Ancient DNA shows domestic horses were introduced in the southern Caucasus and Anatolia during the Bronze Age

Silvia Guimaraes, Benjamin Arbuckle, Joris Peters, Sarah Adcock, Hijlke Buitenhuis, Hannah Chazin, Ninna Manaseryan, Hans-Peter Uerpmann, Thierry Grange, Eva-Maria Geigl

To cite this version:

Silvia Guimaraes, Benjamin Arbuckle, Joris Peters, Sarah Adcock, Hijlke Buitenhuis, et al.. Ancient DNA shows domestic horses were introduced in the southern Caucasus and Anatolia during the Bronze Age. Science Advances , American Association for the Advancement of Science (AAAS), 2020, 6 (38), pp.eabb0030. 10.1126/sciadv.abb0030 . hal-03039029

HAL Id: hal-03039029
https://hal.archives-ouvertes.fr/hal-03039029
Submitted on 3 Dec 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Ancient DNA shows domestic horses were introduced in the southern Caucasus and Anatolia during the Bronze Age

Silvia Guimaraes1, Benjamin S. Arbuckle2, Joris Peters3,4, Sarah E. Adcock5*, Hijlke Buitenhuis6*, Hannah Chazin7*, Ninna Manaseryan8*, Hans-Peter Uerpmann9*, Thierry Grange1#, Eva-Maria Geigl1#

1Institut Jacques Monod, CNRS, University of Paris, Paris; biguimaraes@hotmail.com; thierry.grange@ijm.fr; eva-maria.geigl@ijm.fr
2Department of Anthropology, University of North Carolina at Chapel Hill, NC 27599, USA; bsarbu@email.unc.edu
3ArchaeoBioCenter and Department of Veterinary Sciences, Institute of Palaeoanatomy, Domestication and the History of Veterinary Medicine, Ludwig Maximilian University Munich, Kaulbachstr. 37/111, 80539 Munich, Germany; joris.peters@palaeo.vetmed.uni-muenchen.de.
4State Collection of Anthropology and Palaeoanatomy Munich, Bavarian Natural History Collections, Karolinenplatz 2a, 80533 Munich.
5Department of Anthropology, University of Chicago, 1126 East 59th Street, Chicago, Il 60637, USA; adcock@uchicago.edu
6Groningen Institute of Archaeology, University of Groningen, 9712 ER Groningen, The Netherlands; hijlkebuitenhuis@gmail.com
7Department of Anthropology, Columbia University, 1200 Amsterdam Avenue, New York, NY 10031, United States; h.chazin@columbia.edu
8Scientific Center of Zoology and Hydroecology, Institute of Zoology, National Academy of Sciences of the Republic of Armenia; 7, P.Sevak Str., Yerevan 0014, Armenia; ninna_man@yahoo.com
9Institut für Ur- und Frühgeschichte und Archäologie des Mittelalters, Abteilung Ältere Urgeschichte und Quartärökologie, Zentrum für Naturwissenschaftliche Archäologie, Rümelinstr. 23, 72070 Tübingen, Germany; hans-peter.uerpmann@uni-tuebingen.de
$First co-authors
*These authors are listed in alphabetical order
#Last co-authors
aPresent address: Archaeology of Social Dynamics, Institució Milà i Fontanals, CSIC, Barcelona, Spain. CIBIO-InBIO, Centro de Investigación em Biodiversidade e Recursos Genéticos, Universidade do Porto, Campus Agrário de Vairão, Vairão, Portugal.

Corresponding author: Eva-Maria Geigl, Institut Jacques Monod, Institut Jacques Monod, CNRS, University of Paris, 15 rue Hélène Brion, 75013 Paris, France; telephone +33-157278132; eva-maria.geigl@ijm.fr
Abstract

Despite the important role horses played in human history, particularly in the spread of languages and cultures, and correspondingly intensive research on this topic, the origin of domestic horses remains elusive. Several domestication centers have been hypothesized, but most of these have been invalidated through recent paleogenetic studies. Anatolia is a region with an extended history of horse exploitation that has been considered a candidate for the origins of domestic horses but has never been subject to detailed investigation. Our paleogenetic study of pre- and protohistoric horses in Anatolia and the Caucasus based on a diachronic sample from the early Neolithic to the Iron Age (~8,000 to ~1,000 BCE), encompassing the presumed transition from wild to domestic horses (4,000-3,000 BCE), shows the rapid and large-scale introduction of domestic horses at the end of the third millennium BCE. Thus, our results argue strongly against autochthonous independent domestication of horses in Anatolia.

Keywords: Equus, horse, phylogeography, paleogenetics, ancient DNA, domestication
Introduction

The domestication of the horse ca. 5,500 years ago represents one of the most important technological innovations in the ancient world (1, 2). With the harnessing of horsepower, political, economic and social relationships throughout the ancient world were transformed as horses revolutionized transportation and impacted patterns of trade, warfare and migration (e.g., (1-4)). Archaeological, organic residue and genetic analyses suggest that the domestic horse originated in the Central Asian steppes, then spread into eastern Europe and later into Southwest Asia (SWA) (e.g., (2, 5-11)). In particular, data from the Botai hunter-gatherer culture in Kazakhstan suggest that by the mid to late fourth millennium BCE horses were bitted, milked, selected for the TRPM1 coat-color locus, as well as kept in enclosures and were therefore under intensive management (10, 12-14). A recent study of ancient horse genomes, however, challenged the view that modern domestic horses derived from Central Asia, since “Botai-like” horses were shown to be the ancestors of NE Asian Przewalski’s horses but not the main source of ancient or modern domestic horses (14, 15). Another recent study ruled out a second potential horse domestication center, the Iberian Peninsula, showing that Iberian wild horses went extinct without leaving significant traces in the genomes of modern horses (16). There are, however, two more areas that have been proposed as domestication centers for modern horses: the Pontic-Caspian steppe (17), and Anatolia (8, 18). Although the former has been long hypothesized as the likely source of domestic horses (12, 13, 19), the latter region has been poorly explored regarding its role in horse domestication processes (20-22), despite its long history of wild horse exploitation and its reputation for breeding valuable horses in Classical Antiquity (23).

The origin of the domestic horse in Anatolia, and more generally in SWA, continues to represent a complex archaeological puzzle. A combination of textual, iconographic and archaeozoological data suggest that by the mid to late third millennium BCE, domestic horses were introduced from neighboring mountain regions into Mesopotamia (modern Iraq and NE Syria), where they were often referred to in cuneiform texts as ANŠE-KURRA (“donkey of the mountain”) (24-28). Initially kept only in small numbers, horses rose to prominence across SWA within a few centuries in association with the spread of chariots, a technological innovation of the second millennium BCE (1, 29). Since horse domestication in the Eurasian steppes, a region historically known for its “horse cultures”, likely began in the fourth or perhaps even fifth millennium BCE (17, 30-32), it has long been argued that SW Asian horses are the descendants of these early domesticates, which arrived in the region via poorly understood participation in Pontic-Caspian-Transcaucasian interaction spheres or population movements (33, 34).

Another hypothesis argues that Anatolia played a central role in the transmission of domestic horses into Syro-Mesopotamia (33, 35, 36), and an Anatolian contribution to early domestic populations has been suggested (37-40). Indeed, archaeological data indicate the widespread presence of wild horses (Equus ferus) and also of so-called hydruntines (a subspecies of the Asiatic wild ass named E.hemionus hydruntinus (41)) in early and middle Holocene Anatolia that were regularly exploited (20, 22, 40, 42). The continuity of human-horse interactions from the ninth millennium through the second millennium BCE, when domesticates are known from archaeological contexts, led to the hypothesis that Anatolian wild horses may have been a potential source population for domestic horses. The lack of reliable osteomorphological criteria for differentiating the skeletal remains of wild and domestic horses, however,
has hampered attempts to address the hypothesis of a local domestication. Therefore, the cultural processes and mechanisms triggering the widespread appearance of domesticates in the late third millennium are still elusive. In this study, we take advantage of the abundance of archaeological horse remains from the central Anatolian plateau (20, 42-46) to provide the first rigorous test of the hypothesis of Anatolian horse domestication applying paleogenetics.

Complete present-day mitochondrial genomes have revealed 18 major haplogroups (A-R), the radiation times of which date mostly to the Neolithic and later periods (47). In contrast, studies of the mitochondrial hypervariable region in ancient horses have shown that domestic horses exhibit a much higher amount of genetic variation in mitochondrial lineages compared to cattle, sheep, and pigs (48-51). Furthermore, most mitochondrial lineages observed in domestic horses already existed prior to domestication (52). These analyses of ancient horses did not yield a clear phylogeographic structure that would allow the spatio-temporal origin(s) of horse domestication to be identified. This was interpreted as suggesting that the mobility of wild horses in northern Eurasia allowed constant population reshuffling and repeated recruitment of wild local mares, precluding the establishment of a phylogeographic structure (52, 53).

In contrast, extant domestic horses exhibit remarkably little variation in the male Y chromosome line, with only one haplotype so far identified in modern domesticates, which led to the early claim of a single domestication event for horses (54, 55). Paleogenomic analyses of ancient specimens, however, observed additional male lineages in prehistoric populations prior to domestication and revealed that genetically diverse male founders were involved in early domestication (15, 56). This diversity was subsequently reduced, likely as a result of more directed human selection likely starting in the Iron Age and continuing during Roman times (57), and again during the Islamic conquest and the Byzantine-Sassanid war after 7th – 9th c. CE (16).

Paleogenetic evidence from genetic loci associated with coat color in horses argues for a diversification of coat color starting in the Bronze Age and is considered associated with an early stage of the domestication process (11, 58). Since the appearance of new coat colors is common in domestic taxa compared to their wild counterparts, they provide a useful marker for identifying domestic horses in archaeological assemblages.

Up to now, the origins of domestic horses in Anatolia remained elusive, but careful recovery of horse remains from well-stratified archaeological contexts in Anatolia and in the neighboring Caucasus together with progress in paleogenetic approaches now make it possible to specifically address the processes responsible for the origins of domestic horses in this part of western Asia. For this project, we combined morphological classification of equid remains, which can be hampered by a lack of diagnostic anatomical and/or biometrical criteria (41, 59), with paleogenetic analysis of mitochondrial, Y chromosome DNA, and autosomal DNA markers related to coat color to trace the spatio-temporal dynamics of the emergence of domestic horses in Anatolia. We analyzed over 100 equid remains from 14 prehistoric sites in central Anatolia and the Caucasus covering most of the Holocene (9,000 BCE – 1,000 CE) to gain insights into the origins of domestic horses in Anatolia, a pivotal issue in Near Eastern history.
Results

Based on our extensive experience with ancient animal remains from SWA including Anatolia, we expected DNA to be highly degraded in most samples that we had collected, of which only a few were petrous bones. Indeed, DNA in osseous remains from SWA is notoriously poorly preserved, petrous bones being an exception, although not all of them contain preserved endogenous DNA. For this reason, we decided to rely on a highly optimized metabarcoding approach combining the sensitivity of PCR and the efficiency of next-generation sequencing (NGS) specifically tailored to highly degraded ancient DNA (\cite{60}). This approach has been shown to recover DNA molecules that escape shotgun sequencing (\cite{15}) or sequence capture (\cite{61}) and, if primers are optimized \textit{in silico} and \textit{in vitro}, is highly locus-specific (\cite{60}). Used in combination with methods minimizing contamination (\cite{62}), as well as statistically sufficient replications (\cite{60}) (see also Materials & Methods section as well as Supplementary Materials), it at least equals or can occasionally be superior to DNA capture methods, which are plagued by biases (e.g., \cite{63}). We have employed it for the study of various species, including horses (\cite{15, 60, 64, 65}).

Here, we analyzed 111 equid remains from eight sites in Central Anatolia as well as six sites in the Caucasus dating from the Early Neolithic to the Iron Age (ca. 9,000 - 500 BCE, see TabS1), with a few samples dating to later, historic periods. This approach had been developed and optimized previously to produce reliable data from highly degraded samples (\cite{60, 64, 65}), and it has been used successfully in situations where shotgun sequencing was not effective enough to genotype a large proportion of phenotype-associated SNPs (\cite{15}). We targeted the mitochondrial hypervariable region as well as 18 specific single nucleotide polymorphism (SNP) regions diagnostic for the 18 major mitochondrial haplogroups considered diagnostic in earlier studies (\cite{47} and TabS2). These SNPs are sufficiently diagnostic to recapitulate the essential features of the mitogenome phylogeny (Fig.1). Moreover, we analyzed six regions of the Y chromosome, four anonymous Y-linked fragments and two fragments of the amelogenin gene to evaluate male inheritance (\cite{54, 56, 57} and TabS3). Finally, we chose a set of eight diagnostic SNPs in seven genes associated with the coat color in horses, including basic colors (bay, black, chestnut and grey), diluted phenotypes (silver and cream), spotted or painted phenotypes (overo, tobiano and sabino) and leopard spotting (\cite{11}; TabS4).

Genotyping versus osteological determination

From the 111 analyzed equid remains, 77 (70\%) yielded ancient DNA results and could be genotyped in independent triplicate PCR experiments (TabS1). We obtained 14 different caballine (\textit{E. caballus}) mitochondrial haplogroups previously defined in present-day horses (\cite{47}), as well as a new haplogroup, herein termed X, that belongs to the OPQ subtree (Fig 1, Tab S2). Moreover, from ten specimens we obtained haplotypes characteristic of donkeys (\textit{E. asinus}). Finally, seven specimens yielded haplotypes belonging to \textit{E. hemionus clade H1}, which Bennett et al. assigned to \textit{E. hemionus hydruntinus} (\cite{41}) and are therefore referred to as ‘hydruntine’ below.

Genotyping and osteological determination agreed in 48 of the 57 remains that were assigned osteologically, with various degrees of certainty, to one of the equid species (84\%) (TabS1). In particular, 38 of the 40 remains assigned osteologically to wild or domestic horses showed the corresponding mitochondrial DNA (95\%). Agreement was also obtained for three out of six hydruntine and seven out of
eleven donkey remains. Of the twenty remains that could not be assigned osteologically to one of the aforementioned species, we identified sixteen horses, two donkeys, and two hydruntines through mitochondrial DNA typing. Genotyping and osteological determination disagreed in only four osteologically unambiguously assigned cases: two bone specimens determined osteologically as hydruntines yielded horse mitochondrial DNA, whereas two others classified osteologically as horses carried a donkey and a hydruntine mitochondrial DNA. Finally, one unassigned equid was determined genetically as a hybrid, more precisely a mule, since it carried horse mitochondrial and donkey Y chromosomal DNA (TabS1: specimen CD6189).

Diachronic pattern of maternal lineages

The 12 Anatolian horse remains predating ~4500 BCE carried either the mitochondrial haplogroup P or a new mitotype, termed herein X, not previously identified in modern or ancient horses (Fig. 1-2; TabS1-S2). Neither the P nor the X haplogroup have been documented so far beyond Anatolia in contemporaneous or older samples (48, 50, 66-69). This strongly suggests that these two haplogroups represent the unique signature of a local wild horses native to the Anatolia plateau. After 2,200 BCE, this pattern changed profoundly, with 13 new mitochondrial haplogroups appearing in faunal assemblages from the Bronze and Iron Ages (Fig. 2; TabS1). Among the Bronze and Iron Age specimens, the pre-Bronze Age haplogroup P represents only 6% of the obtained haplotype spectrum (2 out of 33 remains), with both specimens dating to the earliest phase of this period (c. 2,000 BCE). Moreover, in post-3,300 BCE specimens, haplotype X is no longer detected. The novel haplotypes detected in our archaeological sample correspond mainly to haplogroups Q (11 remains), G (5 remains), and N (5 remains), while haplogroups A, B, D, E, H, I, L, Q account for the remaining 20 specimens (TabS1). These results indicate a nearly complete population turnover from the late third millennium BCE onward and corresponds well with iconographic and textual evidence for the appearance and dispersal of horse management in Anatolia and Mesopotamia (26, 28).

In the Caucasus, the earliest specimen that yielded a genetic result dates to the third millennium BCE and corresponds to haplogroup Q. Of the remaining 13 specimens, all excavated from archaeological contexts dating to the second millennium BCE, 11 represent a diverse array of haplogroups including A, B, C, E, FG, G, and Q. Taken together, this change in haplotypes in both Anatolia and the Caucasus is statistically highly significant (Fisher’s test p-value 5.7x10⁻⁶). The remaining two samples were identified as haplogroup P, presumably representing a continuation of the native Anatolian matriline into the Late Bronze Age.

Paternal lineages and hybrids

Paternal lineages were genotyped through six different loci on the Y chromosome (TabS3). Y chromosomal DNA data are less numerous and none were obtained from remains older than the Bronze Age, which must be due to poor DNA preservation. In 19 specimens dating to the Bronze Age or subsequent periods, however, the Y chromosome haplotype could be determined. Of these, 12 belong to E. caballus and six to E. asinus, and one specimen that was identified more generally as asinine (TabS1 and S3).

We could attribute the horse Y chromosomal sequences to two of a total of four horse haplotypes that have been described previously (57): five remains were carriers of haplotype Y-HT-1, which is the major
haplogroup present in modern horses, while four carried the extinct haplotype Y-HT-3 and three could not be determined due to SNPs that did not yield sufficient sequence coverage.

One specimen originating from Çadır Höyük yielded Y chromosomal SNPs corresponding to a jack ass, whereas the mitochondrial DNA corresponded to a horse (TabS1 and S3), thus reflecting the presence of a hybrid (mule) dating to the Iron Age. Interestingly, the mitotype of this individual was L, a mitotype not encountered in SWA before the Bronze Age.

Coat color

We genotyped SNPs associated with coat color variations (11, 58). As discussed above, retrieval of nuclear DNA data in addition to mtDNA requires better ancient DNA preservation. Therefore, as for the Y chromosome, nuclear SNPs could not be genotyped in a reliable manner in samples predating the Early Bronze Age (TabS1 and S4). Altogether, we obtained SNPs from 43 specimens, allowing us to infer the coat color for 33 individuals including, 25 horses, six donkeys, one hydruntine, and one mule. In our data set, we identified the mutant allele for all but two out of the eight interrogated SNPs. In particular, only the mutant alleles for overo and cream were missing while the other six genetic variants were present. Consequently, a large part of the diversity of mutations affecting the coat color already observed in ancient northern Eurasia (11) proved present in Bronze Age horses in SWA (TabS4). Our results allowed us to attribute a coat-color to 25 horses from the Bronze Age and later periods. We identified seven horses with a wild type bay-colored coat, one with a bay-sabino, eight with a chestnut-colored coat, two each with the colors chestnut-tobiano, chestnut-silver, leopard, and black, one with a Bay-tobiano coat-color and one specimen whose DNA preservation was not good enough to discriminate between chestnut and bay (TabS1 and S4). This diversification in the coat color distribution is statistically significant (Fisher’s test p-value 1.25x10⁻³).

As expected, the six donkeys and the hydruntine did not harbor any of the mutant SNPs that humans selected for in domestic horses. The sample that was identified as a mule carried one mutant allele in both the ASIP and MC1R genes, most likely originating from its horse mother, which are associated with a bay-tobiano coat in horses (TabS1 and S4). Importantly, in specimen AC8811 from Early Bronze Age Acemhöyük, a chestnut coat color is combined with mitotype P, representing the local Anatolian wild horse matriline. This combination indicates that local Anatolian mares were incorporated into domestic herds in the Early Bronze Age.
Discussion

Wild and domesticated horses in Anatolia

Our genotyping of 60 ancient horses was designed to elucidate the long-standing question of a local domestication of horses in Anatolia. Our results allow us to conclude that domestic horses were introduced into the Caucasus and Anatolia by at least 2,000 BCE, presumably from the Eurasian steppes. This conclusion is based on the fact that in Anatolia, local horse populations prior to ca. 4,500 cal BCE carried only two mitochondrial haplogroups, P and X, the latter being a previously unrecorded haplotype which belongs to the O-P-Q subtree (Fig. 1). So far, these haplogroups have not been encountered elsewhere in Eurasia in contemporaneous or earlier contexts. Furthermore, haplotype X, likely had a limited temporal occurrence in Holocene Anatolia possibly disappearing after 5500 BCE. The foregoing supports our conclusion that these two haplogroups reflect the local mitochondrial signature of wild horses hunted in Anatolia in the early and middle Holocene (Fig. 2). We propose that the P and X haplogroups evolved independently in Anatolia during the late Pleistocene and early Holocene with little or no gene flow from neighboring wild horse populations due to geographic barriers separating Anatolia from northern Eurasia, namely the Bosporus, as well as the Caucasus and the Zagros mountain ranges. Thus, our study provides the first evidence showing that Anatolia was home to a genetically distinct population of wild horses, which based on archaeozoological findings were widely exploited during the Neolithic and Chalcolithic periods (20, 22, 42). Equid remains from early Neolithic Aşıklı Höyük, Neolithic/Chalcolithic Köşk Höyük, and Late Chalcolithic/Early Bronze Age Çadır Höyük are representative of these wild Anatolian horses.

Around 2,000 BCE, we observe a statistically significant decline in the frequency of this local wild horse mitochondrial signature as the P haplogroup becomes rare and the X haplotype disappears completely. Presumably the low frequency of haplotype X in pre-Bronze Age horse assemblages (2 out of 11 in our dataset) offers an explanation as to why it did not survive in Anatolia into historic times.

Parallel to this, the diversity of maternal lineages in archaeological horse remains from the Caucasus and Anatolia increased drastically from two to fourteen (Fig. 2), all of which identified previously in present-day horses (47) as well as in Eneolithic, Chalcolithic and Early Bronze Age horses from SE Europe and Kazakhstan (50). In these studies, mitochondrial haplogroups display no phylogeographic structure in Eurasia, which is consistent with the absence of significant physical barriers across the vast Eurasian steppes. This suggests that horse populations in the Eurasian steppes were panmictic likely explaining the high diversity of present-day domestic horse populations. It would also account for the rapid diversification observed in our dataset upon introduction in the Caucasus and Anatolia. This sudden appearance of allochthonous lineages coincides with the emergence of iconographic and epigraphic evidence for horses and horse riding at the end of the third millennium BCE and argues for substantial imports of domestic forms and hence against an independent local domestication process (26-28).

Our results obtained from genotyping of alleles associated with variations in coat color reinforces this conclusion, since Bronze Age horses in Anatolia and the Caucasus show mutations corresponding to coat color variants thought to have been selected during domestication in the Eurasian steppes, such as chestnut, black, and silver (11). Mutations associated with coat color dilutions or spotting appeared later in our data set (after 1,200 BCE). Thus, we conclude from our genetic data that the domestic horses
introduced to Anatolia in the Bronze Age carried mutations found earlier in northern Eurasia (11) and therefore predominantly derived from stocks imported from this vast region. Although the ultimate geographic origins of this allochthonous population cannot be defined with the data at hand, the Eurasian steppes north of the Black Sea seems a most plausible candidate.

Interestingly, we found that the local Anatolian P haplotype persisted in Bronze Age domestic horses in Anatolia and the Caucasus at low frequency (8% in our dataset) as well as in present-day horses, suggesting that wild Anatolian mares were incorporated into domestic herds probably very soon after the introduction of domestic horses in the region and prior to their local extinction in the wild. This is in agreement with our observation that at least two out of the four Bronze Age horses carry the P haplogroup. Since these individuals also carry a mutant coat color allele (AC8811 and TS2, Supp. Table 1), it is most likely that they represent domestic horses whose maternal ancestors were recruited from local wild Anatolian populations, which follows a pattern observed elsewhere whereby local mares were proposed to be recruited into domestic herds resulting in high mitogenomic diversity (48, 50). Our results suggest a rapid shift from hunting to herding following the introduction of domestic horses in the Early Bronze Age, which likely correlates with a terminal decline in the wild horse population in Anatolia and the Caucasus indicated by the low frequencies of their remains in faunal assemblages (70, 71).

The diachronic pattern in Anatolia and the Caucasus of the two Y chromosomal haplotypes, Y-HT-1 and Y-HT-3, hints at the possibility of a population dynamics similar to that documented for northern Eurasia. In this latter region, Y-HT-1 constituting the predominant Y haplotype in present-day horses, drastically increased after the onset of domestication reaching fixation in the gene pool of the domestic horse by the Middle Ages, while Y-HT-3 declined over time until its disappearance (57). In our Anatolian dataset, both haplotypes occurred at comparable frequencies around 2,000 BCE, but here too, Y-HT-3 declined over time with the most recent horse that carried Y-HT-3 originating from the Caucasus at around 1,300 BCE. This reduction of Y chromosome diversity is presumably the result of strong selection of stallions (16).

Since northern Eurasia, and in particular the Pontic-Caspian steppe, is currently the most likely origin for the domestic horses brought into Anatolia, there are two possible introductory routes, one via southeastern Europe and one via the Caucasus. The route across the Bosporus has been postulated on the basis of the earliest zooarchaeological evidence for domestic horses in the southern Balkans at the Early Bronze Age site of Kanlıçeğit around 2,600-2,300 BCE (8). The coat colors of ten horses from this site were genotyped and revealed a highly biased distribution of coat color mutations with six out of ten homozygous black horses (a/a) four of which also show the leopard spotting (LP), plus two bay-colored horses with leopard spotting; no chestnut mutation was detected (72). This pattern strongly contrasts with our results in Anatolia and the Caucasus, where chestnut (e/e) is the earliest coat color variant, while black and leopard mutants remain very rare (only 2/25). These differences argue against the introduction of a domestic horse population similar to that found at Kanlıçeğit. Moreover, there is no archaeological evidence for horse management in western Anatolia in the third millennium BCE, providing little additional support for the hypothesis of an early introductory route across the Bosporus (73, 74).

In contrast, our identification of several allochthonous mitochondrial lineages as well as coat color mutations appearing broadly contemporaneously in the Southern Caucasus and in central Anatolia argues in favor of a dispersal route via the Caucasus. Indeed, the abundance of horse bones and images of horses in Maikop culture settlements and burials of c. 3300 BCE in the northern Caucasus led to the suggestion that horseback riding began in the Maikop period (75). In addition, recent studies of ancient human
genomes showed continuous gene flow between Copper Age steppes and Caucasus peoples (3, 4, 76, 77), and later, during the Bronze Age, between Mesopotamia, Anatolia, the southern and northern Caucasus and the steppes (78). This exchange between human groups appears to intensify during a century long period of cooling and desertification, known as the 4.2 ky event (e.g., (79) but see (80)) which may have impacted subsistence strategies and social networks in the steppe zone (e.g., (78, 81)). On present evidence, this climatic event seems to be broadly contemporaneous with the arrival of non-local horse mitochondrial haplogroups and coat colors in the Caucasus and Anatolia and linked to the expansion of horse husbandry and possibly Indo-European languages (e.g., (14, 75), but see (82)). Although the cultural processes initiating the dispersal of horse husbandry south of the Caucasus are currently difficult to address, it may relate to human population movements into the Caucasus and subsequently into Anatolia beginning in the late third millennium BCE.

Other equids

The present study identified a mule, offspring of a female horse and a male donkey, from an Early Iron Age context (dated to 1,100 – 800 BCE) at central Anatolian Çadır Höyük. Mules have been identified based on osteological criteria and more recently on genomic data in the European Iron Age and Roman periods (16, 83-85). Although several specimens have been tentatively identified as mules from Bronze and Iron Age contexts in SWA (86-88), this is the most ancient genomic evidence for a mule in SWA. Since wild donkeys are not native to Anatolia, the jack ass involved must have been a domestic animal (29, 71). The parent mare of the mule from Çadır Höyük was also a domestic animal as it carried an allochthonous mitochondrial haplogroup and contributed two coat color mutant alleles to its offspring. Although not a surprise, the presence of a mule in the Anatolian Iron Age, reflects a new role for domestic horses that emerged as they moved into SWA where domestic donkeys had been utilized since the fourth millennium BCE (89) and where a tradition of equid (donkey x hemione) hybridization emerged in the third millennium BC (28). This situation reflects a true integration of horses into SW Asian equid economies and is an early example of intentional livestock engineering. Being the most expensive species mentioned on a price list for livestock dating to the Hittite period (~1,600 - ~1,178 BCE (90), mules were obviously highly valued.

In addition to horses, our results provide unambiguous evidence that Neolithic, Chalcolithic and Bronze Age people in Anatolia also hunted hydruntines, the hemione subspecies E. hemionus hydruntinus, once inhabiting large parts of Anatolia (41). This is not a trivial result since the osteological identification of the hydruntine is not obvious (41, 59), an observation confirmed in our study illustrating disagreement between genotyping and osteological determination in six cases. At the same time it also confirmed the results of another genetic study presented by us illustrating that no other subspecies of hemione populated the Anatolian plateau (41). As such, the most recent evidence in our data set for an Anatolian hydruntine dates to ~2,200 BCE, which is in agreement with results reported earlier (41). The combined data thus suggest that hydruntines went extinct in Anatolia during the Late Bronze Age more or less at the same time as local wild horses perhaps in response to the same combination of factors including increased aridity associated with the 4.2 ky event, competition for pasture resources with growing numbers of livestock, and hunting pressure perhaps related to practices of elites hunting (28).
Conclusions

Our study of ancient equid remains from Anatolia and the southern Caucasus covering ~9,000 years of the Holocene analyzed the dynamics over time of mitochondrial lineages and tested the hypothesis of Anatolia as a center of horse domestication. We were able to identify two new mitotypes characteristic of local Anatolian wild horses, which were regularly exploited in the early and middle Holocene. However, we identified a pattern of genetic change that does not reflect a gradual process involving the local population but rather a sudden appearance ~2,000 BCE of non-local lineages that are still present in domestic horses. We also show that these imported horses exhibited coat colors that are absent in local wild horses prior to domestication. Moreover, continuation of Anatolian maternal lineage P into the Bronze Age implies some limited incorporation of local mares into domestic herds. These patterns of change indicate that domestic horses were introduced into Anatolia perhaps via the Caucasus region during the Bronze Age and provide, for the first time, a date for the beginning of the exploitation of domestic horses in Anatolia and Transcaucasia. They also argue against local independent domestication of the horse in this region. Indeed, our results strongly suggest that Anatolia was not a primary source for domestic horse lineages but, as observed in other regions, local matrilines were incorporated into herds of imported domestic horses, which were also hybridized with local donkeys to create mules. The ultimate geographic origins of the imported domestic herds remain to be determined but by eliminating Anatolia directs further attention to the adjacent regions of the Black Sea.
Material and Methods

Archaeological bones

The provenance of the analyzed horse specimens is described in the Supplementary Materials paragraph 1 and table S1.

Genetic analysis

Sample preparation, DNA extraction, DNA purification and following pre-PCR procedures were carried out in the high-containment facility of the Jacques Monod Institute in Paris physically separated from areas where modern samples are analyzed and dedicated exclusively to ancient DNA analysis using the strict procedures for contamination prevention previously described (62, 91). Bone specimens were cleaned and ground in a freezer mill, and DNA extracted and purified as previously described (61, 92) (also see Supplementary Materials paragraph 2.1.). Genetic analyses were performed using aMPlex-Torrent, a multiplex PCR assay coupled to NGS optimized to genotype reliably highly damaged ancient DNA at high throughput (60, 64, 65). The approach relies on careful design of primers that were optimized in silico and in vitro and on the systematic replication in triplicates of each PCR (60 and Supplementary Materials). Moreover, we used it in combination with rigorous methods of elimination of DNA molecules contaminating commercially purchased reagents (62), and of previous PCR amplification products with UDG treatment (93), which also removes cytosine deamination products (94). This approach proved extremely useful to retrieve selected markers more efficiently and at much lower costs than with shotgun sequencing (15). Compared to DNA capture approaches, it provides a higher yield since capture approaches tend to leave gaps in the genomic regions that are targeted, which then have to be filled via PCR approaches (61; supplementary information).

To amplify the mitochondrial genome, primers were designed in order to cover the complete mitochondrial genome with minimal primer dimer propensity from a multiple alignment of all the equids mitogenomes sequences present in Genbank in 2014 using the software Oligo 7 as previously described (60). The primers were then tested for efficiency and dimer formation using quantitative real-time PCR (qPCR) and optimized as described before (60) and combined in three different multiplex reactions (see SI paragraphs 2.2.1., 3.1. and 3.4.1.). We proceeded in the same way to analyze the other genetic markers. To analyze the coat color, we chose a set of eight SNPs in seven genes for detecting basic colors (bay, black, chestnut and grey), diluted phenotypes (silver and cream), spotted or painted phenotypes (overo, tobiano and sabino) and leopard spotting, as described previously (11) (see SI paragraphs 2.2.2., 3.2. and 3.4.2.). For the analysis of the Y chromosome, we selected four anonymous Y-linked fragments, Y2B17, Y3B1, Y3B12, Y3B19 and two fragments of the amelogenin gene, AME2 and AME3, as previously described (54) (see SI paragraphs 2.2.3., 3.3. and 3.4.3.). To protect against cross-contamination between samples, we used the UNG-coupled PCR system (see Supplementary Materials paragraphs 2.2.1.-2.2.3.). For each sample, we then performed the six multiplex PCRs with different primer combinations (mtDNA, coat color and Y chromosome) (see SI paragraphs 3.4.). Each extract was amplified in triplicate and the triplicates were pooled so that, at the end, there were three pools for each extract. The PCR products were pooled in a 96-well plate and DNA libraries prepared in an automate Tecan Freedom Evo 100 ligating sample-specific Ion Torrent barcoded adaptors (see Supplementary Materials paragraph 2.3.). After amplification
the size distribution and concentration of the library was assessed on the Agilent 2100 Bioanalyzer. Emulsion PCR and Ion Sphere Particle enrichment were conducted with the Ion OneTouch System (Life Technologies) using the Ion OneTouch 200 Template kit v2 DL (see Supplementary Materials paragraph 2.3.). The DNA library was sequenced on the Ion Torrent Personal Genome Machine (PGM) Sequencer using the Ion PGM 200 Sequencing Kit and Ion 314 semiconductor sequencing chips (Life Technologies) (see Supplementary Materials paragraph 2.3.). Consensus sequences were established typically from several tens to several hundreds of sequences, and only replicated sequences were considered herein.

**Phylogenetic analysis of the mitochondrial DNA**

A maximal likelihood phylogenetic tree was constructed from aligned complete *Equus* mitogenomes using RAXML (95) and a general time reversible nucleotide substitution model, with gamma categories and an estimated proportion of invariant sites. Following reduction of the mitogenomes to the fragment amplified by PCR after primer removal, a phylogenetic tree was constructed using PHYML (96) with four gamma categories and an estimated proportion of invariant sites and a Tamura-Nei substitution model (97).

**Statistical analysis**

The probability (two-tailed p-values) that the changes of mitochondrial DNA haplotype and coat color diversity in Anatolia and the Caucasus after ~2000 BCE could have been observed simply by chance was estimated with the Fisher’s exact test using as nominal variables time (before and after 2000 BCE) and genotype (OPQX vs other haplogroups for mitochondrial haplotypes, Bay vs other colors for coat colors).

Supplementary Material for this article is available at
Supplementary Materials and Methods
Supplementary Tables S1 – S4:
TabS1: samples analyzed with dates and results
TabS2: List of SNPs used for determination of the mitochondrial DNA haplogroups
TabS3: List of diagnostic SNPs for the determination of horse mitochondrial DNA haplogroups in both the complete mitogenomes and in the reduced assay used in the present study
TabS4: List of diagnostic SNPs assayed to determine horse mitochondrial DNA haplogroups
TabS5: List of SNPs obtained from the Y-chromosome assay
TabS6: List of genotypes obtained from the coat color assay
Supplementary Figure S1: Map showing the locations of the archaeological sites in which the analyzed specimens had been preserved
Supplementary Figure S2: Horses showing the coat colors evidenced in the present study

**References:**
1. R. Drews, *Early Riders: The beginnings of mounted warfare in Asia and Europe*. (Routledge, New York, 2004).
2. M. A. Levine, in *Late Prehistoric Exploitation of the Eurasian Steppe*, M. Levine, Y. Rassamakin, A. M. Kislenko, N. S. Tatarintseva, Eds. (Cambridge McDonald Institute, Cambridge, 2005), pp. 5-58.
3. W. Haak et al., Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* **522**, 207-211 (2015).
4. M. E. Allentoft et al., Population genomics of Bronze Age Eurasia. *Nature* **522**, 167-172 (2015).
5. S. Bökényi, The earliest waves of domestic horses in East Europe. *Journal of Indo-European Studies* **6**, 17-76 (1978).
6. N. Benecke, *Der Mensch und seine Haustiere*. (Konrad Theiss Verlag GmbH & Co, Stuttgart, 1994).
7. H.-P. Uerpmann, Die Domestikation des Pferdes im Chalkolithikum West- und Mitteleuropas. *Madrider Mitteilungen* **31**, 109-153 (2001).
8. N. Benecke, in *Equids in Time and Space: Papers in Honour of Véra Eisenmann*, M. Mashkour, Ed. (Oxford, 2006).
9. H. J. Greenfield, in *Horses and humans: The evolution of human-equine relationships*, S. L. Olsen, S. Grant, A. Choyke, L. Bartosiewicz, Eds. (BAR International Series 1560, Oxford, 2006), pp. 221-244.
10. A. K. Outram et al., The earliest horse harnessing and milking. *Science* **323**, 1332-1335 (2009).
11. A. Ludwig et al., Coat color variation at the beginning of horse domestication. *Science* **324**, 485 (2009).
12. D. W. Anthony, D. Brown, Bit wear, horseback riding, and the Botai site in Kazakhstan. *Journal of Archaeological Science* **25**, 331-347 (1998).
13. S. J. Olsen, in *Prehistoric steppe adaptation and the horse*, M. Levine, C. Renfrew, K. Boyle, Eds. (McDonald Institute Monographs, Cambridge, 2003), pp. 83-102.
14. C. Gaunitz et al., Ancient genomes revisit the ancestry of domestic and Przewalski’s horses. *Science* **360**, 111-114 (2018).
15. P. Librado et al., Ancient genomic changes associated with domestication of the horse. *Science* **356**, 442-445 (2017).
16. A. Fages et al., Tracking Five Millennia of Horse Management with Extensive Ancient Genome Time Series. *Cell* **177**, 1419-1435 e1431 (2019).
17. D. Anthony, *The Horse, the Wheel, and Language: How Bronze-Age Riders from the Eurasian Steppes Shaped the Modern World*. (Princeton University Press, 2007).
18. B. S. Arbuckle, Animals and inequality in Chalcolithic Anatolia. *Journal of Anthropological Archaeology* **31**, 302-313 (2012).
19. N. S. Shaler, *Domesticated animals. Their relation to man and to his advancement in civilization*. (Charles Scribner’s Sons, New York, 1895).
20. L. Martin, N. Russell, in *Horses and humans: the evolution of human-equine relationships*, S. J. Olsen, Ed. (British Archaeological Reports International Series 1560, Oxford, 2006), pp. 115-126.
21. A. Ünal, Eski Anadolu’da At Hitiçe Kikkülü at eğitimi metinleri ve “Tavlaya çekmek”le ilgili teknik bir ayrıntı. *Çorum Kültür Sanat: Bilim, Kültür, Sanat, Tarih, ve Turiz, Dergisi* II, 40-66 (2013-2014).
22. B. S. Arbuckle, A. Öztan, in *Archaeozoology of the Near East XII*, C. Çakırlar, J. Chahoud, R. Berthon, S. E. Pilaar, Eds. (Groningen, 2018), pp. 42-59.
23. J. M. C. Toynbee, M. R. Alföldi, D. Misselbeck, *Tierwelt der Antike - Bestiarium romanum*. (Verlag Philipp von Zabern, Mainz am Rhein, 1983).
24. J. Boessneck, A. von den Driesch, Pferde im 4/3. Jahrtausend v. Chr. in Ostanatolien. *Säugetierkundliche Mitteilungen* **24**, 81-87 (1976).
25. M. A. Littauer, J. H. Crouwel, *Wheeled Vehicles and Ridden Animals in the Ancient Near East*. (Brill, Leiden, 1979).

26. D. I. Owen, The first equestrian: An Ur III glyptic scene. *Acta Sumerologica* **13**, 259-273 (1991).

27. J. Oates, in *Prehistoric Steppe Adaptation and the Horse*, M. Levine, C. Renfrew, K. Boyle, Eds. (McDonald Institute Monograph, Cambridge, 2003), pp. 115-125.

28. J. Zarins, *The Domestication of Equidae in Third-Millennium BCE Mesopotamia*. Cornell University Studies in Assyriology and Sumerology (Cornell University, Ithaca, NY, 2014), vol. 24.

29. J. Clutton-Brock, *Horse Power. A History of the Horse and the Donkey in Human Societies*. (Harvard University Press, 1992).

30. S. Bökönyi, in *Die Indogermanen und Das Pferd*, B. Hänsel, S. Zimmer, Eds. (Akten Internationalen Interdisziplinären Kolloquiums Freie Universität Bernfried Schlerath, Budapest, 1994), pp. 115-122.

31. F. E. Zeuner, *A History of Domesticated Animals*. (Hutchinson, London, 1963).

32. H.-P. Uerpmann, *Domestication of the horse: when, where, and why*. L. Bodson, Ed., Colloque d’histoire des connaissance zoologiques (Le cheval et les autres équidés: aspects de l’histoire de leur insertion dans les activités humains) (University of Liege, Liège, Belgium, 1995).

33. A. Sherratt, in *Prehistoric steppe adaptation and the horse*, M. Levine, C. Renfrew, K. Boyle, Eds. (McDonald Institute Monographs, Cambridge, 2003), pp. 233-252.

34. D. W. Anthony, *The horse, the wheel, and language: How Bronze-Age riders from the Eurasian steppes shaped the modern world*. (Princeton University Press, Princeton, NJ., 2007).

35. C. Becker, in *Die Indogermanen und Das Pferd*, B. Hänsel, S. Zimmer, Eds. (Akten Internationalen Interdisziplinären Kolloquiums Freie Universität Bernfried Schlerath, Budapest, 1994), pp. 145-177.

36. H.-P. Uerpmann, The Ancient Distribution of Ungulate Mammals in the Middle East. *Beihfte zum Tübinger Atlas des Vorderen Orients Reihe A (Naturwissenschaften) Nr. 27*, Wiesbaden, Dr. Ludwig Reichert, (1987).

37. S. Bökönyi, in *Equids in the Ancient World Volume II*, R. H. Meadow, H.-P. Uerpmann, Eds. (Dr. Ludwig Reichert Verlag, Wiesbaden, 1991), vol. Beihfte zum Tübinger Atlas des Vorderen Orients Reihe A (Naturwissenschaften) 19/2, pp. 123-131.

38. N. Benecke, in *Archaeozoology of the Near East III. Proceedings of the third international symposium on the archaeozoology of southwestern Asia and adjacent areas*, H. Buitenhuis, L. Bartosiewicz, A. Choyke, Eds. (ARC Publication 18, Groningen, 1998), pp. 172-179.

39. E. Vila, in *Equids in time and space: Papers in honour of Vera Eisenmann*, M. Mashkour, Ed. (Oxbow Books, Oxford, 2006), pp. 102-123.

40. B. S. Arbuckle, Zooarchaeology at Acemhöyük. *Anadolu/Anatolia* **39**, 55-68 (2013).

41. E. A. Bennett *et al.*, Taming the late Quaternary phylogeography of the Eurasian wild ass through ancient and modern DNA. *PLoS One* **12**, e0174216 (2017).

42. H. Buitenhuis *et al.*, in *The Early Settlement at Aşıklı Höyük; Essays in Honor of Ufuk Esin*, M. Özbaşaran, G. Duru, M. C. Stiner, Eds. (Ege Yayınları Istanbul, 2018), pp. 281-323.

43. S. Payne, in *Equids in the Ancient World*, R. H. Meadow, H. P. Uerpmann, Eds. (Dr. Ludwig Reichert Verlag, Wiesbaden, Tübingen, 1991), vol. II, pp. 132-165.
44. H.-P. Uerpmann, in *The salvage excavations at Orman Fidanlığı: A Chalcolithic site in inland northwestern Anatolia.* , T. Efe, Ed. (TASK Vakfı Yayınları, Istanbul, 2001), pp. 187-210.
45. B. S. Arbuckle, Zooarchaeology at Kösk Höyük. *Kazı Sonuçları Toplantısı* 27, 124-136 (2008).
46. B. S. Arbuckle, Chalcolithic caprines, Dark Age dairy and Byzantine beef. *Anatolica* 35, 179-224 (2009).
47. A. Achilli et al., Mitochondrial genomes from modern horses reveal the major haplogroups that underwent domestication. *Proc Natl Acad Sci U S A* 109, 2449-2454 (2012).
48. C. Vila et al., Widespread origins of domestic horse lineages. *Science* 291, 474-477 (2001).
49. T. Jansen et al., Mitochondrial DNA and the origins of the domestic horse. *Proc Natl Acad Sci U S A* 99, 10905-10910 (2002).
50. M. Cieslak et al., Origin and history of mitochondrial DNA lineages in domestic horses. *PLoS One* 5, e15311 (2010).
51. J. Lira et al., Ancient DNA reveals traces of Iberian Neolithic and Bronze Age lineages in modern Iberian horses. *Mol Ecol* 19, 64-78 (2010).
52. S. Lippold, N. J. Matzke, M. Reissmann, M. Hofreiter, Whole mitochondrial genome sequencing of domestic horses reveals incorporation of extensive wild horse diversity during domestication. *BMC Evol Biol* 11, 328 (2011).
53. V. Warmuth et al., Reconstructing the origin and spread of horse domestication in the Eurasian steppe. *Proc Natl Acad Sci U S A* 109, 8202-8206 (2012).
54. G. Lindgren et al., Limited number of patrilines in horse domestication. *Nat Genet* 36, 335-336 (2004).
55. T. Kavar, P. Dovç, Domestication of the horse: Genetic relationships between domestic and wild horses. *Livest. Sci.* 116, 1-14 (2008).
56. S. Lippold et al., Discovery of lost diversity of paternal horse lineages using ancient DNA. *Nature communications* 2, 450 (2011).
57. S. Wutke et al., Decline of genetic diversity in ancient domestic stallions in Europe. *Science advances* 4, eaap9691 (2018).
58. M. Pruvost et al., Genotypes of predomestic horses match phenotypes painted in Paleolithic works of cave art. *Proc Natl Acad Sci U S A* 108, 18626-18630 (2011).
59. E. M. Geigl, T. Grange, Eurasian wild asses in time and space: Morphological versus genetic diversity. *Annals of Anatomy* 194, 88-102 (2012).
60. S. Guimaraes et al., A cost-effective high-throughput metabarcoding approach powerful enough to genotype ~44 000 year-old rodent remains from Northern Africa. *Mol Ecol Resour* 17, 405-417 (2017).
61. D. Massilani et al., Past climate changes, population dynamics and the origin of Bison in Europe. *BMC biology* 14, 93 (2016).
62. S. Champlot et al., An efficient multistrategy DNA decontamination procedure of PCR reagents for hypersensitive PCR applications. *PLoS One* 5, (2010).
63. A. Albrechtsen, F. C. Nielsen, R. Nielsen, Ascertainment biases in SNP chips affect measures of population divergence. *Mol Biol Evol* 27, 2534-2547 (2010).
64. N. M. Coté et al., A New High-Throughput Approach to Genotype Ancient Human Gastrointestinal Parasites. *PLoS One* 11, e0146230 (2016).
65. C. Ottoni et al., The palaeogenetics of cat dispersal in the ancient world. *Nature Ecology and Evolution* **1**, (2017).
66. J. Weinstock et al., Evolution, systematics, and phylogeography of Pleistocene horses in the new world: a molecular perspective. *PLoS Biol.* **3**, e241 (2005).
67. E. D. Lorenzen et al., Species-specific responses of Late Quaternary megafauna to climate and humans. *Nature* **479**, 359-364 (2011).
68. L. Orlando et al., Recalibrating Equus evolution using the genome sequence of an early Middle Pleistocene horse. *Nature*, (2013).
69. J. Elsner, M. Hofreiter, J. Schibler, A. Schlumbaum, Ancient mtDNA diversity reveals specific population development of wild horses in Switzerland after the Last Glacial Maximum. *PLoS One* **12**, e0177458 (2017).
70. J. Boessneck, A. v. d. Driesch, Tierknochen und Molluskenfunde aus Munbāqa. *Mitteilungen der Deutschen Orient-Gesellschaft* **118**, 147-160 (1986).
71. H. P. Uerpmann, *The Ancient Distribution of Ungulate Mammals in the Middle East.*, Beihefte zum Tübinger Atlas des Vorderen Orients. Reihe A (Naturwissenschaften) Nr. 27 (Dr. Ludwig Reichert Verlag, Wiesbaden, 1987).
72. A. Ludwig et al., Twenty-five thousand years of fluctuating selection on leopard complex spotting and congenital night blindness in horses. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **370**, 20130386 (2015).
73. H. P. Uerpmann, in *Troia and the Troad: Scientific approaches*, G. A. Wagner, E. Pernicka, H. P. Uerpmann, Eds. (Springer 2013).
74. C. Çak_rl_r, L. Atici, *Patterns of animal exploitation in western Turkey: from Paleolithic molluscs to Byzantine elephants*. U. Albarella, H. Russ, K. Vickers, S. Viner-Daniels, Eds., *The Oxford Handbook of Zooarchaeology* (Oxford University Press, Oxford, 2017).
75. D. W. Anthony, *The Horse, the Wheel and Language. How Bronze-Age Riders from the Eurasian Steppes Shaped the Modern World*. (Princeton University Press, Princeton and Oxford, 2007), pp. 847.
76. I. Lazaridis et al., Genomic insights into the origin of farming in the ancient Near East. *Nature* **536**, 419-424 (2016).
77. P. de Barros Damgaard et al., The first horse herders and the impact of early Bronze Age steppe expansions into Asia. *Science* **360**, (2018).
78. C. C. Wang et al., Ancient human genome-wide data from a 3000-year interval in the Caucasus corresponds with eco-geographic regions. *Nature communications* **10**, 590 (2019).
79. T. J. Szczęsny, Was the 4.2 ka Event an Anthropogenic Disaster? *Open Journal of Ecology* **6**, 613-631 (2016).
80. P. Voosen, New geological age comes under fire: Massive drought or myth? Scientists spar over an ancient climate event behind our new geological age. *Science* **361**, 537-538 (2018).
81. N. Shishlina, in *Counterpoint: Essays in Archaeology and Heritage Studies in Honour of Professor Kristian Kristiansen*, S. Bergerbrant, S. Sabatini, Eds. (Archaeopress, Oxford, 2013), pp. 53-60.
82. J.-P. Démoule, *Mais où sont passés les Indo-Européens ? Le mythe d’origine de l’Occident.* (Editions du Seuil, 2014), pp. 823.
83. P. L. Armitage, H. Chapman, Roman mules. *The London Archaeologist* **3**, 339-346 (1979).
Acknowledgments. We thank the directors of excavations Harald Haupmann † (Norşun Tepe), Mihriban Özbaşaran (Aşıklı Höyük) for access to samples as well as the Department of Archaeology at Bitlis Eren University, the Çorum Müzesi, Andreas Schachner and T. Emre Şerifoğlu for their support. Permission to export samples was granted by the Turkish Ministry of Culture and Tourism and the Aksaray Museum. Many thanks to Ruben Badalyan and Adam T. Smith for permission to analyze the equids from the Project ArAGATS excavations.

Funding. The study was supported by the National Science Foundation BCS-1311551. The paleogenomic facility of the Institut Jacques Monod obtained support from the University Paris Diderot within the program “Actions de recherches structurantes”. The sequencing facility of the Institut Jacques Monod,
Paris, is supported by grants from the University Paris Diderot, the Fondation pour la Recherche Médicale (DGE20111123014), and the Région Ile-de-France (11015901). Moreover, we acknowledge support from the French national research center CNRS, the National Geographic Society, the American Research Institute in Turkey, the University of North Carolina at Chapel Hill and Baylor University, as well as the Deutsche Forschungsgemeinschaft (DFG) to JP (PE 424/10,1-3). S.E.A. was supported by a Fulbright IIE Grant and the University of Chicago.

**Author contributions:** B.S.A. and E.M.G. initiated and coordinated the project. E.M.G. and T.G. designed research and supervised experiments; S.G. performed experiments; E.M.G., T.G. and S.G. analyzed data; E.M.G., T.G. and B.S.A. wrote the paper with input from J.P.

The authors declare no to have no competing interests.

Data and materials availability: All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors.
Figure 1: Comparative maximum likelihood phylogenetic analyses of horse mitogenome using either the complete mitogenome sequences (left side) or the concatenated mitogenome fragments used for genotyping ancient remains (right side). The nomenclature of the horse mitochondrial haplogroups from A to R is as defined by Achilli et al., 2012 (47). Haplogroup S corresponds to an additional haplogroup obtained when adding Przewalski’s horse sequences not belonging to the F haplogroup (98). The scale bar represents the number of nucleotide substitution per site as indicated. The branches separating horse, hemione and donkey sequences are not drawn to scale as indicated by the intersecting parallels. A magnified view of the O-P-Q subtree of the concatenated fragments is represented in the box on the right side. The magnified view reveals in addition the position of the X sequence found in two ancient Anatolian remains. The numbers by the nodes indicate their corresponding bootstrap values.
Figure 2: Mitochondrial and coat color diversity before (Upper part) and after (Lower part) 2,000 BCE. The left panel presents evolution of mitochondrial haplotype diversity of horses in Anatolia and the southern Caucasus. The right panels represent the evolution of coat color genetic diversity in these two geographic regions in the same time ranges. The area of the circles is proportional to the number of individuals present in each category.

Supplementary Materials

Ancient DNA shows domestic horses were introduced in the southern Caucasus and Anatolia during the Bronze Age.
1. Description of the archaeological sites

The locations of the archaeological sites are shown in Fig. S1.
1.1. Anatolia

Aşıklı Höyük

Hijlke Buitenhuis, Joris Peters

Aşıklı Höyük is an aceramic Neolithic settlement mound located next to the Melendiz River. It covers an area of 3.5-4 ha, rising 15.35 m above the Melendiz Ülain at the north and 13.16 m at the south. Under the direction of Prof. Dr. U. Esin (University of Istanbul), excavations were carried out from 1989 until 2003. Renewed excavations directed by Prof. Dr. M. Özbaşaran (University of Istanbul) started in 2009 and continued since then. Five levels, designated as 1-5 from top to bottom respectively, could be documented so far. The earliest settlement phase excavated thus far is represented by Level 5. Radiocarbon dates allow concluding that early site occupation dates to ~8,400 BCE (1, 2). The faunal remains of the site are currently analyzed by a team of zooarchaeologists including H. Buitenhuis, N. Munro, J. Peters, N. Pöllath, and M.C. Stiner. A most important finding is the fact that caprine management was already practiced by the site inhabitants (3-6) and developed further on during the occupation in the later levels to a full caprine husbandry. So far, two species of wild equids have been evidenced in the archaeofauna. Remains of wild horses (E. ferus) form a small part of this assemblage, whereas the majority belongs to hemione-like animals, likely E. hydruntinus (4, 7). The faunal assemblage is curated on site at the Aşıklı Höyük dighouse, Kızılıkaya village, Aksaray, Turkey. All specimens were exported with the permission of the Aksaray Museum and the Turkish Ministry of Culture and Tourism.

Acemhöyük

Benjamin Arbuckle

Acemhöyük is a large mound located at an elevation of 950 meters on the alluvial fan of the Melendiz river near the Tüz Gölü (Salt Lake) in Central Turkey. The mound is approximately 800 x 600 meters in dimension rising 20 meters above the surrounding Yeşilova plain and represents the remains of a large, fortified Bronze Age urban center. The site has been excavated since 1962 by Dr. N. Özgüç, Ankara University in association with the Turkish Historical Society and the General Directorate of Antiquities and Museums and by Prof. Dr. A. Öztan of the same institution since 1989 (8, 9). Acemhöyük consists of 12 levels with deposits representing Chalcolithic to Medieval occupations. The faunal assemblage at Acemhöyük is dominated by domestic sheep, goats, cattle and pigs with equids representing a tertiary resource (10, 11). The Acemhöyük faunal assemblage is curated on site at the Acemhöyük dighouse, Yeşilova, Aksaray, Turkey. All specimens were exported with the permission of the Konya Museum and the Turkish Ministry of Culture and Tourism. 24 equid specimens were submitted for ancient DNA analysis. Twenty of these samples derive from deposits stratigraphically assigned to levels XI, X, VII, V and IV representing the late Early Bronze Age (EBA) and earliest Middle Bronze Age (MBA) (associated radiocarbon dates c. 2100-1900 BCE). Many of these specimens were identified only to the genus level (Equus sp.) and could potentially represent wild horse (Equus ferus), domestic horse (Equus caballus), wild hemione/hydruntine (Equus hemionus hydruntinus), domestic donkey (Equus asinus), or one of several hybrids (horse x donkey, donkey x horse, donkey x hemione). Four of these specimens (AC14371, AC14379, AC14380, AC14427) derive from equid burials and are identified based on dental morphology and body size as...
donkeys or hemiones. An additional 7 of these late EBA/early MBA specimens were also identified as donkeys or hemiones rather than horses. One specimen, identified as *Equus sp.*, derives from level III, the Assyrian Colony period (c. 1700s BC). Two specimens (*Equus caballus*) are from Hellenistic contexts (level I), while the last specimen (*Equus caballus*) is thought to come from a pit perhaps dating to the Ottoman period.

Çadır Höyük

*Benjamin Arbuckle*

Çadır Höyük is a 32-meter-high mound located in the Kanak Su basin situated in the Yozgat Province of north central Turkey. The modestly sized 240 x 185-meter mound boasts an occupational history spanning 6000 years (ca. 5200 BCE to the thirteenth century CE). This sequence has been explored since 1993 by the Çadır Höyük Archaeological Project led by Ron Gorny and more recently Greg McMahon and Sharon Steadman (12, 13). The Çadır Höyük faunal assemblage is curated on site at the Çadır Höyük dighouse, Peyniryemez, Sorgun, Turkey. Specimens were exported with the permission of the Yozgat Archaeology Museum and the Turkish Ministry of Culture and Tourism. 21 specimens derived from throughout the Çadır Höyük occupational sequence were submitted for ancient DNA analysis. These include two of specimens (CD2017, CD2015) from the Deep Sounding which were recovered from a Middle Chalcolithic context with an associated radiocarbon date of 4500 cal BCE (Beta-418460). Due to this early date, these specimens are thought to represent wild Anatolian horse (*Equus ferus*). Another (wild?) horse specimen (CD1875) was recovered from a Late Chalcolithic context with an associated radiocarbon date placing it in the late fourth millennium BCE (Beta418461: 3350 cal BCE). Six specimens derive from Early and Late Bronze Age contexts; ten specimens come from Iron Age deposits and include horse, donkey and *Equus sp* identifications. Finally, two horse specimens derive from Byzantine deposits.

Boğazköy-Hattusa

*Benjamin Arbuckle*

Boğazköy-Hattusa is an UNESCO World Heritage site located in north central Turkey (Çorum province) that represents the remains of a major urban center dating to the Middle and Late Bronze Age (2000-1200 BCE). Boğazköy is perhaps best known as the capital city of the Hittite empire and has been excavated by the Deutsches Archäologisches Institut under the supervision of Dr. Andreas Schachner. One specimen was provided from Hattusa for ancient DNA analysis. This specimen, an incisor tooth likely representing a domestic donkey, derives from Late Bronze Age (Hittite) deposits (excavation year 2013, context 407, bag#820).

Köşk Höyük

*Benjamin Arbuckle*

The site of Köşk Höyük (originally Köşk-Pınar (14)) represents the remains of a small agro-pastoral village located on the eastern edge of the Bor Plain in central Turkey at an elevation of 1400 meters (15). The site consists of a shallow mound, about half a hectare in area, situated on the edge of a limestone ridge overlooking a permanent spring. The location of the site provides ready access to the resources of the Konya-Ereğli Plain to the west, the foothills and uplands of the Taurus Mountains to
the east and south, and the most direct route from the west and south to the obsidian sources of the nearby Melendiz Mountains. Excavations at Köşk began with a series of soundings done by Dr. Nimet Özgüç (Ankara University) in 1981, followed by a more extensive excavation campaign led by Dr. U. Silistreli, in association with the Nigde Museum from 1983 to 1992, and finally by Prof Dr. Aliye Öztan from 1995 to 2012 (15, 16). The stratigraphy of Köşk has been divided into prehistoric levels V-I. Stratigraphic levels V-II have been radiocarbon dated to between 6200-5500 cal BCE and represent the Early Chalcolithic period in central Anatolia. In this period the faunal assemblage indicates that although domestic sheep and goats are the most abundant mammalian taxa, wild taxa are abundant as well. Wild cattle, boar, as well as hare and foxes were hunted and trapped at the site. However, wild equids including both wild horse (*Equus ferus*) and European wild ass (*Equus hemionus hydruntinus*) are the most abundant wild taxa representing a surprising 15-38% of the fauna from EC contexts (17).

24 specimens from Köşk Höyük were sent for ancient DNA analysis. These specimens derive from the Early Chalcolithic levels IV and III representing the Early Chalcolithic period. 14 specimens were identified based on dental and skeletal morphology as Anatolian wild horse (*Equus ferus*); five specimens were identified as European wild ass (*Equus hemionus hydruntinus*); while the remaining five specimens were identified to genus as *Equus* sp (Arbuckle). All specimens were exported with the permission of the Niğde Archaeology Museum and the Turkish Ministry of Culture and Tourism.

*Troy*

**Hans-Peter Uerpmann**

The excavation of the site of Troy, located close to the Dardanelles under the hill of Issarlik in Western Turkey, has identified a succession of 9 towns and 46 occupational levels. The complex stratigraphy extends over 20 m. The bones used for the present study originate from three different levels, Troy I, Troy IV and Troy VI. Troy I is a small fortified village and corresponds to the first human occupation of the site during the Early Bronze Age. The bone samples analysed in the present study originate from Troy IV and were assigned to *E. caballus*. Troy IV is dated to the Middle Bronze Age and corresponds to the phase of decline of the town whose subsistence mode seems to have been based on agriculture and handcraft rather than trading. Troy VI corresponds to the golden age of the town at the end of the Bronze Age as witnessed by the construction of palaces. Only a fraction of the animal remains found during the ongoing excavations have been studied. Nevertheless, the archaeozoological data of more than 120,000 bone finds are stored in a data-base (18).

*Lidar Höyük*

**Joris Peters**

Southwest of the village Lidar (SE-Anatolia, Turkey) lies the impressive settlement mound of Lidar Höyük, measuring 200 by 240 m at the base and rising 23 m above the Euphrates River terrace. Situated near a major passage across the Euphrates River, the site of Lidar Höyük was already inhabited in the Early Bronze Age. Although occupation lasted until Medieval Times, the site was not occupied permanently, as the archaeological record shows. Excavations carried out between 1979 and 1986 under the direction of Prof. Dr. H. Hauptmann† produced a rich archaeofauna totalling 62,453 vertebrate remains (19). 3.8% of this assemblage
pertains to equids, remains of donkeys outnumbering those of horses and hemiones. Difficulties arose to assign some larger specimens either to (small) horse, mule or Asiatic wild ass. The 1987 excavation produced another sample of equid bone, from which specimens were selected for ancient DNA analysis, including a Bronze Age horse radius collected in Area S47.

1.2. Caucasus

Georgia

Didi Gora

Hans-Peter Uerpmann

This site in the province of Kachetia is located on the plain of the Alazani in Eastern Georgia between the Black and the Caspian Seas. In the past it was probably located on a river bank. It comprises a history of three millennia, from the Early Bronze Age (c. 2800 BCE) to the Early Iron Age (c. 1,000 BCE) and is one of the largest settlement hills during the Neolithic and Bronze Age in this region. Archaeological evidence points to a semi-nomadic, seasonal lifestyle of the inhabitants, to the exploitation of ore deposits and to extensive trading contacts with other regions in the Near East. The Bronze Age layers are poor in archaeological artifacts except for obsidian flints, but very rich in faunal remains.

These faunal remains, which were analyzed in the present study, are very well preserved suggesting a rapid burial. The genetically analyzed equid remains had been attributed to *E. caballus* based on the relative size of the bones and on their weight since the bones were heavily fragmented.

Tqisbolu Gora

The site of Tqisbolu Gora was occupied between 3,800 and 2,800 years ago and was thus partly contemporaneous to Didi Gora. The youngest layers of Didi Gora correspond to Iron Age layers in Tqisbolu Gora. Three equid species were identified on the two sites. The remains of domestic horses (*E. caballus*) are bigger than the remains of the small equids from the site. The latter could belong to *E. hemionus*, *E. hydruntinus* or to the domestic donkey *E. asinus*.

In the present study, four remains were analyzed genetically. Three of them had been attributed to *E. caballus* and one to *E. hydruntinus* based on the relative size of the bones.

Armenia

Lusakert

Nina Manaseryan

The analyzed horse remains originate from the Middle Palaeolithic Cave Lusakert 1 in the Hrazdan valley of Armenia.

Mets Sepasar

Nina Manaseryan

Mets Sepasar is a Bronze Age habitat and necropolis in Shirak. The mountain Mets Sepasar is located on the Shirak plateau in the northwest of Armenia. The archaeological studies, led by L.Yeganyan, started out from the foothills of the mountain. As a result of these excavations an Early Bronze Age
settlement has been discovered. In the middle of the monument, in a volcanic cone, a sanctuary was found showing traces of festivities accompanied by sacrifices.

**Tsaghkahovit and Gegharot**

**Hannah Chazin**

Gegharot and Tsaghkahovit are multi-phase sites located in the Tsaghkahovit Plain, Armenia and have been the site of ongoing archaeological research by the joint Armenian-American Project ArAGATS (21-23). Gegharot was occupied during both the Early and Late Bronze Age. The Late Bronze Age occupation consisted of a walled fortress and nearby cemetery. Tsaghkahovit was occupied during the Late Bronze Age, as well as during the Achaemenid period. The Late Bronze Age occupation at Tsaghkahovit consisted of a walled fortress, extra-mural room blocks, and an associated cemetery. Two of the equid specimens were from Late Bronze Age tombs (Gegharot Kurgan 1 and Tsaghkahovit BC12), and the remainder were from Late Bronze middens excavated in the Tsaghkahovit Residential Complex. Equid specimens from these sites were identified by Hannah Chazin and Belinda Monahan.

**2. Paleogenetic Analysis**

**Rationale of the experimental approach**

Due to the age of the archaeological samples and the climatic conditions of the area and to the fact that most of the bones that were appropriate to answer the scientific questions were not petrous bones, DNA was poorly preserved. Thus, we used a method that we have developed for poorly preserved archaeological specimens (24). This powerful multiplex PCR assay is highly sensitive when optimized primers are used and allows high-throughput genotyping of a large sample number since it is coupled to next-generation sequencing (24). We used it previously successfully to genotype various difficult ancient samples (24-27). This approach is more efficient and cost-effective than sequence capture, in particular for samples with highly degraded DNA, and is informative enough for the specific questions addressed in this study. Here, we optimized the primers *in silico* and *in vitro*, i.e., tested and optimized the PCR conditions prior to the analysis of the ancient specimens. We amplified the ancient DNA extracts in the presence of dUTP and degraded products of previous PCR amplifications using UDG (28). This UDG treatment eliminates in addition the products of cytosine deamination, thus increasing the accuracy of retrieved sequences (28). Moreover, prior to each PCR amplification, we eliminated contaminating DNA molecules in reagents using the UVD decontamination procedure (29). This multi-level decontamination procedure enabled us to produce reliable, authentic amplicons. Finally, we performed independently each PCR amplification three times and sequenced the bulk PCR products using NGS to determine the authentic sequences.

All pre-PCR procedures were carried out in the high containment facility of the Jacques Monod Institute in Paris physically separated from areas where modern samples are analyzed and dedicated exclusively to ancient DNA analysis using the strict procedures for contamination prevention previously described (29, 30).
2.1. DNA Extraction

The external surface of the specimens was removed with a sterile blade to minimize the environmental contamination. For each bone sample, roughly 0.2 g was ground to a fine powder in a freezer mill (Freezer Mill 6750, Spex Certiprep, Metuchen, NJ), which was then suspended in 2mL of extraction buffer containing 0.5M EDTA, 0.25M Di-sodium hydrogen phosphate (Na2HPO4), pH 8.0, 0.14M 2-mercaptoethanol and 0.25 mg/mL of proteinase K and incubated under agitation at 37°C for 48 hours. Blank extractions were carried out for each extraction series. Samples were then centrifuged and the supernatant was purified with a modified protocol of the Qiaquick PCR Purification Kit (Qiagen, Hilden, Germany) using the manifold device, as previously described (30-32).

2.2. Multiplex PCR Amplification

2.2.1. Mitochondrial DNA

Primers were designed in order to cover the complete mitochondrial genome with minimal primer dimer propensity from a multiple alignment of all the equids mitogenomes sequences present in Genbank in 2014 using the software Oligo 7 as previously described (24). The primers were then tested for efficiency and dimer formation using quantitative real-time PCR (qPCR) (24). We selected a total of 20 primer pairs (see paragraph SI 3.1.). To assess the phylogenetic consistency of our molecular assay with the selected fragments, we built a maximum likelihood tree with PHYML (33) for both the complete mitogenomes (GTR + G + I) and the concatenation of the 20 amplified fragments without primers (TN93 + G + I) (34). This confirmed the preservation of the salient features of the tree topology when using only the selected fragments (Fig. 1). We optimized the multiplex PCR as described before (24) and the primer pairs were combined in three different multiplex reactions (see paragraph SI 3.4.1.). To protect against cross-contamination between samples, we used the UNG-coupled PCR system [40]. Each reaction was carried out in a 30 µl reaction volume with a final composition of 50 mM Tris/HCl, 10 mM KCl, 5 mM (NH4)2SO4, 4 mM MgCl2, pH 8.3 (@25°C), 1 mg/ml BSA (bovine serum albumin), 0.15 µM of each primer, 0.25 mM dA/G/CTPs and 0.50 mM dUTP, 2 U of Fast Start DNA polymerase, 0.5 U of uracil-N-Glycosylase (UNG) and 4 µl of DNA. The MgCl2 solution, reaction buffer and BSA were decontaminated by exposing the solutions in UV-pervious tubes (Qubit®, Life Technologies) to UV light for 10 minutes at short distance as described (29). A negative control was performed for each multiplex reaction that was processed in the same way as the samples throughout the whole experimental procedure. The cycling program consisted of 15 minutes at 37°C (carry-over contamination prevention and aDNA damage products through digestion by UNG of dUTP-labeled amplicons and of cytosine deamination products), 95°C for 10 minutes (inactivation of UNG and activation of the Fast Start DNA polymerase), followed by 40 cycles at 95°C for 15 seconds and 60°C for 1 minute and a final extension step at 72°C for 4 minutes. All extracts were tested for inhibition as previously described (35).

2.2.2. Coat color

We chose a set of eight SNPs in seven genes for detecting basic colors (bay, black, chestnut and grey), diluted phenotypes (silver and cream), spotted or painted phenotypes (overo, tobiano and sabino) and
leopard spotting, as described previously (36) (see paragraph SI 3.2.) We optimized the multiplex PCR as described before (24) and all the primers were selected and combined in a single multiplex reaction (see paragraph SI 3.4.2.). Each reaction was carried out in a 30 µl reaction volume with a final composition of 50 mM Tris/HCl, 10 mM KCl, 5 mM (NH₄)₂SO₄, 4 mM MgCl₂, pH 8.3 (@25°C), 1 mg/ml BSA (bovine serum albumin), 0.15 µM of each primer, 0.25 mM dA/G/CTPs and 0.50 mM dUTP, 2 U of Fast Start DNA polymerase, 0.5 U of uracil-N-Glycosylase (UNG) and 4 µl of DNA. A negative control was performed for each multiplex reaction that was processed in the same way as the samples throughout the whole experimental procedure. The cycling program consisted of 15 minutes at 37°C (carry-over contamination prevention through digestion by UNG of dUTP-labeled amplicons), 95°C for 10 minutes (inactivation of UNG and activation of the Fast Start DNA polymerase), followed by 50 cycles at 95°C for 20 seconds, 55°C for 30 seconds, 72°C for 30 seconds and a final extension step at 72°C for 4 minutes.

2.2.3. Y chromosome

We analyzed four anonymous Y-linked fragments, Y2B17, Y3B1, Y3B12, Y3B19 and two fragments of the amelogenin gene, AME2 and AME3, as previously described (37) (see paragraph SI 3.3). We optimized the multiplex PCR as described before (24) and all the primers were selected and combined in two multiplex reactions (see paragraph SI 3.4.3.). Each reaction was carried out in a 30 µl reaction volume with a final composition of 50 mM Tris/HCl, 10 mM KCl, 5 mM (NH₄)₂SO₄, 4 mM MgCl₂, pH 8.3 (@25°C), 1 mg/ml BSA (bovine serum albumin), 0.15 µM of each primer, 0.25 mM dA/G/CTPs and 0.50 mM dUTP, 2 U of Fast Start DNA polymerase, 0.5 U of uracil-N-Glycosylase (UNG) and 4 µl of DNA. A negative control was performed for each multiplex reaction that was processed in the same way as the samples throughout the whole experimental procedure. The cycling program consisted of 15 minutes at 37°C (carry-over contamination prevention through digestion by UNG of dUTP-labeled amplicons), 95°C for 10 minutes (inactivation of UNG and activation of the Fast Start DNA polymerase), followed by 50 cycles at 95°C for 20 seconds, 55°C for 30 seconds, 72°C for 30 seconds and a final extension step at 72°C for 4 minutes.

For each sample, we then performed the six multiplex PCRs with different primer combinations (mtDNA, coat color and Y chromosome) that were pooled at the end in a 96-well plate. Each extract was amplified in triplicate so that, at the end, there were three pools for each extract.

2.3. ION Torrent sequencing

The following steps were performed using an automate Tecan Freedom Evo 100 equipped with a 4 channel liquid handling arm suitable for disposable tips, a gripper to move objects, a double thermoblock and an automated solution for vacuum solid phase extraction. We first performed End Repair with the NEBNext end repair module (New England Biolabs) using 20 µL of multiplex PCR (containing no more than 5 picomoles of amplicon products) in a 50 µL reaction volume with only 0.1 µL of End Repair Enzymes. The reaction was incubated for 30 minutes at 25°C and then purified using the NucleoSpin 96 PCR clean-up kit as recommended by the manufacturer (Macherey-Nagel ref. 740658). The DNA was eluted in 50 µL and 20 µL were used for the ligation of the sample-specific Ion Torrent barcoded adaptors (1 µL of annealed A+P1, 20 µM) using the NEB Next ligation module (New England Biolabs) in a 30 µL reaction volume with 1 µL of Quick Ligase. The reaction was incubated for 30 minutes at 16°C. After the ligation, 60 µL of binding buffer (i.e., NT buffer, Macherey-Nagel) was added and all samples were pooled in a tube before purification on a silica column. The pooled barcoded PCR products were size
selected using either an E-Gel SizeSelect (Life technology) or the Caliper Labchip XT (Perkin Elmer) to purify the PCR products ligated to the adapters. The selected product was subjected to nick repair and amplification with the NEB OneTaq Hot Start in a 40 µL reaction volume containing Ion Torrent primers A and P1 (0.5 µM). The reaction was incubated for 20 min at 68°C (nick repair step). The DNA library was then amplified using the following program: initial denaturation at 94°C for 5 min, (94°C for 15 sec, 60°C for 15 sec, 68°C for 40 sec) for 6 cycles, final elongation at 68°C for 5 min. Products were finally purified with a Qiaquick PCR purification kit (Qiagen). The size distribution and concentration of the library was assessed on the Agilent 2100 Bioanalyzer. Emulsion PCR and Ion Sphere Particle enrichment were conducted with the Ion OneTouch System (Life Technologies) using the Ion OneTouch 200 Template kit v2 DL according to the manufacturer’s protocol. The DNA library was sequenced on the Ion Torrent Personal Genome Machine (PM) Sequencer using the Ion PGM 200 Sequencing Kit and Ion 314 semiconductor sequencing chips (Life Technologies).
3. Information to PCR primers used

3.1. PCR primers for mitochondrial DNA

| Prims   | Primer seq (5' - 3') | Primer coordinates | Amplicon length (bp) | Sequence length (bp) | Gene region |
|---------|----------------------|--------------------|----------------------|----------------------|-------------|
| SNP2F   | GGTACAGCTTTTTAGATACAGGTA | 1554-1578          | 81                   | 34                   | 16S         |
| SNP2R   | GGCTGCTTTTTAAGCCAACATAG | 1613-1634          |                      |                      |             |
| SNP4F   | TTCAACTCCTCTCCCCCTAAACAA | 2749-2769          | 78                   | 37                   | ND1         |
| SNP4R   | AATGCTACGGGCGAGCAAGAT | 2807-2826          |                      |                      |             |
| SNP8F   | AATACGCAAACATCATCATGATAAA  | 3450-3474          | 91                   | 44                   | ND1         |
| SNP8R   | ATGAGTAGAGATCTGCGAGGT | 3519-3540          |                      |                      |             |
| SNP10F  | GAATCGGATTTTAAATATCTACT | 4025-4049          | 73                   | 25                   | ND2         |
| SNP10R  | TCTATGGTCCGGGATTGCTACTT | 4075-4097          |                      |                      |             |
| SNP11F  | TCGGCATATAGGATGAATAAC | 4487-4508          | 120                  | 73                   | ND2         |
| SNP11R  | AAGAGCTGCTGGATAAATGATGAA | 4582-4606          |                      |                      |             |
| SNP13F  | AATAAAAATGACAATTCGAAACAAA | 4872-4896         | 73                   | 25                   | ND2         |
| SNP13R  | GAGGAGTAGGAGGATATAACAA | 4922-4944          |                      |                      |             |
| SNP15F  | GAATTAGGCCAACCCTGGGAC | 5479-5498          | 75                   | 35                   | COX1        |
| SNP15R  | TACGAATGCAAGGGGCGGTTA | 5534-5553          |                      |                      |             |
| SNP17F  | GTATGGGCTCACCACATGTTT | 6220-6240          | 119                  | 76                   | COX1        |
| SNP17R  | TGTTGCTGCTGCTGCTGCTTCT | 6317-6338          |                      |                      |             |
| SNP21F  | TAGGACTCCTCTATTTGAATCTCG | 8008-8030          | 95                   | 49                   | ATP6        |
| SNP21R  | TTGAATAGGAGGATATAGGTG | 8080-8102          |                      |                      |             |
| SNP25F  | CTCACAGCCCTTTTATCTGTTA | 8522-8544          | 76                   | 30                   | ATP6        |
| SNP25R  | GGCTGTTGATTAGCTACTGCGA | 8575-8597          |                      |                      |             |
| SNP26F  | TCCCCCTATCAACACCTCA | 9030-9049          | 83                   | 40                   | COX3        |
| SNP26R  | ACGTTCCTTCTTATAGAGCTAT | 9090-9112          |                      |                      |             |
| SNP29F  | TCCTAACACCCACCTTCACACT | 10057-10078    | 91                   | 47                   | ND4L        |
| SNP29R  | AGTAGGGATAAATCCTAGAGCTG | 10126-10147        |                      |                      |             |
| SNP31F  | ATCCCCCTACTCAGCCCCACTT | 10380-10399        | 71                   | 31                   | ND4L        |
| SNP31R  | GTTTGGCGGTATCTGATG | 10431-10450        |                      |                      |             |
| SNP32F  | CTGCCCACTAATACCTAGATGGAA | 10421-10443        | 79                   | 32                   | ND4L        |
| SNP32R  | GCAGGTGTGATGTAGAGGTGTGT | 10476-10499        |                      |                      |             |
| SNP36F  | GAAAGTATGCAAGCTACTGCTACT | 11654-11676       | 71                   | 25                   | Ser (AGY)   |
| SNP36R  | TCCTTTAAAAAGTTTGAGAGGC | 11702-11724        |                      |                      |             |
| SNP37F  | TCCTCCTCATACTAGTTTCCAC | 11795-11816        | 116                  | 70                   | ND5         |
| SNP37R  | GCATATGAGATGTTTTTTTACA | 11887-11910        |                      |                      |             |
| SNP41sh1F | ATCAGCAATTCCTACATCGG | 14637-14657        | 77                   | 35                   | CytB        |
| SNP41sh1R | TAAGGTTGCTTTTGTCTACTG | 14693-14713        |                      |                      |             |
| SNP41sh2F | GGTGGATTCAGTACGAGAAAAA | 14683-14703 | 79 | 37 | CytB |
| SNP41sh2R | GGGCTGTGATGATGAAGGTA | 14741-14761 | 96 | 52 | D-loop |
| HVR1F | ATTCTTTCCCTAAACGACAAC | 15469-15490 | 121 | 73 | D-loop |
| HVR1R | ATATTGCATGTCAGGGTAT | 15543-15564 | 121 | 73 | D-loop |
| HVR5F | CCGGGGAAAATACGCAAC | 15749-15775 | 121 | 73 | D-loop |
| HVR5R | GAATGGCCCTGAAGGAAAC | 15849-15869 | 121 | 73 | D-loop |

3.2. Primers for alleles involved in coat color

| Colour | Primer name | Gene region | Primer seq (5′-3’) | Amplicon length (bp) | Sequence length (bp) |
|--------|-------------|-------------|--------------------|----------------------|----------------------|
| Chestnut | MC1R-F | MC1R-E1 | GCACTCACCAGTTACTACTTCAGTCAC | 71 | 29 |
| | MC1R-R | | GCACTCACCAGTTACTACTTCAGTCAC | 71 | 29 |
| | ASIP-F | | AAGAAATCCAAAAAGATCAGCATTGAGTACCA | 77 | 38 |
| | ASIP-R | | AAGAAATCCAAAAAGATCAGCATTGAGTACCA | 77 | 38 |
| Cream | MATP-F | MATP (SLC45A2) | GCCATAACATCATCAGTAGAGTGCCCTACATGAAGTACGG | 65 | 24 |
| | MATP-R | | GCCATAACATCATCAGTAGAGTGCCCTACATGAAGTACGG | 65 | 24 |
| Tobyano | KIT13F | KIT (intron 13) | CGTCAATGACTCTTCATGAGGATAACCA | 63 | 18 |
| | KIT13R | | CGTCAATGACTCTTCATGAGGATAACCA | 63 | 18 |
| | KIT16F | KIT (intron 16) | TTTAAATGCTTTCCTTCCTCC | 59 | 15 |
| | KIT16R | | TTTAAATGCTTTCCTTCCTCC | 59 | 15 |
| Silver | SILV11-F | SILV | TCTCTCTCTCCTCCTCAATCTGAAGA | 52 | 9 |
| | SILV11-R | | TCTCTCTCTCCTCCTCAATCTGAAGA | 52 | 9 |
| Overo | ENDRB-F | | CATGACTGCTGCATGCTGTT | 63 | 17 |
| | ENDRB-R | | CATGACTGCTGCATGCTGTT | 63 | 17 |
| Leopard | LP-F | | AGGTGAATGCTGACTGGA | 62 | 22 |
| | LP-R | | AGGTGAATGCTGACTGGA | 62 | 22 |

3.3. Sequences of Y-chromosome primers

| Primer name | Gene region | Primer seq (5′-3’) | Amplicon length (bp) | Sequence Length (bp) |
|-------------|-------------|--------------------|----------------------|----------------------|
| Y2B17-1F | Y2B17 | ATGTCAGGGATGCGACCTGTCGTGCCAACAGACAGGTGT | 74 | 34 |
| Y2B17-1R | | | | |
| | Y2B17-4F | Y2B17 | AGTTGCCGCCTGAAGAGACATTGACRGGAAAAAGGGACAGTG | 76 | 36 |
| | Y2B17-4R | | | |
| | Y2B17-5F | Y2B17 | GCCACCATGTGCCCTTTTTCAGGACACACACACCCCACCTCA | 67 | 28 |
| | Y2B17-5R | | | |
| | Y2B17-6F | Y2B17 | GGCAGACAGCAGCAGCAGCAGCTGAGAGTACGGATGAGA | 64 | 27 |
| | Y2B17-6R | | | |
| | Y3B1-9F | Y3B1 | CTCTCTCTCTAATGTCGCGACAGGTCACBKGCCAGCAGTCATTC | 58 | 16 |
| | Y3B1-9R | | | |
| | Y3B1-10F | Y3B1 | GCTCTCAGCGTCAGGCAGACTGACACTTCACGTGTGTAAGAGA | 56 | 16 |
| | Y3B1-10R | | | |
| | Y3B1-12F | Y3B1 | AATGTCAGGGATGCGACCTGTCGTGCCAACAGACAGGTGT | 73 | 27 |
| | Y3B1-12R | | | |
| Multiplex: | 1 | 2 | 3 |
|-----------|---|---|---|
| SNP2      | SNP29 | SNP8 |
| SNP31     | SNP15 | SNP21 |
| SNP11     | SNP4  | SNP26 |
| SNP17     | SNP32 | SNP36 |
| SNP41sh2  | SNP25 | SNP10 |
| HVR1      | SNP37 | SNP13 |
| HVR5      | SNP41sh1 |

PCR program: 1 cycle 37°C 15 min, 95°C 10 min; 40 cycles 95°C 15 sec, 60°C 60 sec; 1 cycle 72°C, 4 min

### 3.4. Multiplex PCR conditions

#### 3.4.1. Mitochondrial DNA

| Multiplex: | 1 | 2 | 3 |
|-----------|---|---|---|
| ASIP      | KIT13 | KIT16 |
| KIT13     | MATP  | MC1R |
| MATP      | SILV11 | LP |
| MC1R      | EDNRB |

PCR program: 1 cycle 37°C, 15 min, 95°C, 10 min; 50 cycles 95°C 20 sec, 55°C 30 sec, 72°C 30 sec; 1 cycle 72°C 4 min
3.4.3. **Multiplex Y chromosome**

| Multiplex: | 1          | 2          |
|------------|------------|------------|
| P-Y2B17-1  | P-Y3B1-10  |
| P-Y2B17-5  | P-Y2B17-4  |
| P-Y2B17-6  | P-YAme2-20 |
| P-Y3B1-9   | P-YAme2-22 |
| P-Y3B1-12  | P-YAme2-24 |
| P-Y3B12-16 | P-YAme3-26 |
| P-Y3B19-17 | P-YAme3-27 |
|            | P-YAme3-28 |

PCR program: 1 cycle 37°C, 15 min, 95°C, 10 min; 50 cycles 95°C 20 sec, 55°C, 30 sec, 72°C, 30 sec; 1 cycle 72°C 4 min

**Supplementary Tables:**

Table S1: samples analyzed with dates and results

Table S2: List of SNPs used for determination of the mitochondrial DNA haplogroups

Table S3: List of diagnostic SNPs for the determination of horse mitochondrial DNA haplogroups in both the complete mitogenomes and in the reduced assay used in the present study

Table S4: List of diagnostic SNPs assayed to determine horse mitochondrial DNA haplogroups

Table S5: List of SNPs obtained from the Y-chromosome assay

Table S6: List of genotypes obtained from the coat color assay

**Supplementary Figures:**

Fig. S1: Map showing the archaeological sites from where the analyzed samples originated

Fig. S2: coat colors found in the present study
Fig. S1: Map showing the locations of the archaeological sites from which the analyzed specimens originate. Aşıklı Höyük (1), Acemhöyük (2), Çadr Hőyük (3), Boğazköy-Hattusa (4), Köşk Höyük (5), Troy (6), Lidar Höyük (7), Didi Gora (8), Tqisbolu Gora (9), Lusakert (10), Mets Sepasar (11), Tsaghkahovit (12), Gegharot (13)
Fig. S2: Horses showing the coat colors evidenced in the present study. Copyright: Institut du cheval et de l’équitation ifce, France.
