
6Department of Zoology, The Natural History Museum, Cromwell Road, South Kensington, London SW7 5BD, UK
7Department of Veterinary Sciences, Institute of Palaeoanatomy and the History of Veterinary Medicine, Ludwig-Maximilian University, D-80539 Munich, Germany
8Dipartimento di Scienze Storiche, Archeologiche e Antropologiche dell’Antichità, Università degli Studi di Roma ‘La Sapienza’, Via Palestro 63, 00185 Rome, Italy
9Ministry of Agriculture, PO Box 1048, Asmara, Eritrea
10Ethiopian Wildlife Conservation Authority, PO Box 817, Debrezeit, Ethiopia

Genetic data from extant donkeys (Equus asinus) have revealed two distinct mitochondrial DNA haplogroups, suggestive of two separate domestication events in northeast Africa about 5000 years ago. Without distinct phylogeographic structure in domestic donkey haplogroups and with little information on the genetic makeup of the ancestral African wild ass, however, it has been difficult to identify wild ancestors and geographical origins for the domestic mitochondrial clades. Our analysis of ancient archaeological and historic museum samples provides the first genetic information on the historic Nubian wild ass (Equus africanus africanus), Somali wild ass (Equus africanus somaliensis) and ancient donkey. The results demonstrate that the Nubian wild ass was an ancestor of the first donkey haplogroup. In contrast, the Somali wild ass has considerable mitochondrial divergence from the Nubian wild ass and domestic donkeys. These findings resolve the long-standing issue of the role of the Nubian wild ass in the domestication of the donkey, but raise new questions regarding the second ancestor for the donkey. Our results illustrate the complexity of animal domestication, and have conservation implications for critically endangered Nubian and Somali wild ass.

Keywords: donkey domestication; ancient DNA; Nubian and Somali wild ass

1. INTRODUCTION

Domestication of the donkey (Equus asinus) approximately 5000 years ago transformed ancient societies and land-based transport in Africa and Eurasia, allowing the development of mobile pastoralism and ancient overland trade routes and contributing to the growth of the early Egyptian State [1–3]. Today donkeys are essential means of transport for people living in many mountainous, desert and poor regions of the world [4,5]. Little is known, however, about domestication of the donkey. Historically it has been thought that ancient Egyptians domesticated the African wild ass (Equus africanus) although near-eastern domestication has also been suggested [6–8]. Recent research shows the importance of load-bearing donkeys to the earliest pharaohs and emphasizes the slow nature of their morphological, and probably genetic, change during domestication [3]. Research on the genetics of modern donkeys worldwide demonstrated the existence of two distinct mitochondrial haplogroups, termed Clades 1 and 2 [9–11]. Genetic variability in both domestic
maternal lineages was greatest in Africa, therefore mitochondrial results did not provide support for the hypothesis of Asian domestication. Specifically, Beja-Pereira et al. [9] argued that Asiatic wild asses were excluded as progenitors of modern donkeys and proposed two separate domestication events that occurred on the African continent [9]. There are at least three possible African candidates for wild ancestors of the donkey, the Atlas, Nubian and Somali wild ass (figure 1), reflecting the fact that distinct geographical patterning does not exist in modern donkey haplogroups.

Based on the available genetic data, it has been hypothesized that Nubian wild asses were the ancestors of donkeys of Clade 1 and that a relative of the Somali wild ass, probably extinct, was the ancestor of Clade 2 [2,9]. Archaeological data, the distribution of African wild ass, and linguistic data suggest that mobile African cattle herders domesticated the donkey in response to increasing aridity in the Sahara and the Horn [2,9,12].

In order to investigate the relationships of African wild ancestors to domestic donkey clades, additional information is needed on variability within and among ancient and modern wild ass populations [11]. African wild asses have been well documented in at least three regions of Africa, but there has been debate over the extent to which populations represent the remnants of once continuous variability versus distinct subspecies [13–15]. Nubian wild asses (Equus africanus africanus) were still fairly common in the Atbara region and the Red Sea Hills (NW Sudan) during the first half of the twentieth century AD and the Somali wild ass (Equus africanus somaliensis) existed in southern Eritrea, Ethiopia and Somalia (figure 1). The Atlas wild ass (Equus africanus ‘atlanticus’) was once confined to the northwestern part of the continent and probably became extinct in early historic times [13] (figure 1).

There has also been debate over whether the African wild ass once ranged into western Asia. This is complicated by a lack of reliable historic documentation of E. africanus in the region and difficulties in morphological discrimination between E. africanus and E. hemionus. In addition, the recent discovery that early dynastic Egyptian donkeys used for transport at Abydos are morphologically indistinguishable from the African wild ass [3] raises the possibility that faunal specimens attributed to wild ass in Asia could be derived from early domestic donkeys.

Historic populations of wild ass have been distinguished phenotypically by size, and the presence or absence of distinctive leg stripes and shoulder crosses [13] (figure 1). Osteological differences in size and cranial
mortality of variability among these wild populations is not straightforward because there are only a few skeletons available in museum collections. Additional samples are virtually impossible to procure because Somali wild ass are critically endangered today, with perhaps as few as 600 individuals left in the wild [15], and Nubian wild ass have only been infrequently sighted since the 1970s and are therefore considered possibly extinct [16].

The small size and remote distribution of remaining populations also explains why little is known about the genetics of contemporary African wild ass. Prior to this study, there were only five published sequences in GenBank, three from Somali wild ass and two from individuals tentatively identified as Nubian wild ass [9].

To better understand variability in *E. africanus* populations across their former African ranges, additional samples of the remaining Somali wild ass populations as well as ancient DNA (aDNA) data on historic and ancient African wild ass have been obtained. As will be shown, archaeological and museum specimens represent an invaluable genetic repository for African wild ass.

In light of the genetic research that indicates an African origin for both clades of domestic donkey [2,9], the goal of our study was to investigate African settings for domestication of the donkey and to test the current hypotheses that (i) Nubian wild ass were ancestors of Clade 1 domestic donkeys and (ii) a relative of the Somali wild ass was the ancestor of donkeys of Clade 2 [9]. We used aDNA methods to analyse 12 ancient samples from archaeological sites in northeast Africa and Yemen, ranging in age from 3000 years ago to the early Holocene, in addition to nine tissue samples from all known historic Nubian skeletons and two Somali wild ass museum specimens collected between 1880 and 1950 (electronic supplementary material, Background for more historical and taxonomic information on the historic samples). In addition, dried skin and faecal samples from Somali wild ass populations in Ethiopia and Eritrea. These mitochondrial DNA (mtDNA) data from modern, historical and ancient specimens were combined with previously published sequences for network and phylogenetic analysis.

2. MATERIAL AND METHODS

Holocene archaeological (n = 12) and historic museum (n = 11) samples of Nubian wild ass, Somali wild ass and donkey were obtained for this study (electronic supplementary material, table S1) through the comprehensive evaluation of a significant portion of all specimens in existence including: annual camel-based surveys of critically endangered African wild ass conducted by the International Union for Conservation of Nature (IUCN), approximately 12 skeletons of African wild ass held in world museums, and a survey of isolated donkey bones from African archaeological sites. Appropriate permits were obtained for all specimens including CITES permits for all wild ass specimens owing to their status as critically endangered (see electronic supplementary material, Background for more historical and taxonomic information on the historic samples).

Faecal samples from Somali wild ass from Ethiopia (n = 6) and Eritrea (n = 27) were collected across the species habitat range after observation of the animal (electronic supplementary material, table S1). Each sample was stored in white paper envelopes, dried for 24 h, and shipped to CIBIO-Universidade do Porto. In addition, dried skin from five skeletons of animals that died during the drought of 2006 in Eritrea was used for DNA extraction.

Samples were analysed using standard precautions for working with ancient DNA. At the University of Florida and Harvard University, analyses were performed in laboratories dedicated to ancient DNA work in which no previous work on equids had been performed. Ground bone samples were extracted with two different methods, one based on DNA binding to silica [17] and one using phenol/chloroform extraction [18]. Only one sample and the accompanying extraction blank were processed at a given time. Samples that yielded DNA were re-extracted with a minimum of one other sample being processed in between the first and second extraction of the positive samples. Museum tissue, dry skin and faecal samples were extracted with the Qiagen DNeasy tissue kit at the University of Florida and CIBIO.

For the archaeological samples, primers were designed to amplify segments of 56-158 base pairs (bps) of the most variable regions of the control region that specifically distinguish between domestic donkey clades (electronic supplementary material, table S2). Museum tissue, dry skin and faecal samples were amplified in three to four overlapping segments ranging from 158 to 308 bps in length. PCR conditions were as follows: 25 μl reaction with 1x manufacturer’s PCR buffer, 2.5 mM MgCl₂, 200 μM each dNTP, 1 μM of each primer, 1.5 μg BSA and 1 unit AmpliTaq Gold DNA polymerase or a 25 μl reaction with 1x Bioline Short mix and 1 μM of each primer (see electronic supplementary material, table S2 for annealing temperatures). A minimum of three independent PCR amplifications were performed with each primer pair.

PCR amplification products from the ancient archaeological samples were cloned into a TOPO TA vector (Invitrogen) following the manufacturers’ recommendations. Eight to 12 colonies from each amplification product were sequenced and analysed on a Beckman CEQ 8000, following the manufacturers’ recommended protocol for sequencing. Amplification products from the tissue, dry skin and faecal samples were sequenced directly using the forward and reverse primers that were used for the PCR amplification. Products from a minimum of three independent PCR reactions were sequenced in both directions for the historic samples. Additional details on extraction, amplification and sequencing are available in electronic supplementary material, material and methods.

Newly reported equid sequences were used to create both a median-joining network and a phylogeny. These sequences were aligned with Clustalw from MEGA 4 software [19] and compared with previously published sequences: *E. asinus* NC_001788 [20], DQ44878-DQ449023 [10] and AY569462-AY569547 [9]. Sequences of 440 bps were used in the network and phylogenetic construction. In cases where ancient or historic sequences were shorter than 440 bps, but identical to previously published sequences, those sample labels are listed along with the identical (full length) previously published sequences on the network and phylogeny. NHML 1939 yielded a sequence of only 204 bps and is a new sequence, so this sample does not appear in either the network or phylogeny. Median-joining networks [21] were constructed with NETWORK v. 4.5 (http://www.fluxus-engineering.com/). Reticulations were resolved through a maximum-parsimony criterion [22]. Information on the phylogenetic analysis and estimation of
time to most recent common ancestor for each clade can be found in electronic supplementary material, material and methods.

Determination of wild or domestic status of the ancient Uan Muhuggiag magnun, Os Carpale III (specimen Verona 3870, articulated with aDNA no.7), was made on the basis of morphometrical analysis. Size-based identifications of domestic versus wild ass are not reliable for the earliest periods of domestication prior to size decrease [3], but smaller donkeys can readily be distinguished from wild ass in Africa after ca 4000 cal year BP. Greatest breadth and maximum length measurements were made to the nearest millimetre using a measuring board and callipers and following conventions established by von den Driesch [23] (electronic supplementary material, table S3). Uan Muuggiag measurements were compared with those from seven ancient donkeys, nine modern donkeys and 14 wild ass, including three juveniles. The Uan Muuggiag mandible (Verona 3988, aDNA no.8) was aged using incisor dental eruption and wear following the sequence documented for donkeys [24].

3. RESULTS

The collection of samples successfully analysed in this study covers the presumed range of Nubian and Somali wild ass over northeastern Africa and includes modern, historic and ancient specimens spanning a time depth of 3000 years. In total, three of the 12 ancient samples, 10 of the 11 historic samples and 33 modern Somali wild ass samples were successfully amplified and sequenced for the mitochondrial control region (electronic supplementary material, tables S1 and S4; all sequences are available through GenBank, HM622626-HM622669). Final analysed sequences ranged in size from 201 to 440 bps and were composed of amplicons ranging in size from 33 to 440 bps, i.e. some specimens required multiple fragments to construct an informative DNA sequence. Four historic samples were independently processed and their sequences confirmed at the CIBIO-Universidade do Porto, and one ancient sample was independently extracted and amplified at Harvard University (electronic supplementary material, table S1).

Of the ancient samples, one had the maternal genetic signature of horse and is not analysed further here. This specimen was represented by a single tooth and had been provisionally identified as donkey. In actuality, the specimen may belong to a mule, i.e. offspring of a female horse and male donkey. Two ancient specimens from the Uan Muuggiag rock shelter in the central Sahara [25–27], a mandible with one permanent incisor erupted and a trapezoid, were also successfully sequenced. An uninform that was articulated with the trapezoid was directly AMS dated at the University of Oxford to 3160–2975 cal BP (electronic supplementary material, table S5). The Uan Muuggiag sequences matched a historic Nubian wild ass sequence reported below (NHML1904, electronic supplementary material, table S4) and fell in Clade 1, which supports an ancestral role for Nubian wild ass within Clade 1. Furthermore, the Uan Muuggiag sequences were identical to each other, suggesting that both specimens came from a single individual or from maternally related animals. Morphometric analysis of the Uan Muuggiag specimens documented that they were more consistent with those of a small domestic donkey than those expected for either adult or juvenile wild ass, suggesting that the Uan Muuggiag animal(s) was domestic (electronic supplementary material, figure S1).

Nine of the historic samples were from animals identified as Nubian wild ass on phenotypic and geographical grounds. Eight of those specimens yielded five different sequences that fell within Clade 1 (figure 2 and electronic supplementary material, figure S2). One of the five sequences was new (BSZM 1952). Four samples, two pairs of mother and foetus, had identical sequences that matched a haplotype also found in domestic donkeys. They were collected from two areas in close proximity in the Red Sea Hills. Another match to a domestic donkey haplotype was found in an animal from the Tibesti Mountains of the Sahara (RMCA31155). Significantly, the sequence from one sample from the Atbara region in Sudan (NHML1904) exactly matched a modern sequence from the eastern Sudan that had been tentatively identified as Nubian wild ass (H6; [9]), suggesting that Nubian wild ass maternal lineages survived at least until the last decade in the eastern Sudan. The geographical breadth of the successfully assayed specimens confirms the ancient range of the Nubian wild ass in Sudan and northern Eritrea and the presence of animals of Clade 1 not only east, but also west, of the Nile River as far as the central Sahara. The ninth historic sample (BSZM1963) attributed to Nubian wild ass on phenotypic grounds had a sequence identical to a haplotype found in domestic donkeys of Clade 2 (figure 2 and electronic supplementary material, figure S2).

The 33 modern Somali wild ass specimens fell in the same clade as previous Somali wild ass specimens (WH1–4). This clade is well separated from the domestic donkey Clades 1 and 2 and is clearly not ancestral to either clade (figure 2 and electronic supplementary material, figure S2). Only four new haplotypes were found in the 33 specimens analysed and haplotype diversity of the Somali wild ass clade is only 0.7417 ± 0.0444 (compared with 0.9309 ± 0.0102 and 0.8212 ± 0.0268 for Clades 1 and 2, respectively), suggesting that the genetic variability in present-day Somali wild ass is low. The new haplotypes are found in both Eritrea and Ethiopia, and show no geographical structure. The single historic Somali specimen that was successfully amplified came from Berbera, Somalia. Collected around 1886, it showed a sequence identical to that of one of the new Somali wild ass haplogroups (WH1) from Eritrea and Ethiopia (figure 2 and electronic supplementary material, figure S2 and table S4). This result demonstrates a degree of historical continuity in the mitochondrial variability of Somali wild ass within the region over the last 120 years.

We also calculated the coalescence time of each clade, i.e. the time to the most recent common ancestor (TMRCA), as follows: Clade 1: 406 000 years ago (95% confidence interval 105 400–811 300 years), Clade 2: 394 600 years ago (95% confidence interval 86 100–661 300 years), Somali wild ass clade: 359 500 years ago (95% confidence interval 57 600–770 800 years). Although there may be some uncertainty in the dates owing to time dependency [28], these dates clearly predate the domestication time for donkey of approximately 5000 years ago.
4. DISCUSSION

(a) Ancestors for the donkeys

The diversity and geographical variability in historic and ancient DNA together with information on modern donkey mitochondrial genetic variation provide new insights into relations among ancestral wild ass, relations of wild ass to domestic haplogroups, and the process of donkey domestication. Our results demonstrate that Nubian and Somali wild asses are mitochondrially distinct. Furthermore, we show that the historic Nubian wild asses and domestic donkeys of Clade 1 are almost indistinguishable on the basis of mtDNA; five of our historic Nubian wild ass samples had haplotypes identical to domestic donkeys of Clade 1. It is probable that the identical wild ass sequences represent survival of the originally domesticated maternal haplotypes in the wild population, although we cannot rule out the possibility that they were introduced into wild herds by feral female donkeys. Historic specimens collected by naturalists over the last two centuries verify a northern Sudanese and Eritrean distribution of Nubian wild ass in northeast Africa, but the Uan Muhuggiag and Tibesti data suggest that donkeys
of Clade 1 and/or Nubian wild asses were present as far west as the central Sahara in late (pre)historic times.

When our results are combined with 98 previously published haplotypes, the topology of the network (figure 2) provides some interesting perspectives on domestication processes. The Clade 1 topology resembles that found in European and Asian domestic pigs [29] and domestic and wild reindeer [30], with several smaller nodes and wild animals interspersed with domestic. It presents an interesting example of survival of wild populations during the process of domestication giving rise to indistinct differences at the mtDNA level between wild and domestic individuals. This process has been proposed for horse, dog, pig and reindeer [30–33], but we are able to verify this conclusion for donkeys since we have multiple cases of mtDNA sequences shared by wild and domestic specimens. The Clade 1 topology is consistent with a scenario whereby the Nubian wild ass was domesticated in several areas and/or over an extended long period, with multiple recruitments from the wild, similar to the domestication process suggested for dogs and goats [34,35]. The much broader distribution of Nubian wild ass in former times and likely domestication by cattle-herders who ranged widely over the Sahara from as early as 8900–8400 cal BP provide geographical, social and temporal contexts for these processes [36–38]. Moreover, the use of ‘morphologically wild’ donkeys for heavy transport at Abydos in ancient Egypt ca 3000 BC [3] has already illustrated a slow process of donkey domestication with late morphological and genetic change.

Introgression with wild ancestors is especially probable among donkeys because they are not herd animals and are not intensively managed by African pastoralists [2]. Pastoralists who keep donkeys for transport particularly value strength and reproductive potential in their animals and recruit both males and female donkeys to their herds [2]. Furthermore, capture of wild ass through trapping is historically documented and illustrated in African rock art [39–41], but is indiscriminate with respect to sex (for additional information, see electronic supplementary material, Background).

We conclude that donkeys of Clade 1 have a long history in the Sahara, that a Nubian wild ass was their ancestor, and that it is probable there was interbreeding between wild and domestic forms over a long period of time with recruitment of several maternal haplotypes from the wild.

(b) The contribution of the Somali wild ass to the domestic gene pool

Our extended mitochondrial dataset from free-living Somali wild ass shows that Somali wild ass are distinct from Nubian wild ass and domestic donkeys of both Clades 1 and 2 (figure 2 and electronic supplementary material, figure S2). Given the extensive haplotype networks found in Clades 1 and 2, it is surprising to find so few Somali wild ass haplotypes after increasing the sample size by an order of magnitude. The low variation and large sample size of Somali wild ass make it unlikely that additional lineages will be identified and, thus, make Somali wild ass a less probable candidate for the ancestor of Clade 2 than previously thought [9,11]. Furthermore, the equal distance of the three major clades to each other diminishes the possibility that the ancestor of Clade 2 lies in either of the other two clades. The very old coalescence times of the three clades reflect the long period of time before donkey domestication and suggest that substantial genetic structuring, fragmentation and/or geographical isolation of wild ass mitochondrial variation may have developed prior to domestication. As a result we cannot rule out the possibility that wild ass in northeast Africa may have had additional, yet unrecognized, genetic substructure and particularly that Clades 1 and 2 may both have Nubian-like wild ass ancestors (see electronic supplementary material, Background for additional details).

In addition, the observation that wild ass/donkey mitochondrial variation may have undergone significant reductions over time also raises the possibility that the ancestor of Clade 2 belonged to an extinct population. Archaeological data suggest the Holocene ranges of African wild ass were substantially more extensive, the presence of wild ass in the central and eastern Sahara being evidenced by rock engravings [42] and skeletal remains [2,26,43], which is consistent with our genetic results. Thus, there are several possibilities for the geographical origin of the wild ancestor of Clade 2. In addition to northeastern Africa, candidates include the ancient range of the Atlas wild ass in the Maghreb [13] and the coast of Yemen, where specimens identified as early domestic donkeys or wild ass have been excavated ([44]; see also [2]); however, our ancient samples from these regions did not yield genetic material. Additional aDNA research in Africa and Asia as well as Y chromosome or nuclear genetic data on donkeys and extant African wild ass are needed to pinpoint the locus of domestication of Clade 2.

(c) Patterns of domestication and conservation implications

The findings presented in this study clarify the role of the Nubian wild ass in the domestication of the donkey but raise new questions regarding the second ancestor for the donkey. Evidence for domestication of several Nubian haplotypes, multiple recruitments from the wild, and ongoing gene flow in Clade 1, contrasts with a simpler domestication process starting from fewer ancient founders for Clade 2. These distinct patterns fit with recent research on other livestock species showing multiple domestication events with differing histories, social contexts and timelines [9,11,29,30,35,36,45,46]. Our findings also have several implications for conservation. (i) Nubian wild asses are distinct from Somali wild asses based on mtDNA, a result that indicates the need for separate management of Nubian and Somali populations. (ii) The finding that maternal lineages of the Nubian wild ass may have survived in the eastern Sudan until the 1990s implies that Nubian wild asses are not extinct or became extinct very recently, and reinforces the need for surveys and management plans for eastern Sudan and northern Eritrea. (iii) Extant Somali wild ass in Eritrea and Ethiopia shows an absence of geographical structuring of genetic variation as well as low haplotype diversity, which may reflect past bottlenecks in ancestral populations in which short-term crashes wiped out many haplotypes. These results suggest that captive breeding
populations of Somali wild ass already sample much of the available mitochondrial diversity. Our findings represent a valuable contribution to debates regarding variability, phylogeny and management of extant but critically endangered African wild ass [13–15]. Our research also underscores the need for further studies of the nuclear and Y chromosomal DNA of extant populations and for more specimens for aDNA analysis.

This research was supported by National Science Foundation grant BCS-0447369 (FM), FCT grants (SRFR/BPD/26802/2006, SRFR/BPD/17822/2004, and PTDC/BIA/BDE/64111/2006) (ABP) and grants from Wildlife Trust, Saint Louis Zoo, Basel Zoo, Liberec Zoo and Sea World and Busch Gardens Conservation Fund (PDM). The authors would like to thank four anonymous reviewers for many useful comments. We thank V. Costa for help during faecal DNA extraction and T. Schilling for assistance with figure 1. We are indebted to F. Alhaique, the late A. Bietti, L. Chaix, L. Khalidi, C. Lewis, R. Meadow, W. van Neer, T. Webber, W. Wendelen and to the late A. Bietti, L. Chaix, L. Khalidi, C. Lewis, R. Meadow, W. van Neer, T. Webber, W. Wendelen and to the institutions and individuals listed in electronic supplementary material, table S1 who generously provided material or assisted in collection of samples for this study.

REFERENCES


