UNIVERSITY OF EXETER

ORE Open Research Exeter

TITLE

Ancient genomes revisit the ancestry of domestic and Przewalski's horses

AUTHORS

Gaunitz, C; Fages, A; Hanghoj, K; et al.

JOURNAL

Science

DEPOSITED IN ORE

26 February 2018

This version available at

Supplementary Materials for

Ancient genomes revisit the ancestry of domestic and Przewalski's horses

Charleen Gaunitz,[1]* Antoine Fages,[1,2]* Kristian Hanghøj,[1,2] Anders Albrechtsen,[3] Naveed Khan,[1,4] Mikkel Schubert,[1] Andaine Seguin-Orlando,[1,2,5] Ivy J. Owens,[6,7] Sabine Felkel,[8] Olivier Bignon-Lau,[9] Peter de Barros Damgaard,[1] Alissa Mittnik,[10] Azadeh F. Mohaseb,[11,12] Hossein Davoudi,[12,13,14] Saleh Alquraishi,[15] Ahmed H. Alfarhan,[15] Khaled A. S. Al-Rasheid,[15] Eric Crubézy,[2] Norbert Benecke,[16] Sandra Olsen,[17] Dorcas Brown,[18] David Anthony,[18] Ken Massy,[19] Vladimir Pitulko,[20] Aleksei Kasparov,[20] Gottfried Brem,[8] Michael Hofreiter,[21] Gulmira Mukhtarova,[22] Nurbol Baimukhanov,[23] Lembi Lõugas,[24] Vedat Onar,[25] Philipp W. Stockhammer,[10,19] Johannes Krause,[10] Bazartseren Boldgiv,[26] Sainbileg Undrakhbold,[26] Diimaajav Erdenebaatar,[27] Sébastien Lepetz,[11] Marjan Mashkour,[11,12,13] Arne Ludwig,[28] Barbara Wallner,[8] Victor Merz,[29] Ilja Merz,[29] Viktor Zaibert,[30] Eske Willerslev,[1] Pablo Librado,[1] Alan K. Outram,[6†,+] Ludovic Orlando,[1,2†,+]

Correspondence to:
Alan K Outram (A.K.Outram@exeter.ac.uk), Ludovic Orlando (orlando.ludovic@gmail.com)

This PDF file includes:

    Materials and Methods
    Supplementary Text
    Figs. S1 to S34
    Tables S1 to S7
    References 1 to 174

Other Supplementary Material for this manuscript includes the following:

    Tables S8 to S18 hosted online as Excel files

Material and Methods

<u>Section 1. Sample information</u>

Horse archaeological remains were processed in the Ancient DNA facilities of the Centre for GeoGenetics, Copenhagen, Denmark, through the procedure described in section 5 and sequenced at low sequencing depth (generally ~5-10 million reads) at the Danish High-Throughput DNA Sequencing Centre. As part of this procedure, a total of 42 ancient horse samples spanning a time range of ~5,500 years ago (ya) to ~1,100 ya, recovered from 17 palaeontological and archaeological sites across whole Eurasia, showed moderate to high endogenous DNA content (0.17-0.78), following limited shotgun sequencing efforts. These were selected for further analyses, including deeper sequencing of the screening DNA libraries as well as additional DNA libraries that were constructed to limit the redundancy of the sequencing data. The following section describes the sites and the archaeological contexts associated with all remains analyzed in this study (also see Table S1).

Unless indicated otherwise, AMS-radiocarbon dating of the samples was performed at the Keck Carbon Cycle Accelerator Mass Spectrometry Laboratory, UC Irvine. Bone or tooth pieces between 729 mg and 2,832 mg were sampled in the bone laboratory in the Ancient DNA facilities of the Centre for GeoGenetics and sent for subsequent dating. Sample preparation backgrounds were estimated, and subtracted, based on measurements of ^{14}C-free calcite. Calibration was carried out using OxCalOnline *(27)* and the IntCal13 calibration curve. The amount of ultrafiltered collagen obtained for sample Ridala_Rid1_2717 was incompatible with radiocarbon dating. Calibrated dates are provided in Table S2.

Botai (Kazakhstan)

The Eneolithic site of Botai in Northern Kazakhstan, dated to ~3,500BC, represents the earliest evidence for horse domestication *(8)*. Botai people had a semi-sedentary lifestyle; they lived in pit-house settlements and produced a wide variety of artifacts, including copper tools, bell-shaped pottery and thong-smoothers. Over 300,000 horse bone fragments have been excavated in the four Botai settlements of Botai, Roshchinskoe, Krasnyi Yar I and Vasilkovka IV, suggesting the central role of horses in their economy *(28)*. Evidence for horse domestication include the discovery of a large number of handcrafted thong-smoothers, the presence of biting damage on horse premolars – suggesting harnessing – as well as the detection of mare milk processing residues in pottery *(8)*. A total of 20 samples sequenced in this study were excavated from the Botai site in 2015 and 2016. They consist of petrosal bones mainly, but also teeth and metapodials.

Borly 4 (Kazakhstan)

The settlement of Borly 4 is located in the northeast of Kazakhstan in the Pavlodar region (51°49'28.9"N 77°56'42.4"E). The site occupies the headland of the second lakeside terrace of the western shore of the lake. An area of 175 m² has been studied at the site, and revealed archaeological materials from the Middle Neolithic to the Early Iron Age. The main cultural layer is Eneolithic and contains dwellings and other structures including a large number of stone, ceramics and copper artifacts, as well as bones of horses and cattle *(29, 30)*. Samples Borly4_PAVH6_5012, Borly4_PAVH8_4978 and Borly4_PAVH9_4977 were obtained from the

Eneolithic dwelling No. 1 in 2012, and Borly4_PAVH4_4974 and Borly4_PAVH11_5015 from dwelling No. 2 in 2013.

Grigorevka (Kazakhstan)

Sample Grigorevka4_PAVH2_1192 was obtained in 2012. It was selected from a horse cranium, without the mandible, deposited on a possible alter on the western side of the burial of a teenager at Grigorievka 4. According to the radiocarbon dates (see Table S2), this burial refers to the Kimak period. Cultural affiliation is, however, not possible, since the period of the early and middle Middle Ages in the region remains poorly understood.

Postillionstraße Haunstetten (Germany)

The cemetery and archaeological site Positillonstraße is situated in Haunstetten close to the city boundary of Königsbrunnen, Augsburg (Bavaria, Germany). The excavation was carried out in 1992 and revealed 40 inhumation graves, all dating back to the Early Bronze Age (EBA) *(31, 32)*. Three of the burials were surrounded by circular enclosures and were probably covered by a burial mound. All skeletons were found in long-oval pit holes in a crouched position and were either pointing North (males) or South (females) with both genders facing east. The horse burial ("grave 34", feat. no. 17) was found alongside a grave within the row of the three burials surrounded by circular enclosures next to a burial pit containing a female skeleton (2029 BC to 1911 BC). The horse skeleton was lying on its left side with the skull pointing North, indicating that it was a stallion (the sex was confirmed by DNA analyses). Radiocarbon dating, conducted by the Curt-Engelhorn-Zentrum Archäometrie gGmbH, Mannheim in 2015, showed that the horse skeleton was intrusive in the EBA layer, and dated back to Roman times (MAMS 23733: 4 AD to 123 AD).

Bateni (Khakassiya Republic)

The sample Bateni_Rus14_3318 belongs to the Late Bronze Age Karasuk culture (1500-800 BC), which followed the Andronovo culture in the South of Siberia. It covered an area from the Aral Sea to the Yenisei river on the East and to the Altai mountains and Tien Shan in the South. Karasuk communities are known to be farmers who practiced a mixture of agricultural and stockbreeding of cattle, sheep and horse. They are especially known for their metallurgy, in particular for their daggers and knives *(33, 34)*. The bone sample was excavated by M.P. Gryaznov between 1963 and 1964 close to the Bateni settlement, Republic of Khakassiya. The excavation site is not accessible anymore due to the flooding of the Kranoyarsk Reservoir.

Uushgiin Uvur (Mongolia)

The two horses sequenced in this study, UushgiinUvur_Mon84_3123 and UushgiinUvur_Mon86_3039, were sampled from two skulls excavated from the late Bronze Age complex of deer stones and ritual mounds, facing the Ulaan Uushig Mountain, in Northern Mongolia. The Uushgiin Uvur Deer Stone Complex was dated at 1312-810 BCE and was named after the carving of deers and other animals - including horses - on the stone plinths, which represents a unique mortuary landscape tradition. The deer stones often stand on top of buried remains from horses sacrificed without any artifacts as offerings, including cervical vertebras, hooves and heads facing East *(35)*. With its 30 deer stones, ranging from 133 to 375 cm high, arranged in a row from North to South and surrounded by many ritual structures with solid or ring-

shaped rock coverings, the Uushgiin Uvur Deer Stone Complex displays one of the largest number of deer stones in Eurasia.

Syrgal (Mongolia)

Syrgal is an archaeological site located in the Mongolian Altai (around 2,100 m above sea level) associated with the Pazyryk culture, a culture of nomadic people traveling on horseback and dating to the Iron Age (III-IV century BC). The sample Syrgal_Syr1t1c3_2317 from Syrgal was unearthed from the skeleton of a horse buried together with one human and four other horses.

Gol Mod II (Mongolia)

Gol Mod II is a large cemetery, located north of the Khangai mountain range in Central-North Mongolia. The site, initially excavated in 2001, 2002, 2004 and 2005 and then fully excavated by 2011, is dated to 3rd century BCE - 1st century AD and is associated with the Xiongnu culture *(36)*. The Xiongnu were nomadic peoples of the steppes of North-East Central Asia, who expanded across the steppes of Central Asia on horseback and whose economy relied on herding. The sample GolModII_Mon28_1988 was excavated from the enormous main tomb of the cemetery, built for an aristocratic figure of the Xiongnu culture, and surrounded by a 48 m long, 3.5 m tall rock wall on the east, a 48 m long, 2.2 m tall rock wall on the west, and a 46 m long, 4m tall rock wall on the north. This is the largest burial structure ever found in Central Eastern Asia, and the deepest grave ever dug (23 m). As such, more than 20 horse remains - including the sample GolModII_Mon28_1988 - and five to six goat or sheep remains were discovered ten meters below ground level. Other discovered items include horse-drawn cart remains from iron and bronze carts, tools, gold and silver jewels and decorations, a jade mirror, vases, bronze and glass bowls.

Dunaújváros-Kosziderpadlás (Hungary)

The sample Dunaujvaros_Duk2_4077 was excavated from the Middle Bronze Age settlement of Dunaújváros-Kosziderpadlás (or Kosziderpadlás) in Hungary, which was a major site of the Early-Bronze Age Nagyrév, Vatya and Kisapostag cultures *(37, 38)*. The site is located in the city of Dunaújváros, on the right bank of the Danube, and represents an oval area of 290 by 185 meters surrounded by ramparts. Excavations revealed the presence of hearths and pits within houses, as well as hidden or buried bronze hoards. The site, where over 550 cattle, 275 sheep/goat, 250 dogs, 200 pigs, and 150 horses were excavated (NISP), was probably an agricultural and trade centre where livestock and crops were exchanged for bronze articles *(39)*. Radiocarbon dating of the Dunaujvaros_Duk2 sample sequenced in this study indicated a date of ~2,139 –1,981 cal. BC (IntCal13, Table S2) *(27)*.

Garbovat (Romania)

The archaeological site of Gîrbovăţ (Romania) is located about five kilometers South-East of the town Ghidigeni, and 20 kilometers north of Tecuci. The site belongs to the Bronze Age Noua Culture, which mainly occurred in the northern parts of Romania and Moldova and in southwestern Ukraine *(40)*. Gîrbovăţ is especially rich on faunal remains. In total 6,773 animal bone fragments (Number of Identified Specimens, NISP), including 6,599 of domestic origins (97.3%), such as the sample Garbovat_Gar3_3574 sequenced in this study, were retrieved during the excavations between 1960 and 1962 led by A. Florescu. Horse bones represent only a minority of the total amount of cattle, goat and sheep remains. Archaeologists hypothesized that the horses

found in this context may have participated to the gene pool that underpinned the improvement process by which the Geto-Dacians have constituted the so-called "elite" horse.

Fresnes les Montauban and Actiparc (France)

The first French sample FMontauban_GVA126_1872 was part of a horse skeleton found in a well located at Fresnes les Montauban, Northern France, and dated to the Roman period, around II-III AD. The well is thought to carry some ritual dimension. The second French sample, Actiparc_GVA310_2319, was excavated from Actiparc, Saint Laurent-Blangy, Northern France. This site is associated with the final Iron Age La Tène culture and the Roman period, and dates back to 100 BCE-100 AD. Equine bones and teeth found on this site were found in pits and ditches, and while some of the samples were identified as residual food rest, the majority of the horse remains are considered to be part of the Gallic cavalry.

Tepe Hasanlu (Iran)

The horse sample TepeHasanlu_2405_2886 originates from the site of Tepe Hasanlu, excavated between 1956 and 1977 during ten archaeological campaigns under the general direction of Robert Henry Dyson, Jr. *(41)*. Hasanlu is located in the Solduz Valley in the Western Azerbaijan province of Northwestern Iran. The site consists of two distinct topographic zones: a High Mound which is called the 'Citadel Mound', surrounded by a Low Mound, also known as the 'Outer Town' *(42)*. The site was intermittently occupied during ten different cultural periods from the Late Neolithic (period X) to the Ilkhanid dynasty (period I) *(43)*. The Citadel of Hasanlu was destructed and fired during a battle around 800 BC. Extensive excavations resulted in the discovering of burned buildings, thousands of artifacts in closed contexts, and a large quantity of faunal remains. TepeHasanlu_2405_2886 was discovered within the deposit of Burned Building IV/V in the Citadel, which relates to the Iron Age II (1050-800 BC) (Campaign 1974, Op. X33, Stratum 5, Locus 4B, Lot 14). Horse harnesses, gear and skeletons were also found in the other contexts of Buildings IVE and V, which are thought to have been used as stables and storerooms at the time of destruction *(42)*. The TepeHasanlu_2405_2886 sample originates from an anterior first phalanx which was confirmed to belong to a horse by standard metrical analyses.

Zhanaturmus (Kazakhstan)

The sample Zhanaturmus_Issyk1_1143 originates from an almost complete horse skeleton which was unearthed during the excavation of a modern Muslim cemetery in June 2011 in the village of Zhanaturmus (near the city of Yesik), Enbekshikazakh district, Almaty region, Kazakhstan. The horse remains were found in a barrow oval grave pit at a depth of 1.38 m. Right next to it, close to its western wall, at a depth of 2.10 m a looted tomb of a warrior has been found. Burial type and objects excavated, including belt bronze plaques and fragments of a short metal sword are characteristic of the early Turkic culture of Semirechie (8th-9th century AD). The cultural affiliation has been confirmed by radiocarbon dates of the horse remains (777-971 AD calibrated years). The Turkic burial rituals found at this site show similarities to other Turkic sites excavated in Central, Southern, Eastern Kazakhstan Altai and Tuva.

Tachti Perda (Georgia)

The sample TachtiPerda_TP4_3604 was excavated from the multi-layered prehistoric settlement site of Tachti Perda, located close to Dedopliscqaro, a town in Kakheti, eastern Georgia.

The site was constantly populated throughout different time periods, starting from the late middle Bronze Age (16th century BCE) to the late Iron Age (7th century BCE), and is especially rich in ceramics, stone and metal tools. Due to the strategic and topographical location of Tachti Perda, archaeologists consider that the site was an economic and cultural cornerstone for the whole region *(44)*.

Ridala (Estonia)

The sample Ridala_Rid1_2717 was recovered from a fortified settlement on a moraine ridge close to the coastal zone of Saaremaa Island (west Estonia), which was at the time of the settlement a coastal island. Excavations were carried out in 1961 by Aita Custin and in 1963 by Artur Vassar. The archaeological site covers an area of around 4,500 square meters and was dated to the 8th-7th centuries BCE. Until now, approximately one tenth (435m²) of the area has been excavated. A total of 2,020 bone fragments have been recovered, 75% of which belonging to domestic animals (sheep/goat, pig, cattle, and horse) and 25% to wild animals (seals). Sheep and goat bone fragments are the most frequent while horse bones represent the least frequent of all faunal remains. The horse remains recovered from this site belong to eight different individuals. Two of the specimens were foals, two were slaughtered before the age of three and the other four individuals were between two and four and a half years old. The presence of remains from exclusively juvenile individuals suggests their use for food consumption *(45-47)*.

Yenikapi (Turkey)

The Yenikapi archaeological site is located in Istanbul and covers an area of ~50,000 square meters *(48)*. The excavations led by the Istanbul Archaeological Museum in 2004 revealed that it was a large Byzantine harbour, the harbour of Theodosius. A total of 20,881 bone specimens were identified, of which 32.6% were horses, including the skull sample Yenikapi_Tur172_1695 sequenced in this study *(49)*.

Shahr-I Qumis (Iran)

Shahr-i-Qumis is a very large archaeological site in north eastern Iran located today in a very arid environment. It consists of a series of isolated mounds with a spread of surface pottery over an area of about 28 kilometers. It was excavated during four seasons between 1967 and 1978 co-directed by J. Hansman and D. Stronach *(50)*. Remains found at this site date back to the Parthian and Sasanian periods and even Seljukid. Surprisingly, some structures were converted to secondary, funerary purposes. One of the highlights of the site was the discovery of a very large quantity of equine skeletons *(51)*. The sample ShahrIQumis_AM181_1694 originated from a petrous bone of one these horse skeletons. Given its strategic location on the Great Khorasan Road and the Silk Road and descriptions of Greek historians, Shahr-i-Qumis is identified as the ancient city of Hekatompylos (hundred gates) and capital of the Parthian Empire. However, new radiocarbon dates on the faunal remains suggest an occupation from the 8th century BCE to 8th century AD.

Comparative dataset

We compared the 42 genomes sequenced in this study to a panel consisting of 18 ancient and 28 modern genomes previously published, as well as the genomes of two modern donkeys: one domestic donkey and one Somali wild ass, used as outgroups in our analyses (Tables S8 and

S9), *(15, 20, 21, 52-64))*. The former includes three pre-domestic horses from a now-extinct population encompassing from the Late Pleistocene to the Early Holocene (~42,000, ~16,000 and ~5,000 years old, respectively); one mare associated with the Sintashta culture, dated to ~3600 ya; 11 Scythian stallions from Berel' burial site in Kazakhstan; two individuals from Arzhan I in Russia (2,825-2,760 ya); one individual from 19th century AD Yakutia in Northern Siberia, and; one Przewalski's historical sample from the 19th century AD. One additional stallion from Berel' of relatively limited genome coverage (Berel_BER11_L_2300) was included in analyses of mitochondrial DNA, Y chromosome and Mendelian loci, but excluded from all other analyses given its high relatedness to another Berel' individual (Berel_BER10_K_2300). The comparative panel of modern horse genomes consists of six Przewalski's horses and 22 domestic horses spanning 18 domestic breeds, including one individual from the Arabian, the Connemara Pony, the Duelmener, the Franche-Montagnes, the Friesian, the Hannoverian, the Saxon-Thuringian Heavy Warmblood, the Icelandic, the Jeju pony, the Marwari, the Morgan, the American Quarter, the Shetland pony, the Sorraia, and the Standardbred breed, as well as two Mongolian individuals, two Thoroughbreds and three Yakutian horses.

Section 2. The horse domestication debate in relation to the Botai culture

Horse domestication is widely viewed as being revolutionary to human societies, transforming transportation systems, trade, migration and methods of warfare *(3, 5, 8)*. Additionally, two recent ancient genomic studies *(26, 65)* indicate that the development of pastoralism, particularly in relation to the Yamnaya Culture of the Pontic Steppe, was responsible for significant human migrations within Eurasia around 5,000-6,000 years ago that could be involved in the spread Indo-European languages and culture. The formation and expansion of such societies could be heavily linked to horse riding, pastoral economies, the use of wheeled vehicles and new forms of bronze metallurgy *(2)*. Understanding the earliest development of horse husbandry should be considered to be one of the more significant questions within prehistoric archaeology, since the enhanced mobility it allowed took humans into the incipient phases of globalisation. The steppes of Eastern Europe and Central Asia represent a significant crossroads between major centres of civilization in the East and West and these regions became key in the initial phases of the '*Trans-Eurasian Exchange*' *(66, 67)*.

One of the key candidates for early horse domestication is the Eneolithic Botai culture of Northern Kazakhstan, despite over 30 years of scientific controversy *(5, 6, 68, 69)*. The following text summarizes the main arguments of this debate in relation to the findings presented in this study.

Northern Kazakhstan has two major ecological zones with a 'forest steppe', made up of a patchwork of grassland with stands of birch and pine trees, in the North, and relatively treeless, semi-arid steppe in the Central region. The environment was similar in prehistory, with some temporal variation regarding tree coverage. Pine generally became more extensive from the 4th Mill. BCE up until the Iron Age *(70)*. Cereal agriculture did not become a common practice in this region until the Soviet period. The Neolithic (so named because it possesses ceramics rather than because the people were farmers) of Northern and Central Kazakhstan appears to have had an

economy based upon hunting, gathering and fishing, with a stone tool industry based principally on blade technology. Settlements tend to be rather ephemeral, often comprising little more than scatters, with no clear evidence for farming or husbandry.

Approximately ~5,500 years ago, with the onset of the Eneolithic, the Botai Culture displayed major changes in economy, settlement type and material culture (69, 71). The use of ceramics increased and stone tool traditions shifted towards bifacial flaking and ground stone tools. The Botai Culture has large settlements that can consist of over a hundred semi-subterranean pit houses. Most significantly of all, there is a sudden and intense focus on exploiting horses, and their bones make up the vast majority of faunal assemblages. At the eponymous/type site of Botai itself, horses make up 99% of the animal bones (3, 28). The Central Asian steppes would have been the home to a significant population of wild horses that were hunted for millennia, alongside other wild species. In the Botai Culture, however, there is a sudden focus on horses alongside the introduction of substantial settlements and notable changes to material culture. Since its discovery in the early 1980s, there has been much debate in relation to whether the horses at Botai were hunted or herded. It has been argued that there was no clear body size change in the animals that would be expected if they had undertaken domestication based upon some other domestication processes (9), whilst alternative arguments have suggested that there was domestication that did bring about a reduction in lower leg robusticity (8). However, such arguments, on either side of the debate, related to bone metrics, are rendered moot by the genetic identification of Botai horses as the ancestors of Przewalski's horses. Whilst it is possible to discuss metrical changes within this particular horse lineage by comparing ancient Botai horses with modern Przewalski's horses, the lack of clearly identifiable specimens of definitively wild early Przewalski's horses means that existing metrical and shape comparisons to other horse lineages do not necessary speak to the question of domestication. Other arguments against horse domestication include the observation that the herd structure at Botai is not optimized for meat production (72). However, herd structures would not be optimized for meat exploitation if the horses concerned were also being utilized for secondary products such as milk, riding or traction (6, 68).

The circumstantial arguments for horse husbandry at Botai are strong (28, 68, 69). A significant change in species focus is a frequent indicator of domestication (68), and, in this case, that re-focussing is accompanied by very significant increase in settlement size. The more settled nature of Botai villages perhaps also implies greater control over subsistence resources (28) and the presence of significant quantities of dung within settlements (73) alongside the existence of possible coral areas, provides evidence in favor of animal husbandry within the Botai culture. The confirmation, in this paper, of the presence of Eneolithic enclosures at the Botai site, is supportive of animal husbandry and inconsistent with the requirements of hunter-gatherers. The absence of a 'schlepp effect' (28, 68) at Botai, caused by selective transportation of hunted carcasses back to a settlement from a kill site (74), along with some evidence for pole-axing (28), suggest that animals were slaughtered on or near the site, again suggesting husbandry and control. Botai tool kits also appear to be more associated with hide working than hunting (28). Evidence for the harnessing of horses was presented in the form of the existence of bit wear traces on the lower second premolars (10) in a form now referred to as type 1 bit wear, which is identified by the presence of a bevelled facet on the tooth (6). This earlier work on bit wear has been augmented by the recent application of different methods to positively identify the presence of type 2 (parallel band of wear down the

front of the 2nd mandibular premolar) and type 3 (pathology of diastema) bit wear on Botai horses *(6, 8)*. The existence of three independent lines of bit-wear evidence provides a strong indication of control and possible riding of horses. Furthermore, Botai ceramics have been found to contain the absorbed lipid residues of two types of equine fats that have been identified on the basis of carbon and deuterium stable isotope ratios as being horse apidose fat and mare's milk fat *(8)*. The milking of mares is an unambiguous indicator of taming and husbandry.

The discovery, reported in this paper, that Botai horses are ancestral to Przewalki's horses does not affect the archaeological, pathological or chemical residues lines of evidence discussed above. It does, however, mean that arguments based upon bone metrical data currently lack appropriate comparative frameworks for identifying changes related to horse domesticity in prehistory. This equally affects arguments both for and against domestic status. On the other hand, the identification of a *TRPM1* haplotype associated with leopard spotting (when carried at the heterozygous state) and stationary night blindness (when carried at the homozygous state), is suggestive of the relaxation of selective pressures against an allele that would likely have existed in the wild. Clearly, this negative selective pressure could be removed through human husbandry, or, indeed, there could be positive human selection for leopard spotting. The means for such control and husbandry is also evidenced within this study in the form of archaeological survey and excavation of a potential corral enclosure at Botai.

Of the possible routes to animal domestication that have been defined *(75)*, the *'prey route'* appears to fit the evidence at Botai best. In such a scenario, peoples familiar with hunting a species changed their relationship with the hunted species over time, displaying more interest in protecting the future viability of the herd and controlling its movements. The evidence at Botai suggests that this husbandry extended to the keeping and slaughtering of at least some horses within settlements and some use of secondary products including milk and harnessing. Such horse pastoralism allowed settlement size to grow and become more sedentary. However, the evidence as it stands need not invoke the full range of equestrian activities associated with later steppe cultures or any form of intensive animal breeding. At Botai, it is possible to envisage a relationship more akin to reindeer domestication *(68)*, though there are clear environmental differences.

Section 3. Archaeological Analyses

Survey and excavation at Botai with specific reference to an enclosure feature

Accurate total station survey, complemented by GPS and lensatic compass sightings relating to a permanent datum, has provided precise spatial control at Botai since 2008, and permitted the accurate integration of data from the extensive magnetic gradiometry survey, the coring and geochemical survey which complemented the soil micromorphology study, and past and present excavations. The margin of error in resurvey over the years has been ±0.1 meter for every 100 meter, with the resurvey structured so as to mitigate the possibility of cascading error.

Magnetic north (11° 19' E of true north) was used so as to remain consistent with excavation alignment over the past decades. The gradiometer survey was undertaken in 2008 using a Bartington Grad601-2 dual-array fluxgate gradiometer with data logger, travelling in a zigzag pattern for maximum efficiency of ground coverage. Traverses were set a 1 meter intervals, with readings every 0.125 meter within grid squares of 30 x 30 meters. The maximum north to south extension of the survey area was 360 meters and 540 meters east to west. Within that, the total area surveyed was 16,200 square meters covered by 129,600 readings. The data were processed and graphically rendered using GeoPlot4 software. The full survey is presented in Fig. S1. The dark sub-circular anomalies are Botai pit-houses that are often obvious from the surface, and many of these have been excavated. Such houses often have pits close by. An interpretation of the survey highlighting potential archaeological anomalies, as well as known natural and modern features and positions of old excavation trenches, is presented in Fig. S2. While other potential enclosure features can be identified in the survey, this diagram only highlights the anomaly that has been subject to further investigation.

Relevant to the issue of prehistoric horse husbandry, such an enclosure could have served as a livestock corral and is indicated by both linear trench and posthole-sized geophysical anomalies, and a relative increase in phosphorus levels, encompassing ~500 square meters located with its most extreme coordinates from the survey datum at 17 meters west to 6.75 meters east and 91-118 meters north. Excavations to explore this area were undertaken in 2011, 2012 and 2013 and the positions of these in relation to the geophysical anomaly are presented in Fig. S3. The first was a narrow 5 x 1 meter excavation unit (Fig. S3, excavation unit B) cutting across the apparent boundary of the enclosure. This revealed a distinct prehistoric trench-like feature in cross-section, first identified in plan as a markedly more compact layer of sediment, as well as several clear postholes in profile, and one 20-centimeter-diameter posthole in plan, the fill of which contained horse bones, lithic material, and the distinctive pinkish clay sediment which is associated with the prehistoric living surface at Botai. The northeast section of this trench is presented in Fig. S4. It is notable that the larger trench and postholes also appeared to contain postpipes in profile. A postpipe is created where a wooden post set in a wider hole or trench rots away in situ or is removed and infilled by different sediment to the packing material thus showing the original size and position of the post.

The following year, more extensive excavations, including a 6 x 6 meters block opened to the northeast of the test trench revealed further satisfactory ground-truthing of the magnetic gradiometry survey (Fig. S3, excavation unit A), with a clear linear pattern of three postholes, one of which contained horse ribs aligned almost as if used as shims to steady a post, while another contained an articulated equine lower limb as part of its fill. The latter may be an obvious example of a quick, purposeful posthole in-fill, in keeping with responsible horse husbandry, which limits open postholes and similar hazards to equine limb fracture. These bones produced a radiocarbon date range of 3,629-3,370 cal. BCE (2σ, UBA-33488), confirming that these features are contemporaneous with the Botai settlement. This posthole alignment follows through into excavation unit B. The more continuous trench feature continues in parallel to the line of postholes, suggesting that the enclosure was repaired, rebuilt, or reinforced over time. This trench perhaps represents the merging of close set postholes, or a palisade trench, and has likely been re-cut and re-set several times. The excavated trench line and posthole alignment can be seen in Fig. S5.

Further small test trenches excavated in 2013 (see Fig. S3) confirmed the presence of linear features along the line of the geophysical anomaly and in several places supported the interpretation that the feature had been recut on slightly different alignments over time. The adjacent house excavation trench (Fig. S3) showed two lines of small trenches approaching the house from the northeast and meeting with it. Once again, this most likely represents re-cutting of the feature, rather the simultaneous existence of a double line.

It is possible to identify a number of other potential posthole alignments and other enclosure lines. However, each will need careful ground-truthing. One particularly obvious further candidate lies to the immediate east of the enclosure already documented, just to other side of the adjoining pit house. This feature (highlighted in Fig. S6) is slightly less distinct but of similar appearance and shape, though slightly larger. It is notable that this area is avoided by the otherwise dense arrangement of Botai pit-houses.

Section 4. Morphological Analyses

We followed Outram and colleagues *(8)* and compared a total number of 262 ancient and 112 extant horses for five individual morphological measurements of the third metacarpal (MC-III). These correspond to a sub-selection from 12 key measurements outlined by Vera Eisenmann (full details available at *(76)*: the Greatest length (measurement #1), the Breadth at the middle of the diaphysis (measurement #3), the Proximal breadth (measurement #5), the Distal supra-articular breadth (measurement #10) and the Distal articular breadth (measurement #11). For a total number of 234 specimens, the morphological measurements were compiled from previous publications, directly downloading data from Vera Eisenmann's website *(76)*. This includes: *(i)* 24 Upper Paleolithic specimens from Jaurens, France; *(ii)* two Mesolithic specimens from Mirnoie, Ukraine; *(iii)* four specimens from Beijkush, Ukraine, associated with the so-called Archaic period; *(iv)* two specimens from Tripolie, Ukraine; *(v)* one Early Bronze Age specimen from Roucadour, France; *(vi)* three Bronze Age specimens from Avignon, France associated with the Bell Beaker culture; *(vii)* one Bronze Age specimen from Soleb, Egypt; *(viii)* one Iron Age specimen from Portugal; *(ix)* twenty Iron Age specimens from Sredni Stog, Ukraine; *(x)* twelve Hellenistic specimens from Olvia, Ukraine; *(xi)* three specimens from Stebliev, Ukraine, dating back to the 5th century BCE; *(xii)* one Scythian specimen from Rusalka, Ukraine, dating back to the 4th century BCE; *(xiii)* thirty one Iron Age specimens from France and Switzerland and associated with the La Tène culture; *(xiv)* three specimens from Skel'ka, Ukraine, spanning the period from the 1st century BCE to the first century AD; *(xv)* one Roman specimen from Demianov, Ukraine, dating back to the period spanning the 3rd to the 5th century AD; *(xvi)* one Roman specimen from Didova Khata, Ukraine, dating back to the period spanning the 3rd to the 4th century AD; *(xvii)* four Gallo-Roman specimens from Compiègne and Ribemont, France; *(xviii)* one specimen from Vyshgorod, Ukraine, dating back to the 10th century AD; *(xix)* six specimens from Voiyn, Ukraine, spanning the period from the 11th to the 13th centuries AD; *(xx)* one Turkish specimen from the 13th-14th century AD, and; *(xxi)* 109 modern individuals, including 58 domesticates and 51 Przewalski's horses. The morphological data collected by Outram and colleagues *(8)* for an additional number of 110 specimens was also used. These include: *(i)* 18 Eneolithic specimens from Botai, Kazakhstan, and associated with the Botai culture; *(ii)* 12 Tersek specimens from Kozhai, Kazakhstan, and

associated with the Tersek culture, a culture contemporary and similar to the Botai culture, except that it was located further west and less horse-dominated *(8)*; *(iii)* 41 Eneolithic specimens from Kumkeshu, Kazakhstan, also associated with the Tersek culture; *(iv)* 36 specimens from Kent, Kazakhstan, associated with the Bronze Age period (1,300 BCE – 900 BCE), and; (v) three modern specimens, including two domesticates and one Przewalski's horse. Finally, we collected morphological measurements for 31 Magdalenian specimens, including one from the Monruz archaeological site (Switzerland, *(77, 78))*, one from Le Grand Canton (Parisian Basin, France, *(77, 78))*, ten from Le Queroy (Charente, France, *(79))*, two from Trou Da Somme and one from Trou De Chaleux (Belgium), 13 from Gonnersdorf (Germany, *(80))*, and three from Pekarna (Czech Republic, *(81))*. All raw morphological measurements are provided in Table S10. The distribution of robustness indices, represented by the ratio of the Breadth at the middle of the diaphysis (measurement #3), the Proximal breadth (measurement #5), the Distal supra-articular breadth (measurement #10) and the Distal articular breadth (measurement #11) over the Greatest length (measurement #1) are provided as boxplots in Fig. S7 for 18 geographical and/or temporal categories.

Section 5. Molecular Analyses: DNA extraction and sequencing

DNA extractions were performed in the ancient DNA facilities of the Centre for GeoGenetics, University of Copenhagen (Denmark) using strict ancient DNA guidelines and following the extraction method referred to as Method Y in *(82)*, with slight modifications. DNA was extracted from 130-780 milligrams of bone or tooth powder. First, a pre-digestion step at 37°C for 1h in 4 ml of lysis buffer consisting of EDTA 0.45M, N-lauryl Sarcosyl 0.5% and Proteinase K 0.25 mg/ml was carried out. Then, the undigested pellets were digested overnight in a freshly prepared lysis buffer at 42°C, and DNA extracts were recovered using the supernatant of the second digestion. For the vast majority of the DNA extracts, an aliquot of 23µl was incubated at 37°C for 3 hours with 7µl of USER enzyme (NEB®) mix, a mixture of Uracil-DNA Glycosylase and Endonuclease VIII, to reduce the impact of post-mortem cytosine deamination on subsequent downstream analyses.

Blunt-end Illumina libraries were built following two different protocols, referred to as A and B. Method A followed the procedure described by Meyer and Kircher *(83)*, as modified in *(82)*. Method A is distinct from Method B in that we introduced a unique index of seven nucleotides within adapter P5 and within adapter P7 before their ligation to DNA templates. Each library template thus contained two specific index sequences before further PCR indexing.

After library preparation, a 20X dilution of each library was subjected to quantitative real-time Polymerase Chain Reaction (qPCR) on a Roche LightCycler 480 Real-time PCR System to determine the minimal number of PCR cycles required for subsequent amplification, as described in *(82)*. The numbers of amplification cycles performed during the PCR of each individual DNA library are shown in Table S11. These amplifications were carried out as described in *(83)* in 25µl final reaction volumes and using 1 unit of AccuPrime™ Pfx DNA polymerase, 3 to 6µl of unpurified DNA library and with an overall concentration of 0.2µM of each primer, including the InPE1.0 primer and one custom PCR primer. The latter includes a 6-bp sequence tag index used for sequence demultiplexing. Indexed DNA libraries were purified using Minelute columns

(Qiagen(C)), eluted in 25µl of EB (10 mM Tris-Cl, pH 8.5) + 0.05% Tween, and quantified on the Tapestation 2200 instrument (Agilent technologies). Purified DNA libraries were then pooled together and sequenced on either Illumina HiSeq2500 (method A) or Illumina HiSeq4000 (method B) at the Danish National High-Throughput DNA Sequencing Centre. Alignment of the sequencing data against the horse reference genome EquCab2 *(84)* are available on the European Nucleotide Archive, project number PRJEB22390.

For the majority of samples sequenced in this study, we performed multiple independent amplifications of DNA libraries with various index primers (Table S11A and Table S11B). The goal was to limit the overall clonality of sequencing reads (that can sharply increase with the amount sequencing data) by tracking down the clonality of each sequenced library and adapting the sequencing effort allocated to each library accordingly.

Section 6. Computational Analyses

6.1 Read alignment, rescaling and trimming

One to seven DNA libraries were generated per sample (Table S12), the vast majority of which (154/156, ~98.7%) from ancient DNA extracts that were treated with the USER ™ enzyme mix in order to limit the impact of nucleotide mis-incorporations at deaminated cytosines in downstream analyses *(85)*. Sequencing reads were processed for each individual library through PALEOMIX version 1.1.1 *(86)* with default parameters, excepting that seeding was disabled. In PALEOMIX, sequencing reads were first trimmed for known adapter sequences, low quality termini and filtered out when shorter than 25 nucleotides, and then aligned against the horse mitochondrial reference sequence (Genbank Accession Nb. NC_001640, *(87)* and the horse nuclear reference sequence (EquCab2; *(84)*) appended for 2,797 Y-chromosome contigs (concatenated in a single scaffold, using 500 N bases as separators *(88)*. Read trimming was performed with AdapterRemoval2 (*89)* while read mapping is carried out with BWA version 0.5.9-r26-dev *(90)*, disabling seed, disregarding alignments showing mapping qualities inferior to 25. PCR duplicates were then removed based on 5' read coordinates for Single End sequencing reads, and both read starts and read ends for (collapsed) paired end data. Finally, reads were locally realigned around indels using the IndelRealigner procedure from GATK *(91)*.

The presence of nucleotide mis-incorporation profiles characteristic of ancient DNA data was verified at the library level using 100,000 randomly selected reads and mapDamage2 *(92)*. The expected increase of C→T (G→A) mis-incorporation rates at read starts (read ends) was observed for both for USER-treated and non-USER treated data, given that the experimental conditions used while treating ancient DNA extracts with the USER ™ enzyme mix were chosen to avoid removing all mis-incorporations, following the work from Rohland and colleagues *(93)*. Additionally, the genomic positions preceding read starts were enriched in purines for non-USER read alignments (Fig. S8A and S9A), in agreement with depurination driving most post-mortem DNA fragmentation) *(94)*. For USER-treated read alignments, these positions were enriched in cytosine residues (Fig. S8B and S9B), in line with the excision of deaminated cytosines by the

sequential activities of Uracil DNA glycosylase and Endonuclease VIII enzymes present in the USER mix.

We further applied the following computational procedure to limit the impact of remnant nucleotide mis-incorporations in downstream analyses. First, we identified the first nucleotide position from read starts (ends) where the mis-incorporation rate for all mis-incorporation types except C→T (G→A) is below 0.001. To achieve this, we used the mis-incorporation counts of each substitution type, as provided by mapDamage2 *(92)*. Read positions located upstream were flagged for trimming. Second, using the same nucleotide mis-incorporation profiles as above but disregarding positions flagged as trimmed, we identified the read position from which nucleotide mis-incorporation rates are approximately uniformly distributed across read positions. To determine such positions, we first constructed a linear model between the cumulative distribution of C→T mis-incorporations and read positions, and then identified the minimal value for the model residual towards read ends. As mapDamage profiles are provided by default along 70 nucleotide positions, read ends were defined here to span positions 36 to 70 (minus the amount of positions identified as to be trimmed in the first step). The minimal model residual provides a conservative estimate for how fast C→T mis-incorporations should accumulate along successive read positions. We thus identified the first read position from read start where the accumulation was no longer superior or equal to this threshold. The region comprising read start to this position accumulates an abnormal excess of C→T mis-incorporations and was thus flagged for base quality rescaling *(92)*. The same procedure was implemented for rescaling read ends, excepting that: (1) read starts (first 36 read positions minus the bases flagged for trimming at read starts) were used instead of read ends to define the acceptable threshold, and; (2) G→A mis-incorporations were used instead of C→T mis-incorporations. Once all positions that required trimming and/or base quality rescaling were identified, we first ran the rescaling procedure and then trimmed aligned reads. Rescaling and trimming was applied to positions where reads were effectively aligned against the reference genome (deletions and soft-clipped bases were disregarded). For each sample, all individual rescaled and trimmed BAM alignments were merged together using the MergeSamFiles command from Picard Tools *(95)*.

Per genome sequencing error rates were calculated on each resulting read alignment file, following the procedure described in previous work *(15, 20, 21, 53)*. The full procedure from read alignment to base rescaling, and alignment trimming was applied to the sequencing data underlying a total of 28 modern and 18 ancient horse genomes previously published, as well as the genomes of one domestic donkey and one Somali wild ass, used as outgroups in our analyses *(15, 20, 21, 52-64)*. The average sequencing error rate is provided in Fig. S10 and S11 for two minimal thresholds of base quality scores and for all 42 novel ancient genomes characterized in this study. The corresponding individual error rates observed for each substitution type are provided in Tables S13 and S14.

6.2. Organellar genomes

6.2.a. Mitochondrial genome

Read alignment BAM files were produced, base-quality rescaled and trimmed following the procedure described in section S6.1. For each individual sample, the mitochondrial haplotype sequence (fasta format) was reconstructed using the paleomix vcf_to_fasta command applied to a unique VCF variant file. The VCF file was obtained using the samtools mpileup command, disregarding sites showing base quality strictly inferior to 25, and further filtered using bcftools (version 1.1) for sites showing coverage strictly inferior to three. Each individual fasta file was copied into a unique multifasta alignment file containing all other mitochondrial haplotypes produced in this study as well as 169 previously published *(15, 21, 53, 55-63, 96-101)*, representing a total of 211 mitochondrial sequences (see Table S15). Polymorphic sites were further visually checked and validated for each putative polymorphic site using the samtools tview command. Starting from the multialignment file, a total of six sequence partitions were generated, including first, second and third codon positions for all 13 protein coding genes; ribosomal RNAs; transfer RNAs, and; the control region. A Maximum Likehood (ML) tree was reconstructed using RAxML (multithread version 8.2.4, *(102)*) and applying a different GTR+GAMMA substitution model to each individual partition. Node support for the best ML tree was estimated from a total of 100 bootstrap pseudo-replicates (Fig. S12). The tree was rooted following the root placement obtained from BEAST analyses, applied to the same six partitions but selecting the best substitution model for each partition using modelgenerator (version 0.85, *(103)*) and a Bayesian Information Criterion (note that some of such models are not presently implemented in RAxML and could, thus, not be selected in the analyses described above). For BEAST analyses (v1.8.0) *(104, 105)*, we considered both a Bayesian Skyline (linear change) and an extended Bayesian Skyline demographic model. Additionally, we considered a strict and a relaxed LogNormal correlated molecular clock for each individual partition. The best of the four models investigated was selected using the aicm procedure as implemented in Tracer (version 1.6), disregarding 25% of 617 million generations (sampling frequency 1/2,500). The posterior distribution for the time to the Most Common Recent Ancestor (MCRA) of all mitochondrial haplotypes considered was estimated from the same subset of generation sample in Tracer (version 1.6, *(106)*, Table S3). The final consensus tree was produced using TreeAnnotator (v1.8.0, *(104),* Fig. S13) while the Bayesian Skyline demographic profile (LogNormal correlated) was generated in Tracer (version 1.6) and plotted with R *(107)* (Fig. S14).

6.2.b. Y-chromosome

We extracted all reads mapping against the concatenated Y-chromosome contigs described by Wallner and colleagues *(88)* and three non-repetitive parts of the Y-chromosome (Genbank Accession Nb. AC215855.2 and JX565703). These were prepared as a single scaffold consisting of 2,797 contigs covering 1,685,390bp, and separated by 500 Ns. We used ANGSD (0.917, last accessed march 6th, 2017) *(108)* to generate a fasta consensus based on the majority rule for each stallion individual. We disregarded reads showing mapping quality scores inferior to 25 and bases showing quality scores inferior to 20. Additionally, we used Base Alignment Quality (-baq 1), adjusted mapping qualities for excessive mismatches (-C 50) and removed reads with multiple best hits (-uniqueOnly 1). Individual consensus sequences were generated both considering all sites covered at least once (-setMinDepthInd 1) and all sites covered at least twice (-setMinDepthInd 2). Each set of consensus sequences was merged into a multifasta alignment file that was further used to reconstruct an ML phylogenetic tree with RAxML (multithread version 8.2.4) *(102)* under

the GTRCAT substitution model. A total of 100 bootstrap pseudoreplicates were performed to assess the robustness of each individual node of the tree (Figs. S15 and S16). Additionally, for each set of consensus sequences, we scored the relative amount of missing data for each Y-chromosomal haplotype, relative to the one showing maximal coverage. We then only considered those individuals with no more than 50% data missing and reconstructed an ML phylogenetic tree in RAxML using a GTRCAT substitution model. This tree was used in the Evolutionary Placement Algorithm (EPA) *(109)* together with the full multifasta sequence alignment to assess the most likely phylogenetic placement of those individuals showing no more than 50% data missing. Importantly, the EPA procedure revealed that the African wild ass (*Equus asinus* Donkey_0233A, *(64)* that was used as outgroup was difficult to place with accuracy as 128 nodes (90 nodes) showed likelihood weights ranging between 0.007164 and 0.007506 (0.009970-0.010656) when considering the multifasta alignment file generated from all sites covered at least once (twice). Therefore, the whole set of analyses was replicated excluding the sequence of the African wild ass (*Equus asinus* Donkey_0233A, *(64))*. The resulting phylogenetic trees were in agreement, regardless of the methodology used and are shown as Fig. S17 (minimum coverage = 1) and Fig. S18 (minimum coverage = 2).

6.3. Autosomal chrosomomes

6.3.a. Filters for genomic analyses

We used ANGSD v0.917 (*108*) to conduct all genomic analyses with these filtering options:
- Minimum base quality: 20 (-minQ 20)
- Minimum mapping quality: 25 (-minMapQ 25)
- Remove all no primary, duplicated, and unmated reads (-remove_bads 1)
- Remove all reads showing multiple hits (-uniqueOnly 1)
- Downgrade quality scores around indels (-baq 1)
- Adjust mapping qualities in regions with excessive mismatches (-C 50)

We explored the impact of base alignment quality (BAQ) and extended BAQ (eBAQ) *(110)* on low coverage samples. We down-sampled five genomes, including two ancient and three modern horses. Using ANGSD, we estimated the heterozygosity levels on chromosome 31, from the original depth-of-coverage to 1.0X (Fig. S19). We found that considering eBAQ results in filtering fewer variants than the original BAQ. This is especially true for coverage lower than 10.0X, where the eBAQ procedure leads to an over-estimate of the nucleotide diversity. The original BAQ, instead, produces a more stable heterozygosity estimate from 2.0X coverage and upwards. As the original BAQ is comparable across samples of different coverages, we used this procedure throughout all analyses.

As calling genotypes from low coverage sequence data (< 3.0X depth of coverage) is challenging, we applied two different methodological approaches. The first is based on generating pseudo-haploid genomes by random sampling alleles. This provides unbiased genetic affinity

estimates. Yet, certain population genomic analyses implicitly leverage on the levels and patterns of nucleotide heterozygosity, most notably including inbreeding and relatedness estimates *(111, 112)*, TreeMix population trees *(12)*, drift-based ancestry test *(14)*, and positive selection scans. Therefore, we implemented a second approach, which is based on each individual genotype probability, for 60 samples with a depth of coverage greater or equal to 3.0X.

6.3.b. Mendelian traits

We investigated the genotypes at 43 loci associated with phenotypically important traits, determining coat-color variation, body size, diseases, and racing performance *(113-141)*. For all horses included in this study, we surveyed their shotgun sequencing data. For Berel' and Arzhan stallions, we also integrated capture-enrichment sequencing data generated in *(21)*, targeting ~5,000 loci that include those Mendelian traits *(142)*.

We calculated the genotype likelihoods with ANGSD v0.917. Genotype likelihoods enabled to estimate the probability of the allele associated with the phenotype. The likelihood of the causative/associated allele was conservatively downscaled to zero if solely supported by a single read. Resulting allele probabilities are represented in Fig. S20.

In addition, the C->T mutation, at genomic coordinate chr1:108,249,293, has been traditionally associated with the Leopard spotting complex (LP), and with a nonprogressive visual deficit in dark conditions, known as stationary night blindness (CSNB) *(143)*. The chr1:108,249,293 mutation was present in a single heterozygous individual at Borly4 (Borly4_PAVH11_5015) but three Botai horses (Botai_1_5500, Botai_4_5500 and Botai_D4_5500) were homozygous, possibly suggesting limited visual capacity at night. However, we investigated four additional SNPs strongly associated with LP and CSNB across different breeds *(144, 145, 146)*, including one modifier locus located in chromosome 3, *RFWD3*, that alters the extent of white in LP horses *(18)*. In these additional SNPs, Botai_1_5500, Botai_4_5500 and Botai_D4_5500 were found to carry the ancestral alleles, supporting thus non-LP and non-CSNB phenotypes. Botai_5_5500, however, was heterozygous for the three mutations around *TRPM1*, most consistent with the presence of only one copy of the full haplotype associated with LP and CSNB. This individual was thus most likely leopard spotted but did not suffer CSNB.

It has also been suggested that a retroviral insertion (LTR) within *TRPM1* is causative for *CNBS* and *LP*, which disrupts gene transcription by incorporating pre-mature poly-A tails *(17)*. As this LTR insertion is not present in the reference assembly, we edited the sequence of EquCab2.0 *(84)* by inserting the 1,378 bp of LTR into *TRPM1* (starting at coordinate chr1:108,297,929). Then, raw sequencing reads were mapped against this region, without quality filters and without removing PCR duplicates. Relaxing mapping quality filters ensured high sensitivity in recovering read candidates potentially mapping to the LTR insertion. These candidate reads, however, could equally map to other genomic regions. To control for specificity, therefore, we finally remapped these read candidates against the whole reference sequence, as modified to include LTR. Finally, we counted the number of reads with a minimum mapping quality of 25 that extended over more

than 10 bp inside the LTR insertion breakpoints, located at coordinates chr1:108,297,929 and chr1:108,299,306 of the modified reference.

It is worth mentioning that the sequence of the LTR insert is assumed to be the same as that determined for modern domesticates. But if the origin of this LTR insertion predates the split between the first and second domestic clade, it likely accumulated a large amount of mutations, especially taking into account that is a retrovirus of 1,378 bp, where only a few small poly-"A" motifs are sufficient to provide phenotypic variation. Taken together, this makes the genetic identification of the LTR causative for *CNBS* and *LP* challenging in ancient samples. We nonetheless found that Botai_5_5500 showed reads supporting both LTR insertion termini.

A similar approach was carried out for *TBX3 (147)* and *MSTN (148)*, two genes that underlie the wild-type Dun color and sprint racing performance, respectively. No Botai or Borly4 horse was found to carry the SINE insertion at the *MSTN* promoter. Results for the Dun loci can be found in Table S4.

6.3.c. Genetic load

Modern domestic horse genomes carry an excess of harmful mutations, relative to the wild, archaic horse lineage *(20)*. The genomes of 14 ancient Scythian horses revealed that these deleterious mutations accumulated during the last 2,300 years, later than expected under the '*cost of domestication*' hypothesis *(21)*.

The latter statement assumes as correct the date that generated more consensus for horse domestication so far; the ~5,500 kya associated with Botai culture/site *(8)*. We have shown, however, that archaeological evidence in Botai does not correspond to the domestication event that lead to modern horse domesticates. Exploiting the 42 ancient genomes generated in this study, we decided to re-investigate the '*cost-of-domestication*' hypothesis. Additionally, we explored whether the horse genetic load changed during the time interval when Botai-type horses returned feral and become the Przewalski's horses. The latter experienced a massive bottleneck during the 20th century AD *(14)*, which likely limited the effectiveness of negative selection to filter out deleterious mutations *(149)*.

In order to estimate the deleterious load of each individual genome, we followed the same approach as in *(21)*. This requires the probability that sites are mutated, and the fitness (deleterious) effects of these mutations. For the former, we relied on the genotype likelihoods estimated by ANGSD (-doGlf 4), with all quality filters described above. For the latter, we assumed that mutations in evolutionarily constrained positions are more likely deleterious. We used PhyloP scores as a proxy for evolutionary constraint, measured from an alignment of 46 placental species *(150)*. Averaging the probabilities of being mutated (genotype likelihoods) by their corresponding fitness consequences (PhyloP score), and then dividing by the total number of analyzed positions, we estimated mean load estimates per genome. Of note, we conditioned on homozyous positions, in order to avoid specifying dominance coefficients (*i.e.* recessive, intermediate, *etc*). To more precisely discriminate between homozygous and heterozygous genotypes, we only analyzed sites

with a minimum depth-of-coverage equal or higher than three, a threshold shown to provide stable load estimates.

6.3.d. Pseudo-haploid genomes using a random allele sampling approach

Based on the number of read observations passing all quality filters, one allele was then randomly selected at every genomic site for every horse. Biallelic sites observed as transitions prior to sampling where disregarded, in order to remove potential residual of post-mortem DNA damage. We allowed up to three individuals with missing data per site leading to a total of 14,050,354 segregating sites available for downstream analyses. With this dataset, three analyses were conducted, namely a Principal Component Analysis (PCA), and phylogenetic inference based on pairwise genetic distances, and admixture analyses.

A similar approach was carried out for calculating outgroup-f_3 and f_4 statistics, with the addition of the Somali wild ass, *Equus somalicus* as an outgroup (Somali_0226A) *(64)*. These analyses were based on a total number of 16,204,406 nucleotide transversions.

6.3.e. Outgroup f_3 statistics

In order to quantify the amount of shared genetic drift between pairs of individuals, we calculated the f_3 statistics in the form (X, Y; Outgroup), where X and Y represent all possible pairwise combinations (n=3,828) of horses included in this study. We used the dataset, with 16,204,406 nucleotide transversions, including the outgroup (*E. somalicus*). The f_3-statistics was computed using an in-house C++ programme, available upon request. Results are shown in Fig. S21 for the whole set of individuals.

6.3.f. Admixture Graph based on f_4 statistics

Admixture graphs extend classical phylogenetic inference by allowing migration edges amongst nodes, accounting for the modelling of more complex evolutionary histories, including population splits and gene flow. Following the genetic affinities recovered by PCA analysis and phylogenetic inference (section 6.3.i), we defined nine populations, one for archaic horses, four within the Botai/Borly4 clade, and four for the second domestic clade (DOM2).

The four populations within the Botai/Borly4 clade include Botai (n = 20), Borly4 (n = 5), modern Przewalski's (n = 6) and a Przewalski's horse from the 19th century (n = 1). The Przewalski's horse from the 19th century, Przewalski_Paratype_118, was treated separately as it predates the last bottleneck experienced by this lineage and the well-documented introgression of domestic ancestry *(15)*.

Modern domesticates within the second domestic clade (DOM2) were split into two subgroups. Subgroup A consists of Yakutian_0163A, Yakutian_0170A, Yakutian_0171A, Mongolian_0153A, Mongolian_0215A, Jeju_0275A and Tumeski_CGG101397_192 (n = 7). The latter is a historical specimen from the 19th century AD, but shows higher affinities with modern

Yakutian horses *(52)*. Subgroup B embraces all remaining 16 modern domesticates. Outgroup-f_3 analysis and phylogenetic inference revealed Dunaujvaros_Duk2_4077 as divergent to all horses from the second domestic clade (Fig. 2A-B). Therefore, although archaeological evidence indicates Dunaujvaros_Duk2_4077 was a domestic horse, this specimen was treated independently.

Using the unsupervised brute-force approach implemented in admixturegraph *(16)*, we searched for the topology best fitting the f_4 estimates (Fig. S22). Botai and Borly4 were consistently found as ancient members of the Przewalski's lineage, and Dunaujvaros_Duk2_4077 as the most basal specimen within the second domestic clade. Optimizing f_2 parameters revealed that the genetic drift from Dunaujvaros_Duk2_4077 to the rest of ancient domesticates is large (f_2 = 0.00026) with regards to the limited period of time separating the age of this sample and the node ancestral to all remaining members of the DOM2 lineage. Following the ANGSD recipe in *(151)* with our filtering scheme, we found no major difference in the genome-wide levels of heterozygosity between Dunaujvaros_Duk2_4077 (0.0012) and the rest of ancient domestic horses (mean individual heterozygosity = 0.0014), ruling out a severe population decline as plausible explanation for this f_2 estimate. The divergence of Dunaujvaros_Duk2_4077 within the DOM2 lineage is rather suggestive of genetic ancestry from an unknown 'ghost' population within this sample.

In a D-statistics (((X, Dunaujvaros_Duk2_4077), Archaic), Donkey) configuration, where X represents any ancient or modern domestic horse, introgression from a 'ghost' population into Dunaujvaros_Duk2_4077 can reduce the number of ABBA counts (where A and B refer to ancestral and derived alleles, respectively), provided that the ghost population is distantly related to the three archaic horses previously described. A relatively low number of ABBA counts could then lead to negative D-statistics (BABA patterns are not affected by the ghost population in absence of admixture with X), which could be misinterpreted as evidence for gene flow between archaic and X horses. To avoid misleading conclusions due to unsampled ghost populations, we excluded Dunaujvaros_Duk2_4077 prior further testing of six alternative models (Fig. 3A-F), including:

(i)   no admixture,
(ii)   recent introgression from domesticates into Przewalski's horses,
(iii)   introgression from Botai ancestors into the lineage leading to ancient domesticates,
(iv)   introgression from Borly4 ancestors into the lineage leading to ancient domesticates,
(v)   introgression from Botai ancestors into the lineage leading to ancient domesticates, plus recent introgression from domesticates into Przewalski's horses,
(vi)   introgression from Borly4 ancestors into the lineage leading to ancient domesticates, plus recent introgression from domesticates into Przewalski's horses.

On top of these admixture graphs, we ran Markov Chain Monte Carlo (MCMC) chains with 200,000 iterations and thinning interval of 100. Posterior probabilities were visually

inspected, and the burn-in period was set to 120,000, the MCMC iteration corresponding to when model (iv) reached convergence. Model comparison was carried out based on the calculation of Bayes factors, as implemented in the '*model_bayes_factor_n*' function of the admixturegraph package *(16)*. Posterior means were used to summarize parameters of the best-fit model.

6.3.g. TreeMix trees

As TreeMix v1.13 requires genotype calls, we randomly sampled genotypes according to their posterior probabilities, using an in-house C script available upon request (beagle2treemix). Only positions sequenced at least once in each individual were considered, resulting in a total of 3,723,580 nucleotide transversions.

We applied TreeMix with the following parameters:

- A round of global rearrangements to refine suboptimal maximum likelihood inference (-global).
- A block resampling procedure of –k 500 SNPs (nucleotide transversions), which overpasses DNA fragments ~300 kb.
- Based on previous studies *(20, 52)*, topology was rooted in the lineage separating archaic horses from the rest (-root Batagai_5155, Taymyr_CGG10022_42758, Taymyr_CGG10023_16056).
- Turn off sample size correction to reduce over-fitting (-noss).

We explored from 0 to 5 migration edges (-m). The variance explained by each model is shown in Table S5. As increasing migration edges barely improved the model fit, only the population tree with no migration events is presented, in Fig. S23.

6.3.h. Inbreeding and relatedness

Horses establish natural herd structures characterized by a dominant stallion and multiple mares, as well as inbreeding avoidance. Horse populations showing signs of inbreeding and/or relatedness might be symptomatic of human-directed management. Identifying and filtering out related horses is also crucial methodologically, to avoid biases caused by non-random allele sampling.

We quantified the inbreeding (F) and kinship (θ) coefficients within Botai (n = 20) and Borly4 (n = 5) horses. We first calculated the posterior genotype probabilities with ANGSD, allowing up to 5% of missing sites. We used the genotype panel as input for ngsRelate *(111)* and ngsF v1.0.0 *(112)*. Expectation-maximization algorithms were used for likelihood optimization, with convergence criterions set, in both cases, to a likelihood improvement lower than 10^{-8}.

No Botai or Borly4 horses were found to be inbred, but two individual pairs were related, namely Botai_1_5500 and Botai_4_5500, and Botai_6_5500 and Botai_F_5500 (Table S6).

Botai_1_500 and Botai_F_5500, which showed an average depth-of-coverage of the genome inferior to that of their related counterparts, were excluded from subsequent analyses.

6.3.i. Genetic affinities

We analyzed the genetic affinities between the 42 samples generated in this study and a comparative panel of 46 published modern and ancient horses. First, we conducted a Principal Component Analyses (PCA) using ngsCovar [152], applying a minor allele frequency (MAF) filter of 0.1 (-maf 0.1). We also excluded Botai_1_5500 and Botai_F_5500, as they were inferred to be related to Botai_4_5500 and Botai_6_5500, respectively (Table S6). Results for the first three PCA components are shown in Fig. S24.

A phylogenetic reconstruction based on pairwise genetic distances was additionally carried out. We used ngsDist [153] to compute the pairwise genetic distances, upon which a neighbor-joining (NJ) tree was built with FastMe) [154]. The tree was rooted on the branch leading to Batagai_5155, Taymyr_CGG10022_42758, and Taymyr_CGG10023_16056 following [20] and [52]. Clade support values are based on 100 pseudo-replicate bootstraps. For these three analyses, we used the dataset with 14,050,354 nucleotide transversions, excluding the outgroup.

6.3.j. D-statistics

D-statistics, usually represented in the form $(((H_1, H_2), H_3), Outgroup)$, is a widely used to test for the presence of admixture between H_1-H_3 or H_2-H_3. We computed the D-statistics following [21]. This approach involves different filtering, because D-statistics is especially sensitive to genomes with moderate error rates (Figs. S10 and S11). Briefly, we stratified the read observations per position into A, C, G, and T's including their strand information leading to eight groups in total. For every group, we excluded half the read observations with lowest base qualities prior to random sampling of one allele. Only transversions were considered to reduce biases introduced by post-mortem DNA damage. The D-statistics is calculated in the form (ABBA-BABA)/(ABBA+BABA), where A and B refer to ancestral and derived alleles, respectively. We used a jackknife approach (10Mb window size) to obtain Z-scores. As Z-scores follow a standard normal distribution, we transformed them into p-values, which were in turn corrected for multiple testing using the Holm procedure [155]. Figures S25–S32 contain all configurations of the D-statistics, together with their Z-scores.

6.3.k. Drift-based tests for direct ancestry

At short time-scales, most of the allele frequency changes are due to genetic drift, which results in a loss of heterozygosity over time. By estimating drift parameters from a 2D site frequency spectrum (SFS), Rasmussen and colleagues [14] developed a simple method to test whether an ancient sample belonged to a population that was directly ancestral to a more recent one. The alternative considered consists of the ancient sample belonging to a population that diverged from the population that was effectively ancestral to the more recent one.

We estimated the 2D-SFS from the posterior probabilities previously calculated for the TreeMix analyses (section 6.3.g.), conditioning on sites with a MAF greater or equal to 0.1. The latter was done to filter out *de novo* mutations or sequencing errors, often found at low frequencies. Equations in *(14)* were implemented to search for the maximum likelihood drift parameters, using the R optim function as implemented in the R programming language. Multiple starting values were used to avoid local maxima during likelihood optimization. Results can be found in Table S16.

6.3.l. Impact of C to A (and G to T) error rates on genomic analyses

Although the overall error rates were inferred to be low, especially for nucleotide transversions, we noticed that a limited number of ancient genomes showed moderate C→A (and their reverse complementary G→T) substitution rates (Figs. S10 and S11). We evaluated whether this difference in overall genome quality could impact the characterization of the population affinities. After excluding C→A (and G→T) mutations, in addition to transitions, we found consistent results for the genome-wide tree inference and for the outgroup-f_3 statistics (Fig. S33). This indicates that our analyses are robust to the difference of genome quality present in our dataset. Of note, nucleotide transversions for the PCA and the direct ancestry test were filtered out for minimum allele frequencies greater than 0.1, which implicitly removed sequencing errors, as these often segregate as singletons. Similarly, the f_4 and D-statistics are relatively robust to sequencing errors, as they rely on ABBA-BABA patterns, and hence on doubletons.

6.3.m. Selection scans

Due to their untamed behavior, Przewalski's horses have been traditionally considered wild. Ironically, we found that they descend from horses that represent the first archaeological evidence of horse husbandry, ~5.5 kya *(8)*. During their feralization, traits promoted during horse husbandry could have been counter-selected in the wild, yielding genomic signatures of adaptation. We thus investigated the genomic changes associated with the return of Przewalski's horses to ferality.

In order to exploit the population and temporal structure present in our data set, we applied the LSD framework *(21)*. LSD scans for genomic regions that experienced significant changes in allele frequency along the branches of a given population tree. Based on their genetic affinities (Section 6.3.i and Fig. S22), the population tree was assumed to be the one displayed Fig. S34. Only genomes with a depth-of-coverage greater than 3.0X were considered. Together with the wild ass *E. somalicus* as outgroup *(64)*, this resulted in 61 genomes.

In LSD, changes in allele frequencies are directly estimated from a collection of gene trees, each representing a different genomic region. To build these local gene trees, we calculated the posterior genotype probabilities for SNPs segregating within the 61 samples. Only positions sequenced at least once in each individual were considered. We additionally disregarded SNPs representing nucleotide transitions, as 32 out of the 61 genomes belong to ancient samples, and post-mortem DNA damage could create spurious shifts in allele frequencies that resemble

signatures of adaptation. After filtering, this resulted in one beagle file per chromosome, making a total of 25,269,918 nucleotide transversions.

We sliced each beagle file into genomic windows of 10 kb, overlapping every 5 kb. The resulting 423,151 windows contained, on average, 32.33 nucleotide transversions. We next filtered out the 11,518 windows that showed less than ten nucleotide transitions. For the remaining 411,633, we calculated the pairwise genetic distances using ngsDist v1.0.2, and a NJ gene tree with fastme v2.1.4.

From these 411,633 local gene trees, we calculated the normalized LSD score (LSDnorm) in the branch immediately ancestral to Przewalski's horses, a time frame encompassing their feralization. To assess LSDnorm significance, we generated 1,000,000 replicates of 35 randomly selected SNPs, provided they represent nucleotide transversions at four-fold degenerate positions. Four-fold degenerate positions delineate the maximum LSDnorm score expected under neutrality. Indeed, mutations at four-fold sites do not involve an aminoacid replacement, and background selection in Przewalski's horses barely reduced neutral diversity at genic regions *(156)*. Empirical *p*-values were calculated for each LSDnorm score as the fraction of neutral replicates with LSDnorm scores strictly superior to the one observed. LSDnorm scores with a *p*-value lower than 0.0001 were considered significant. This revealed 342 windows potentially targeted by positive selection, overlapping the protein-coding exons of 107 genes (Table S7). Many of these gene candidates have been previously described as selected in the horse *(15, 20, 52, 157, 158)* and associated with other domestic animals, including with dairy and meat productivity (*159-171*).

We investigated whether those 107 genes are thus enriched in functional categories, relative to a reference set of background genes. For the latter, we used all protein-coding genes overlapping windows successfully analyzed by LSDnorm. Since mouse and human models represent the '*gold-standard*' reference for functional annotations, we relied on horse:human and horse:mouse single-copy gene orthologs, as defined by Ensembl Genes 88. Two different enrichment methods were applied, namely WebGestalt (*172*) and Bgee (*173*). Results can be found in Tables S17 and S18 respectively.

6.4. Species and Sex identification

The Zonkey computational package was applied with default parameters, as implemented in the PALEOMIX pipeline *(86)* to identify possible first generation hybrids among the samples as well as to determine the sex of each specimen (*174*). All the samples selected for this study were confirmed as pure horses, 31 of which are stallions and 11 are mares.
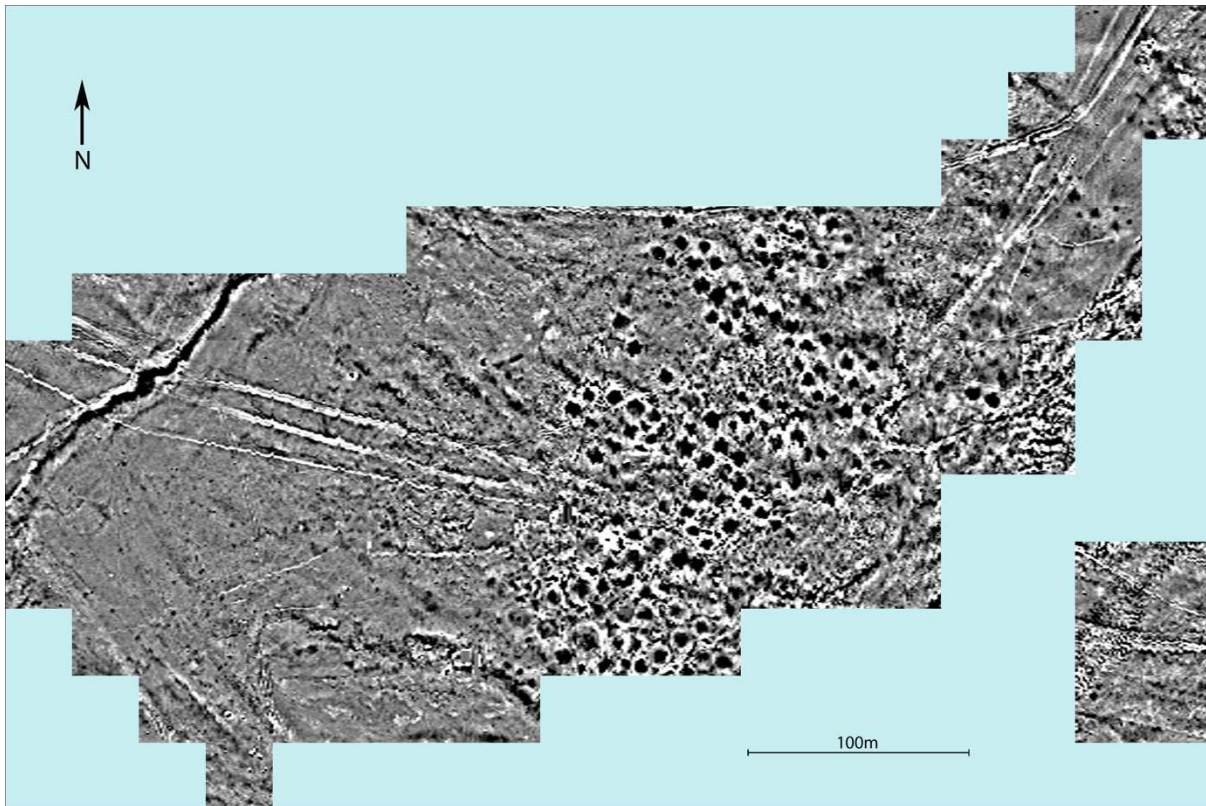
Supplementary Figures

Fig. S1.
Magnetic gradient survey of Botai following processing in GeoPlot4. We used the following parameters during processing in GeoPlot4: zero mean grid, zero mean traverse, clipping to archaeological sensitive range, selected destaggering to correct zig-zag effects and interpolation.
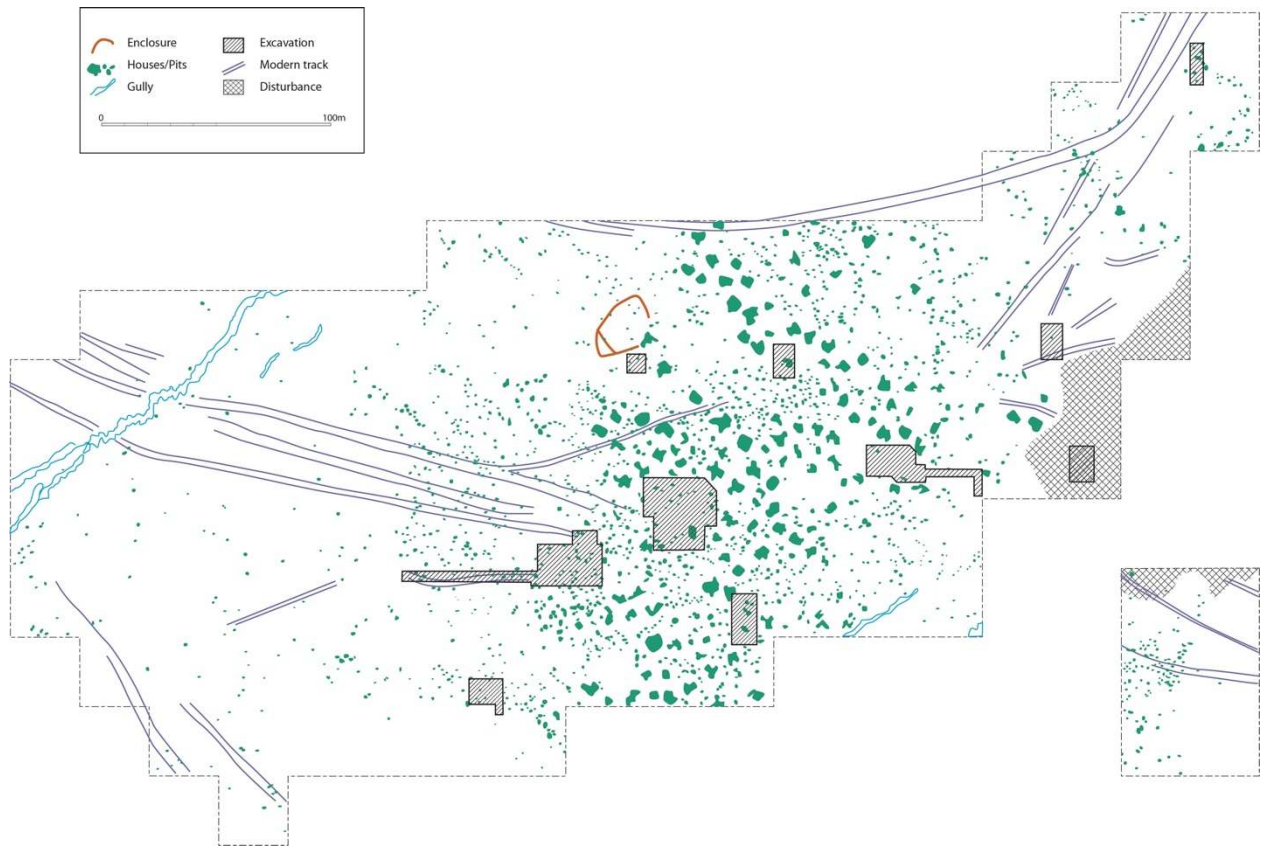
Fig. S2.
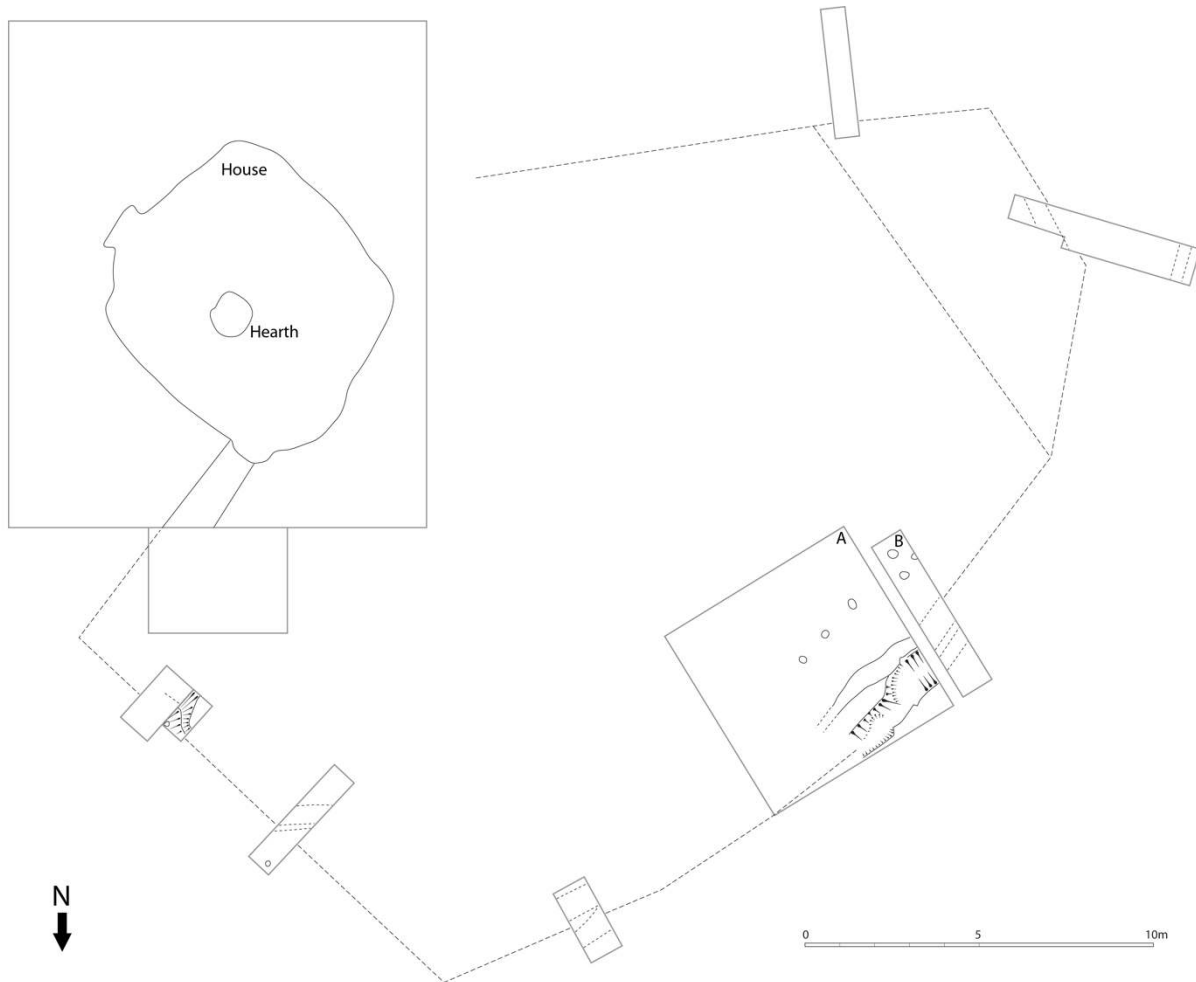Interpretation of the magnetic gradiometer survey with key.

Fig. S3.
Map of the excavation at Botai, showing the centre line of the geophysical anomaly as a dashed line, the excavation trench of an adjacent house excavated in 2013 and ground-truthing trenches excavated in 2011-13. Excavation unit A and B are represented by further illustrations.
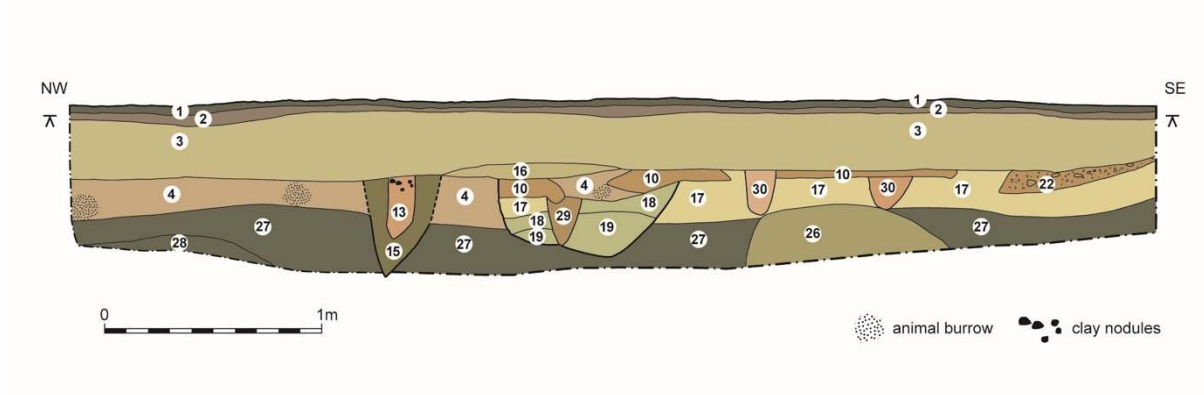
Fig. S4.
Detailed sediment profile of the northeast wall of 5 x 1 meter trench A from 2011 excavations cuts across the boundary of the probable corral area identified by magnetic gradiometery, exposing filled post-moulds bisected by excavation (fill sediments labeled P1, P2, P3). Post-mould P1 sits within a wider posthole, whilst there is infilling of a wider trench surrounding post-mould P2 that consists of multiple sediments immediately to the left and right of P2. The distinction between sediments labeled E and I is also significant as it is probably due to the different conditions on the prehistoric ground surface exterior (E) and interior (I) of the enclosure during its period of most intense or prolonged use. For excavation purposes, sediments were differentiated by color, texture, structure, compactness, inclusions, plasticity, and toughness -- attributes cumulatively differentiated and illustrated here by color.

Fig. S5.
Photograph looking northeast over excavation unit A. The line of the excavated trench feature and parallel postholes is shown.
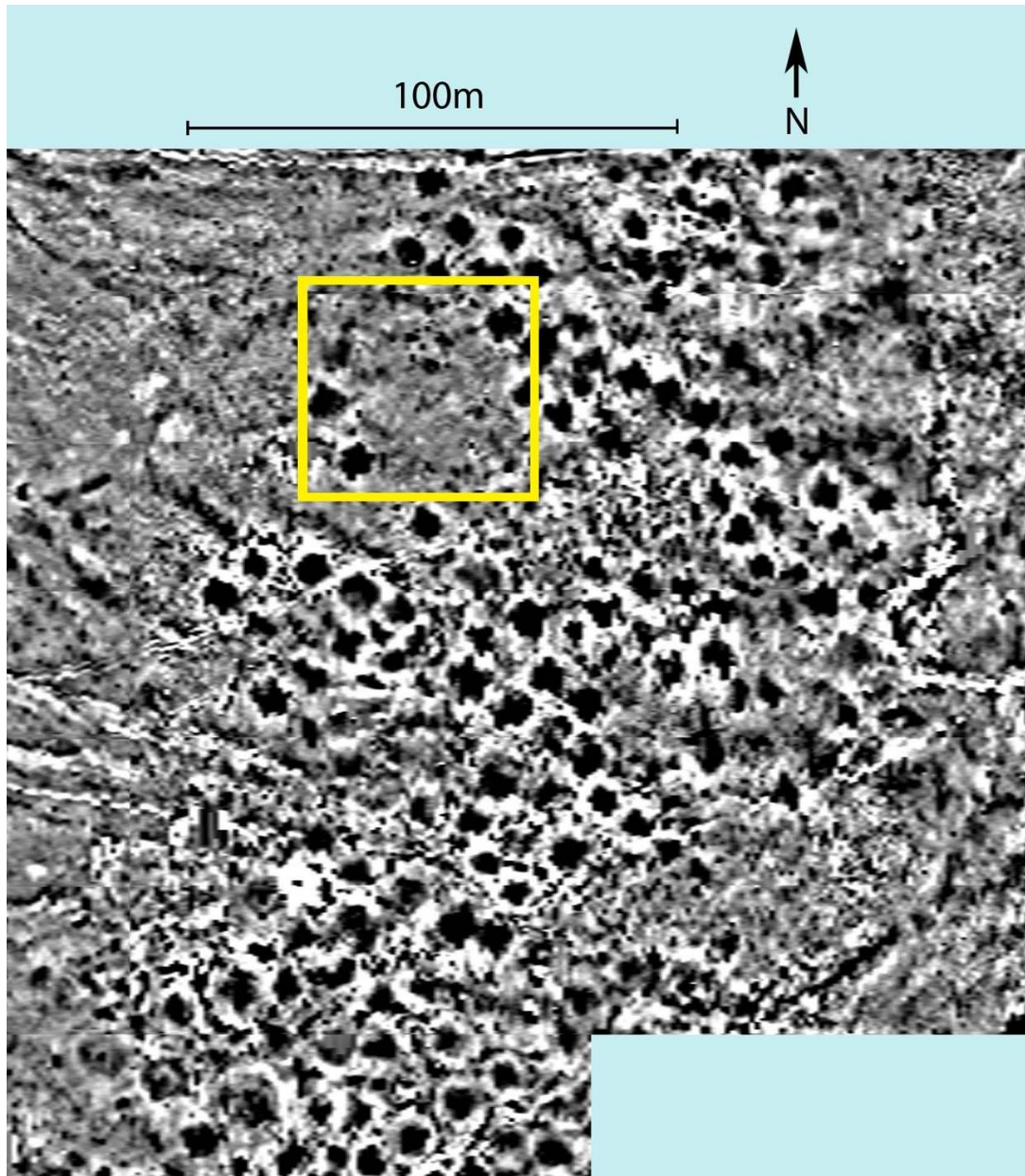
Fig. S6.
The central part of the gradiometer survey of Botai. The area of another potential enclosure highlighted with a box is shown.

Fig. S7.
Morphological indices of the robustness of horse third metacarpals.
(A) Ratio of the Breadth at the middle of the diaphysis over the Greatest length (Bd/GL)
(B) Ratio of the Proximal breadth over the Greatest length (Pb/GL)
(C) Ratio of the Distal supra-articular breadth over the Greatest length (DSb/GL)
(D) Ratio of the Distal articular breadth over the Greatest length (DAb/GL)

Fig. S8.

DNA fragmentation and nucleotide mis-incorporation profiles for sample Haunstetten (CGG_1_017139). Each panel provides DNA fragmentation a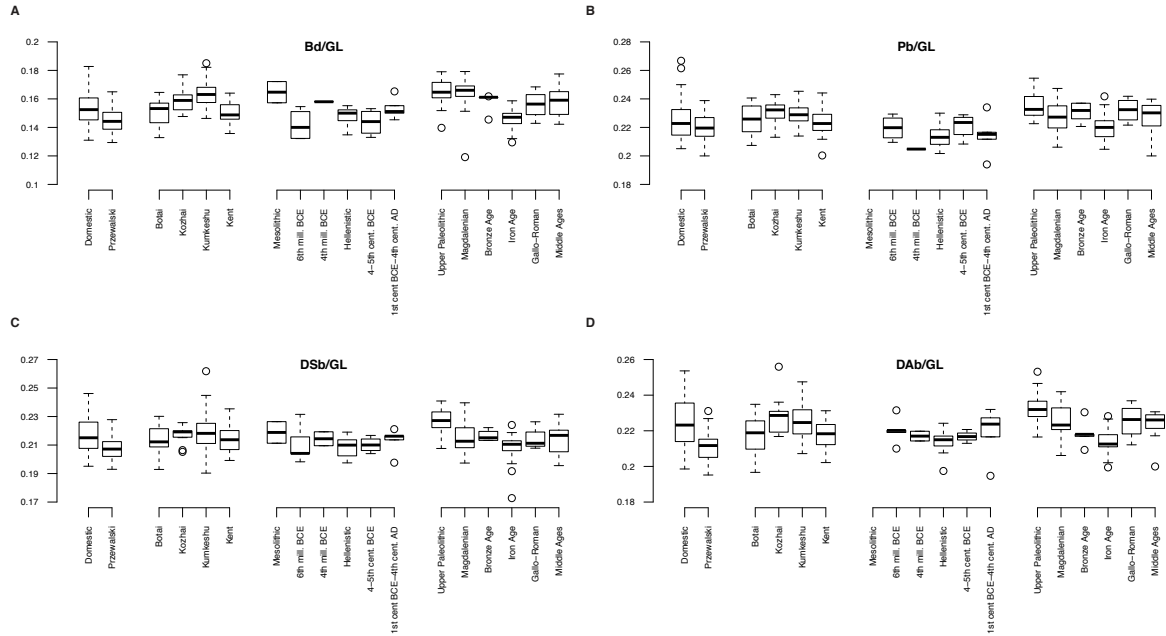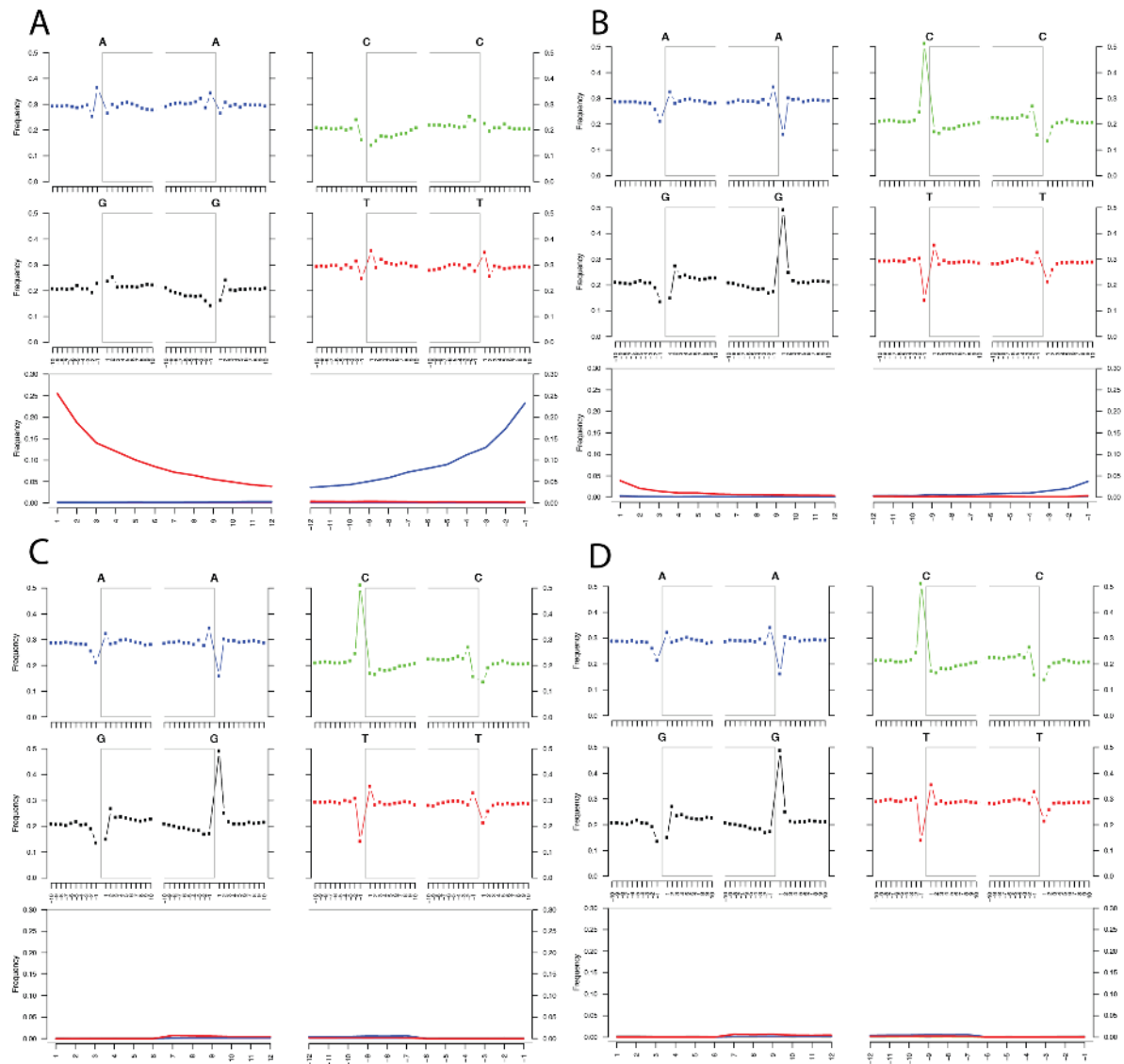nd nucleotide mis-incorporation profiles, as recovered from mapDamage2 *(92)*. DNA fragmentation profiles represent the base composition within the first/last 10 read positions and within the 10 genomic positions preceding and following read termini. Nucleotide mis-incorporation profiles provide C→T (red), G→A (blue) and other (grey) mis-incorporation rates along the first/last 10 read positions. (A) DNA library not USER-treated. The data generated from this library, which shows an excess of nucleotide mis-incorporations, was not included in the final read alignment file used for population genomic analyses. The same holds for the few DNA libraries constructed from raw ancient DNA extracts (ie. not partially USER-treated). (B) DNA library USER-treated but neither rescaled nor trimmed for positions showing elevated mis-incorporation rates. (C) Same library as in B, but following base quality rescaling for positions showing elevated nucleotide mis-incorporation at cytosines that were deaminated post-mortem. (D) Same library as in C, but following trimming for further end positions showing elevated transversion mis-incorporation rates. Here, panels C and D are similar as no trimming was found to be necessary.
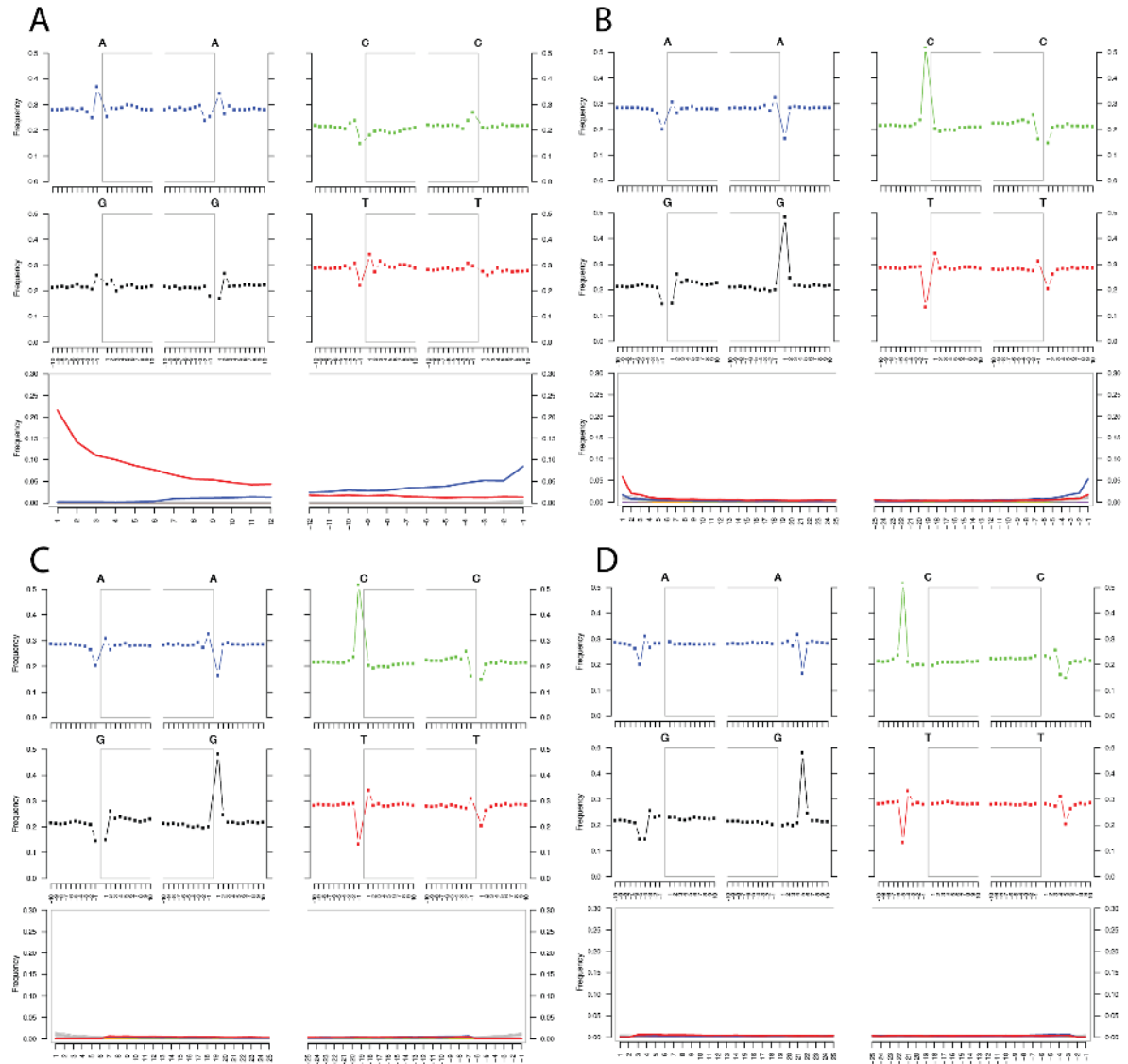
Fig. S9.
DNA fragmentation and nucleotide mis-incorporation profiles for sample NewBotai46 (CGG_1_017038). See Fig. S8 for captions. Here, panels C and D are different due to further trimming for four read positions at both ends. This drastically reduced the impact of sequencing errors on downstream analyses.
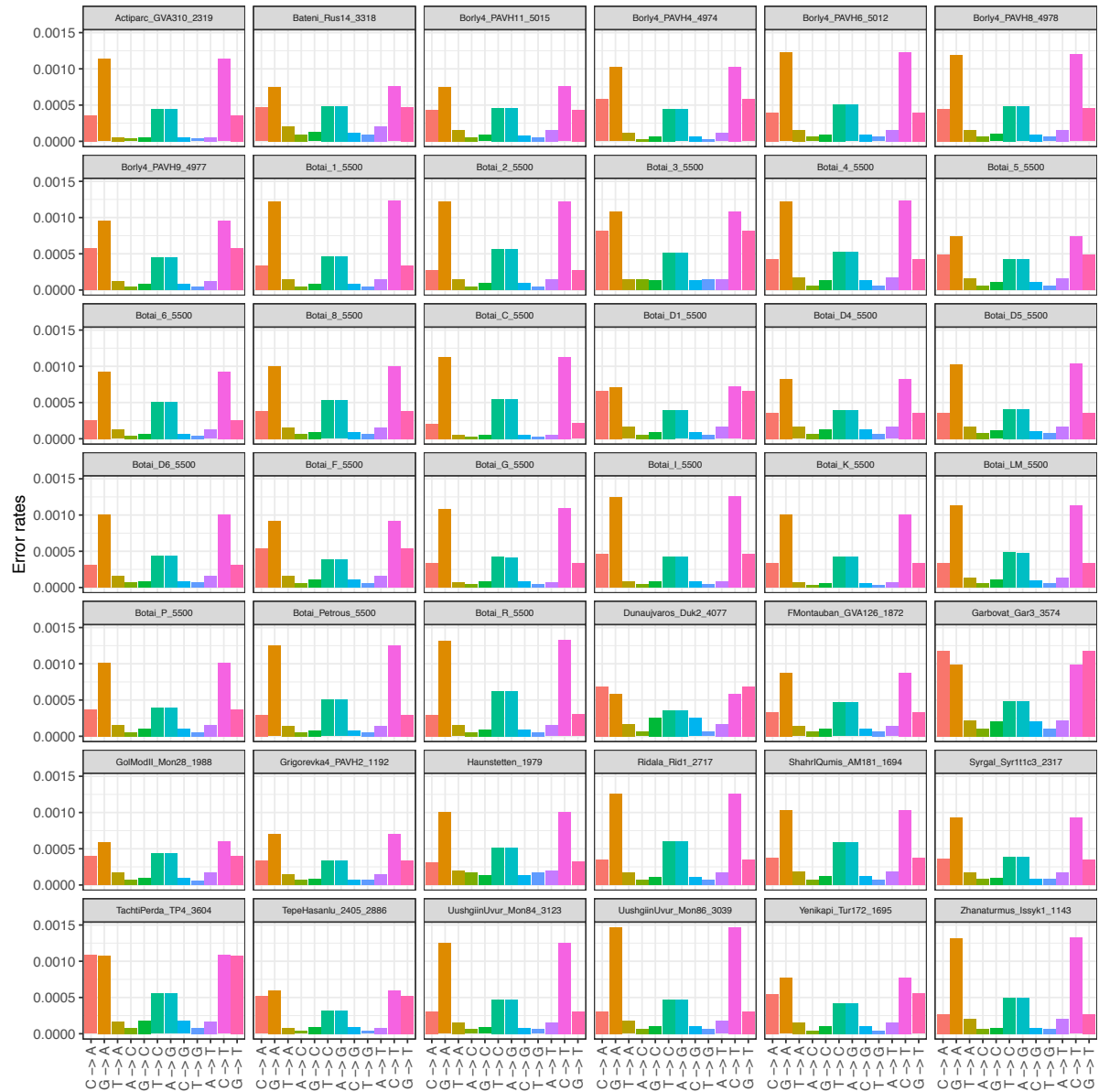
Fig. S10.
Error profiles of the 42 ancient genomes characterized in this study (average error rates per site are provide for each substitution type). Reads showing mapping quality scores strictly inferior to 30 were disregarded. Bases showing quality scores strictly inferior to 20 were disregarded.
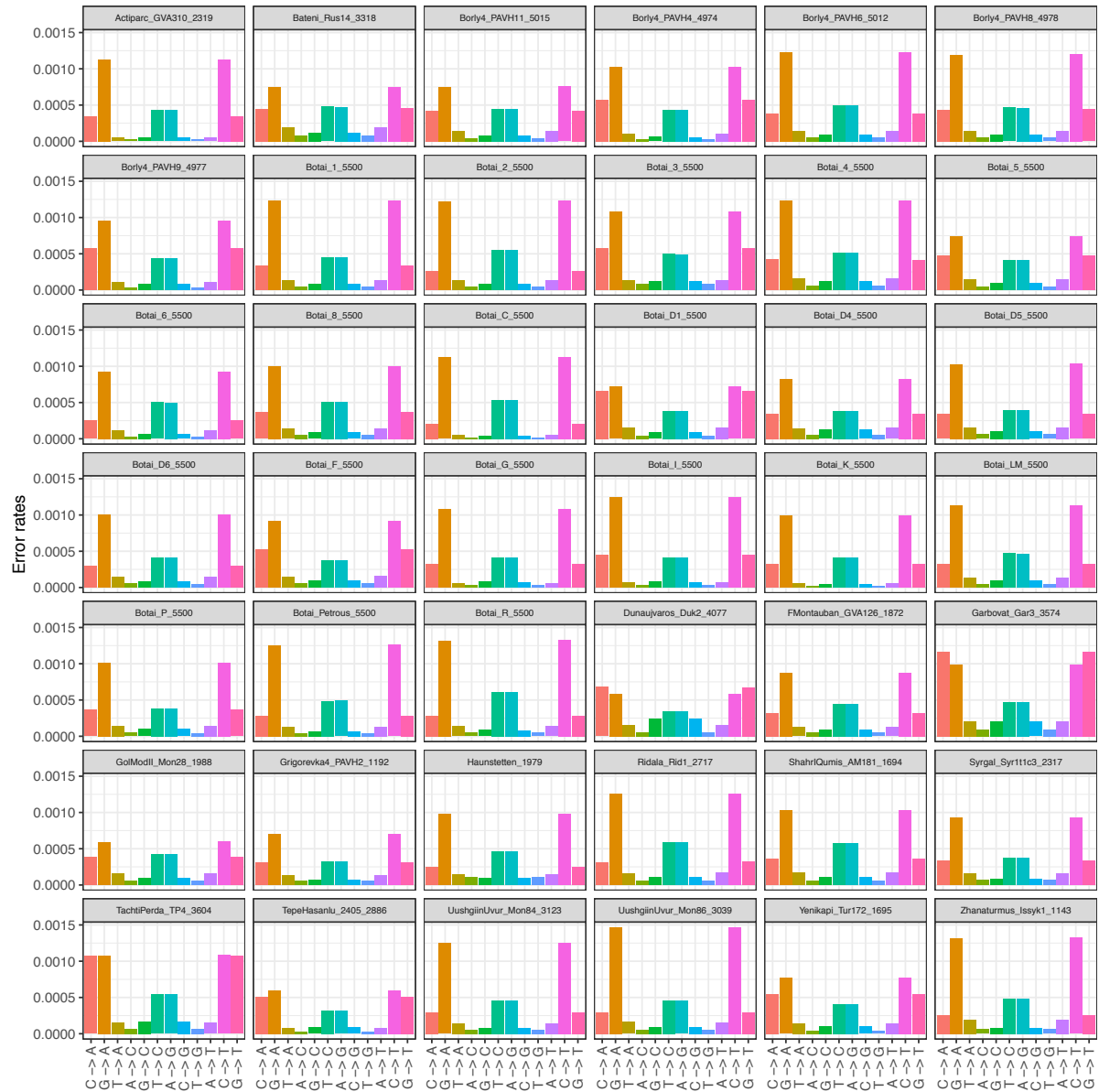
Fig. S11.
Error profiles of the 42 ancient genomes characterized in this study (average error rates per site are provide for each substitution type). Reads showing mapping quality scores strictly inferior to 30 were disregarded. Bases showing quality scores strictly inferior to 30 were disregarded.
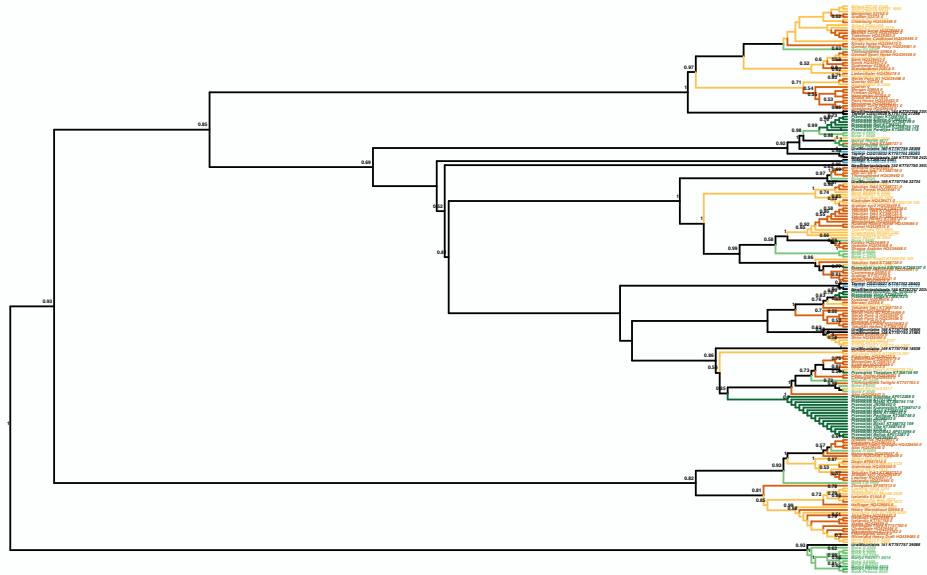
Fig. S12.

Best Maximum Likelihood tree retracing the phylogenetic relationships between 211 mitochondrial genomes. The tree was reconstructed using RAxML, applying independent GTR+GAMMA substitution models to six sequence partitions. Node supports are indicated as fraction of 100 bootstrap pseudo-replicates. The root was manually forced according to the root identified in BEAST phylogenetic analyses. For the specimens analyzed in this study, tip labels are composed of individual sample names, their reference number in the laboratory database as well as their age (years ago, from 2017). Tip labels for previously published sequences report the same information when available, and any other metadata available from the sequence repository.
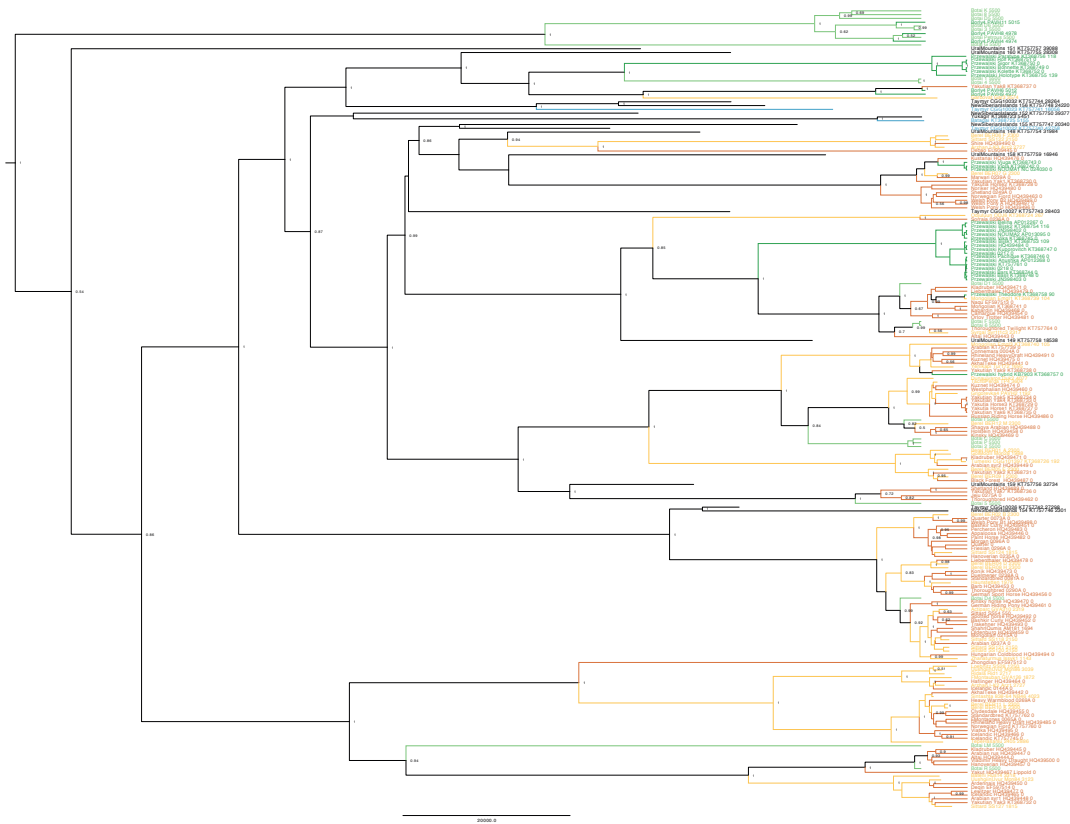
Fig. S13.
Maximum clade credibility tree retracing the phylogenetic relationships between 211 mitochondrial genomes. The tree was reconstructed using a total number of 617 million MCMC states in BEAST, disregarding the first 25% as burn-in, and assuming a LogNormal autocorrelated clock model and a Bayesian Skyline demographic model. The substitution models applied to the six sequence partitions were the HKY+G8+I model (1st codon position, 3,802 sites), the HKY+I model (2nd codon position, 3,799 sites), the TrN+G8 model (3rd codon position, 3,799 sites), the HKY+G8+I model (transfer RNAs, 1,518 sites), the HKY+G8+I model (ribosomal RNAs, 2,555 sites) and the TrN+G8+I model (control region, 961 sites).
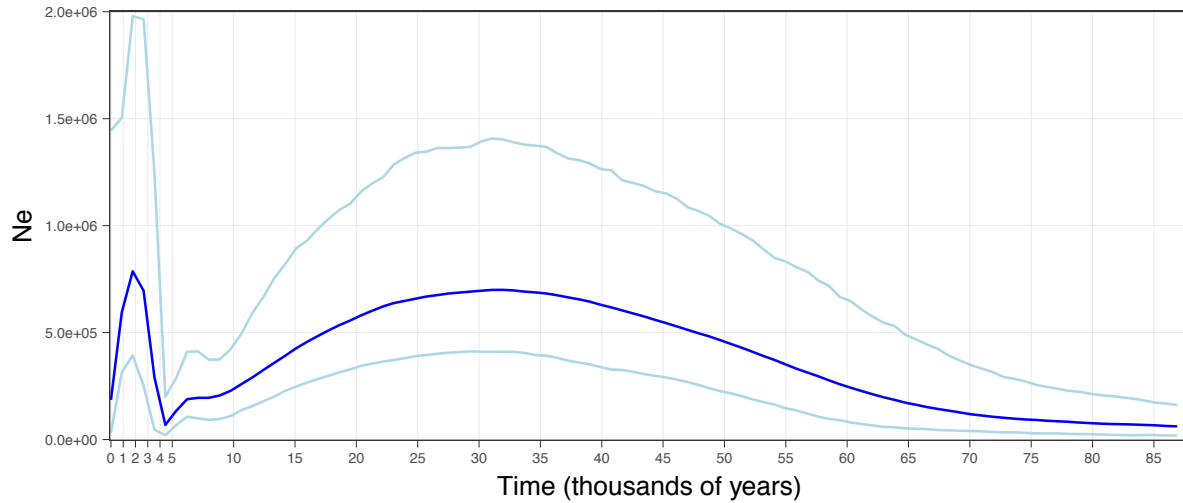
Fig. S14.
Horse Bayesian skyline demographic profile. The demographic trajectory was reconstructed from the sequence variation present in 211 mitochondrial genomes, partitioned according to six gene/sequence categories, assuming a LogNormal autocorrelated clock model and using tip-dating (the age of each individual sequence is reported in the last field of the tip labels, Fig. S13). The final trajectory was obtained from a total number of ~617 million MCMC states in BEAST, disregarding the first 25% as burn-in.
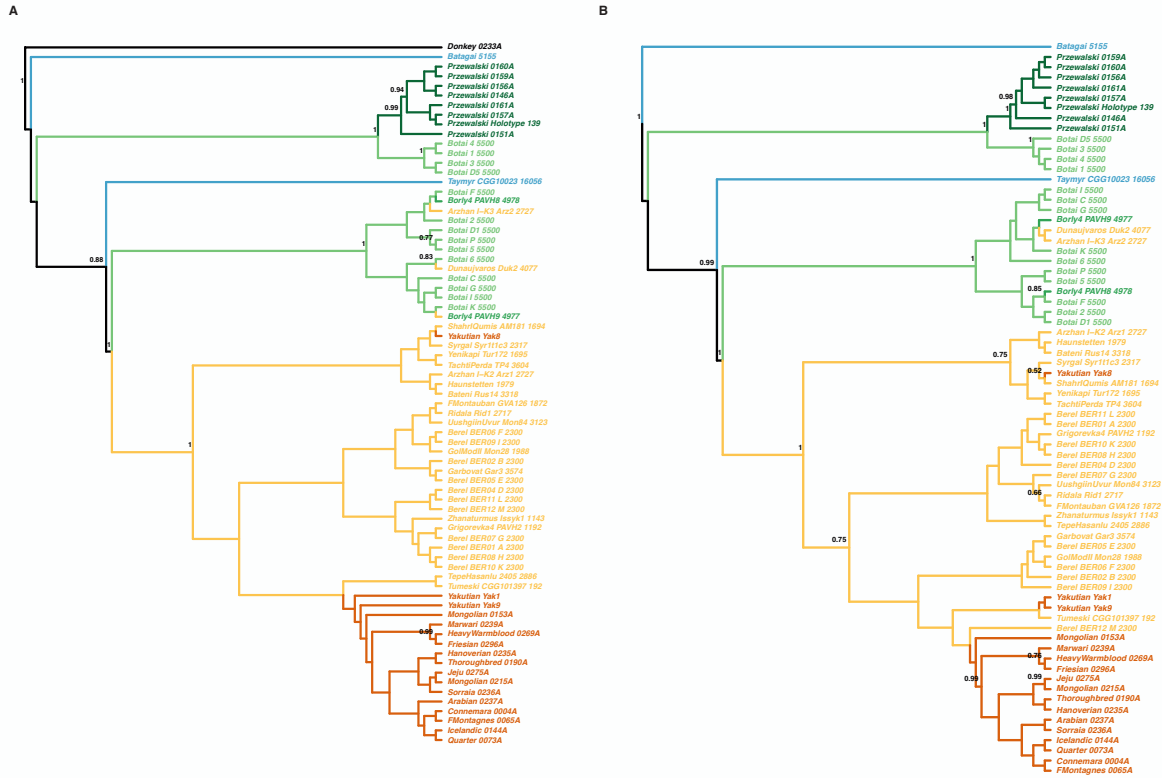
Fig. S15.

Best ML Phylogenetic tree (GTRCAT substitution model) of Y-chromosome haplotypes (minimum coverage = 1). Node supports are calculated from 100 bootstrap pseudo-replicates. Bootstrap supports inferior to 50% are not shown. (A) Tree rooted on the African wild ass, *Equus somalicus*. (B) Tree rooted on the 5 kyrs-old sample ARUS_0223A_Batagai from Yakutia, North East Siberia *(52)*, based on the analysis disregarding the African wild ass Y-chromosome haplotype.
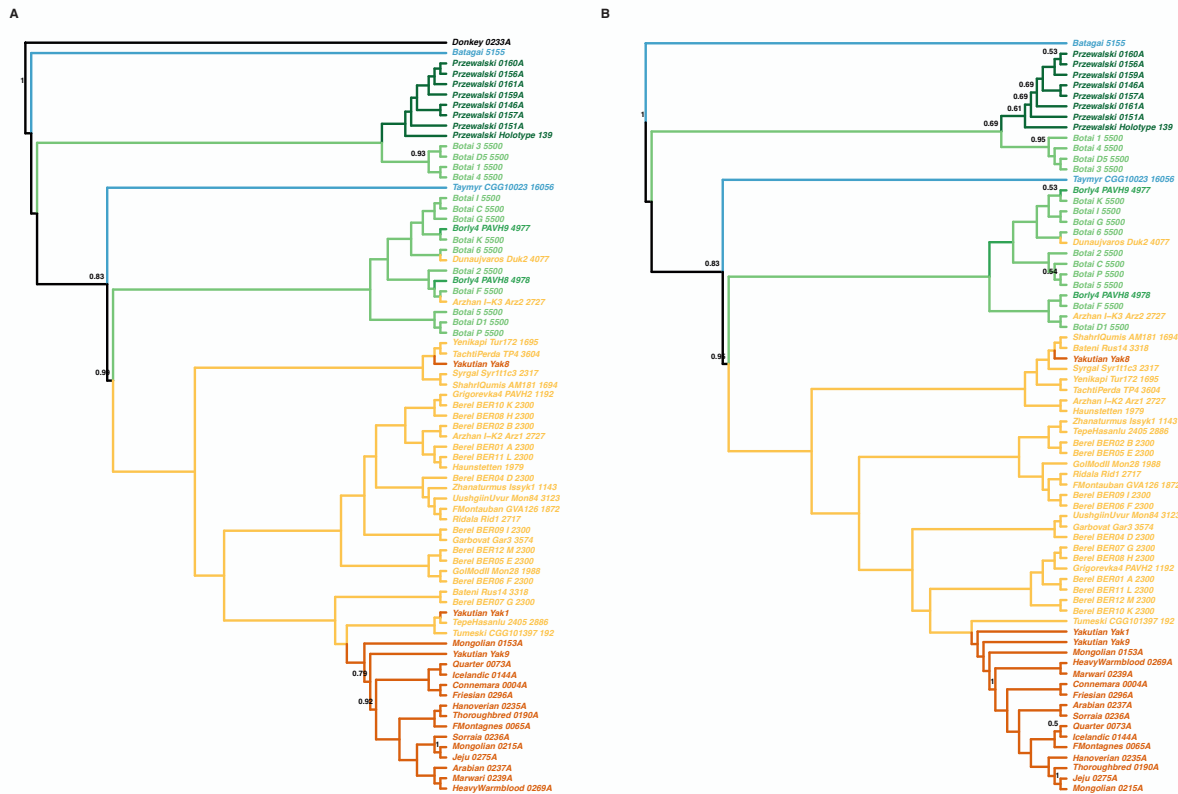
Fig. S16.
Best ML Phylogenetic tree (GTRCAT substitution model) of Y-chromosome haplotypes (minimum coverage = 2). Node supports are calculated from 100 bootstrap pseudo-replicates. Bootstrap supports inferior to 50% are not shown. (A) Tree rooted on the African wild ass, *Equus somalicus*. (B) Tree rooted on the 5 kyrs-old sample ARUS_0223A_Batagai from Yakutia, North East Siberia *(52)*, based on the analysis disregarding the African wild ass Y-chromosome haplotype.
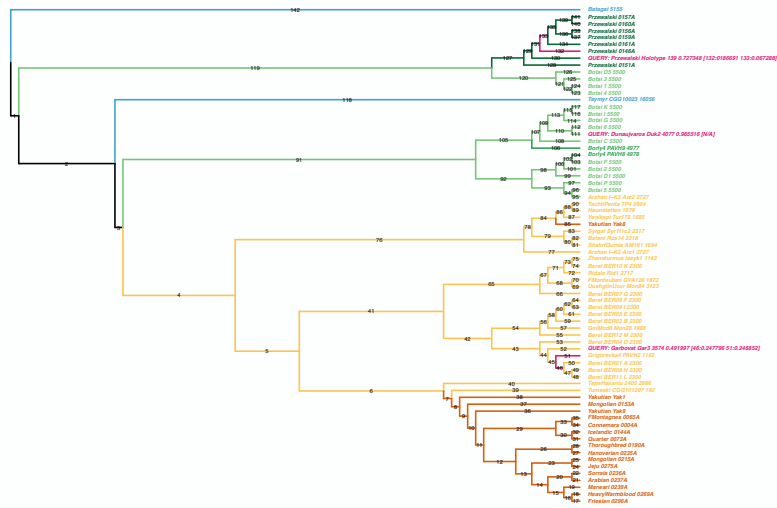
Fig. S17.

Evolutionary Phylogenetic Placement (EPA) of individual stallions showing low coverage data (minimum coverage = 1). Individuals that were placed using the EPA procedure are indicated with a QUERY label. The likelihood weight support for the placement indicated by the tree topology is provided as a frequency after the sample name. Alternative nodes, together with their respective likelihood weight support for their placement, are reported in brackets. Branch labels are reported and those branches supporting an alternative placement of at least one query sequence are colored in pink.
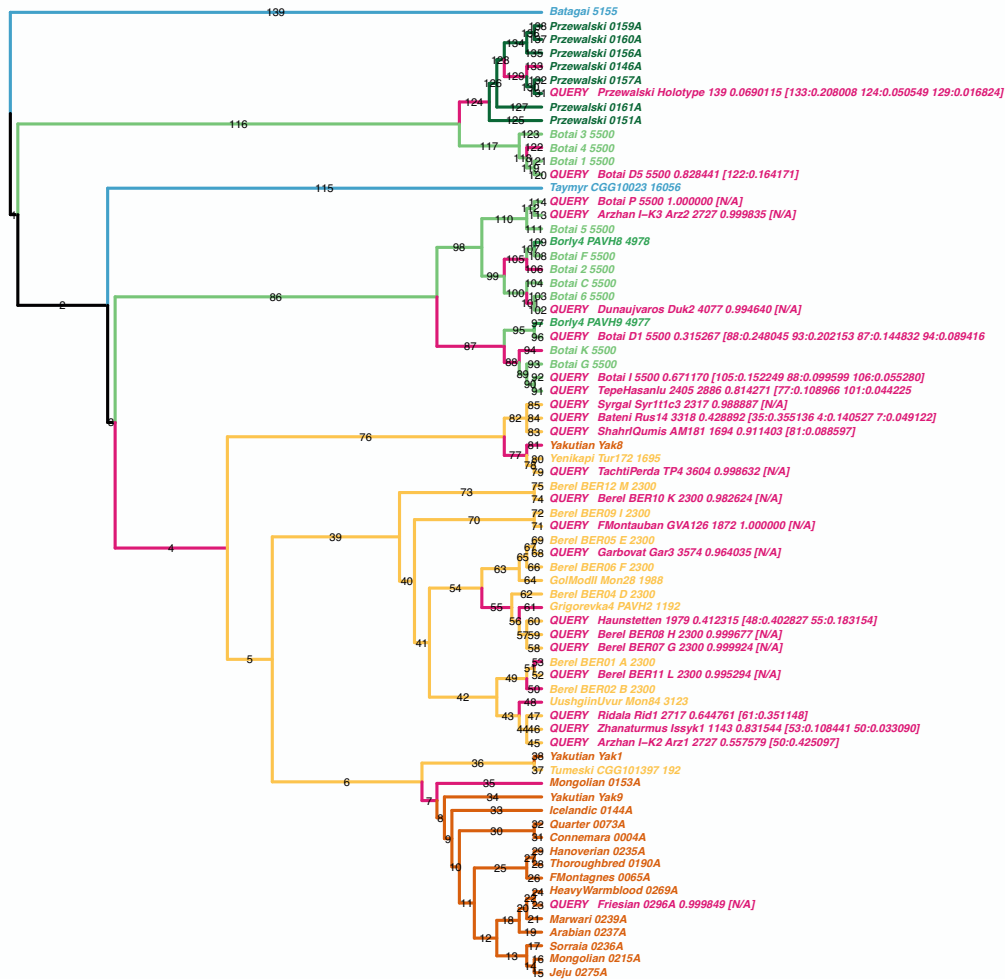
Fig. S18.
Evolutionary Phylogenetic Placement (EPA) of individual stallions showing low coverage data (minimum coverage = 2). See Fig. S17 for captions.
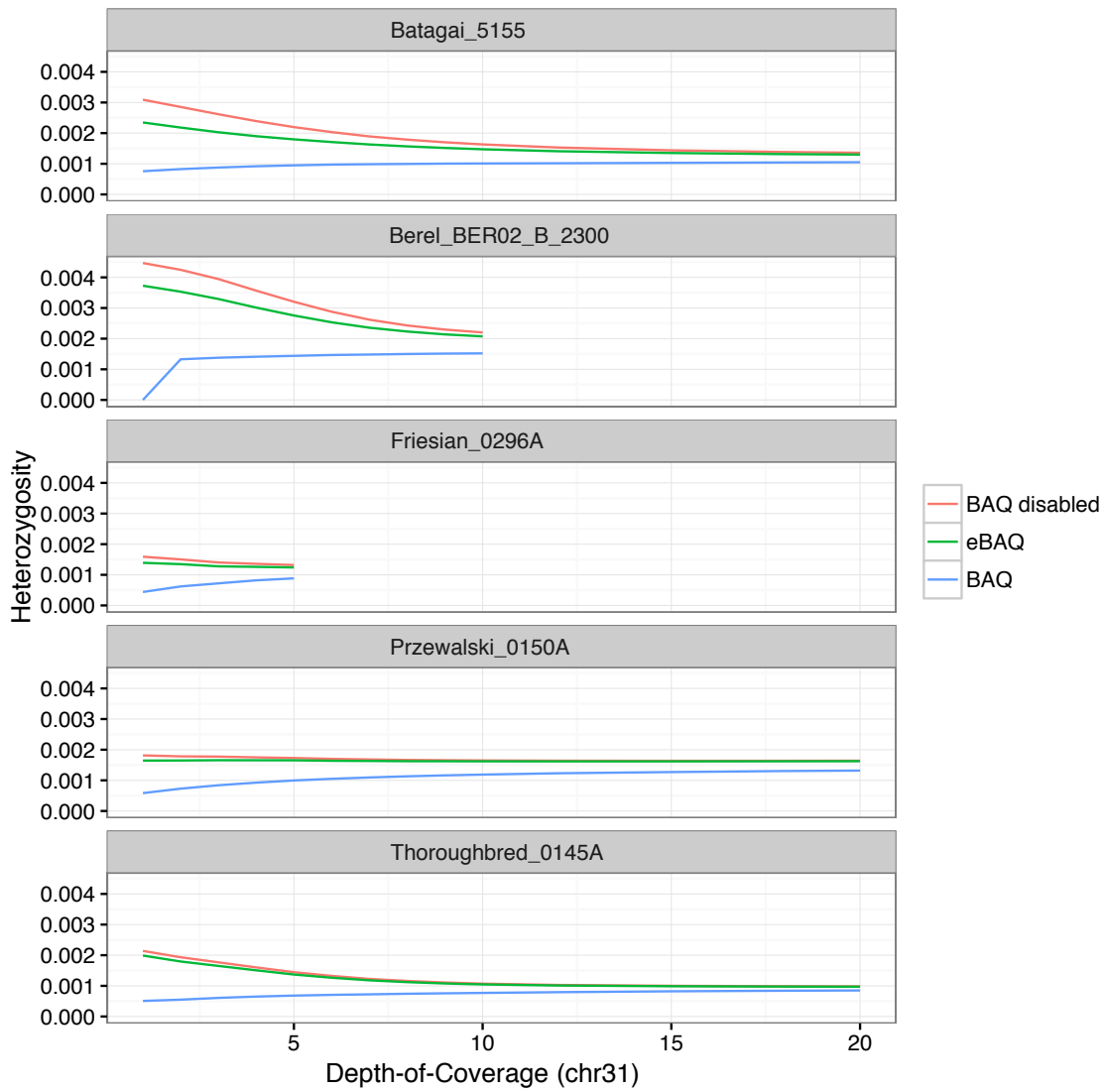
Fig. S19.

Comparison of heterozygosity estimates using Extended base alignment quality (eBAQ), original base alignment quality (BAQ), and with BAQ disabled for three modern and two ancient horses, down-sampled from their original depth-of-coverage to 1.0X.
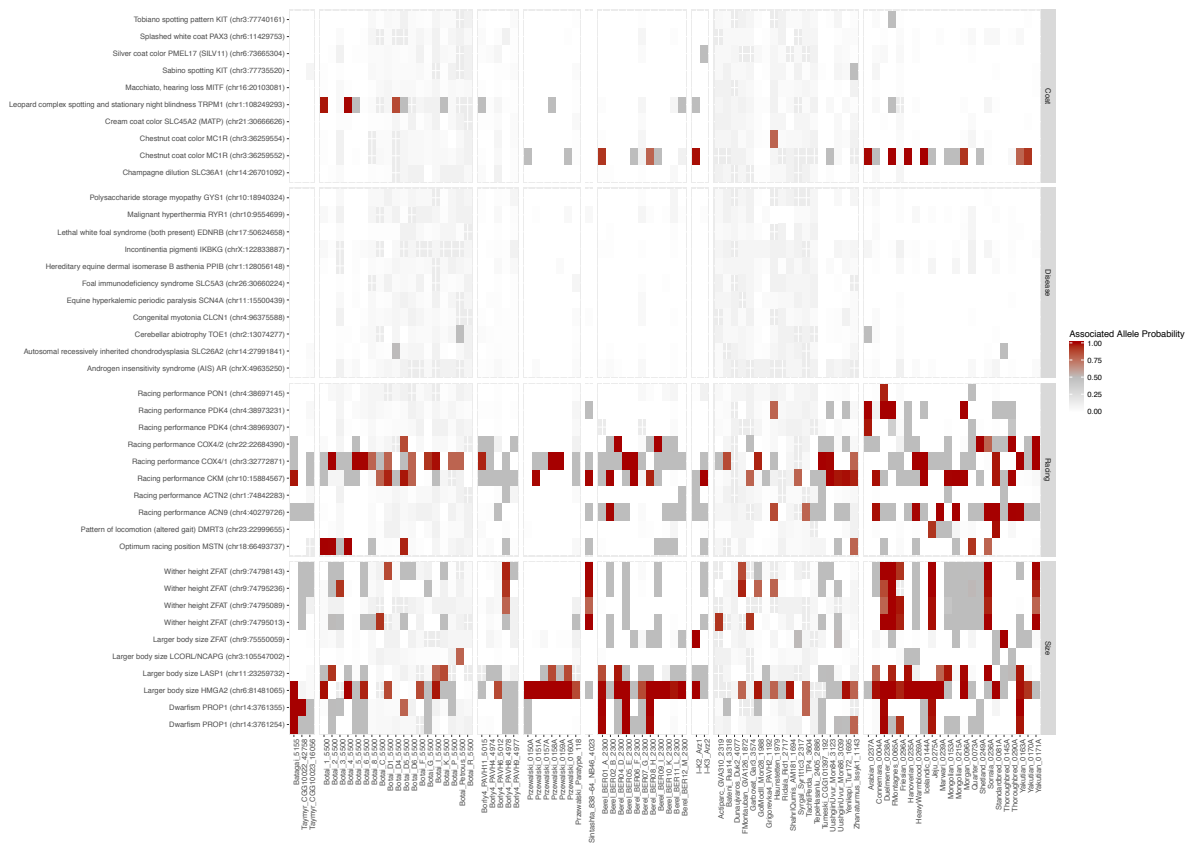
Fig. S20.

Heatmap where the gradient color represents the probability that a horse (x-axis) carries the allele causative of/associated with one of the 41 phenotypic traits (y-axis). Absence of the allele is represented in white, and its presence as homozygous is shown in red. Intermediate grey colors symbolize heterozygous sites. Phenotypes supported by a single read are conservatively downscaled to white. Grey crosses indicate non-genotyped positions.
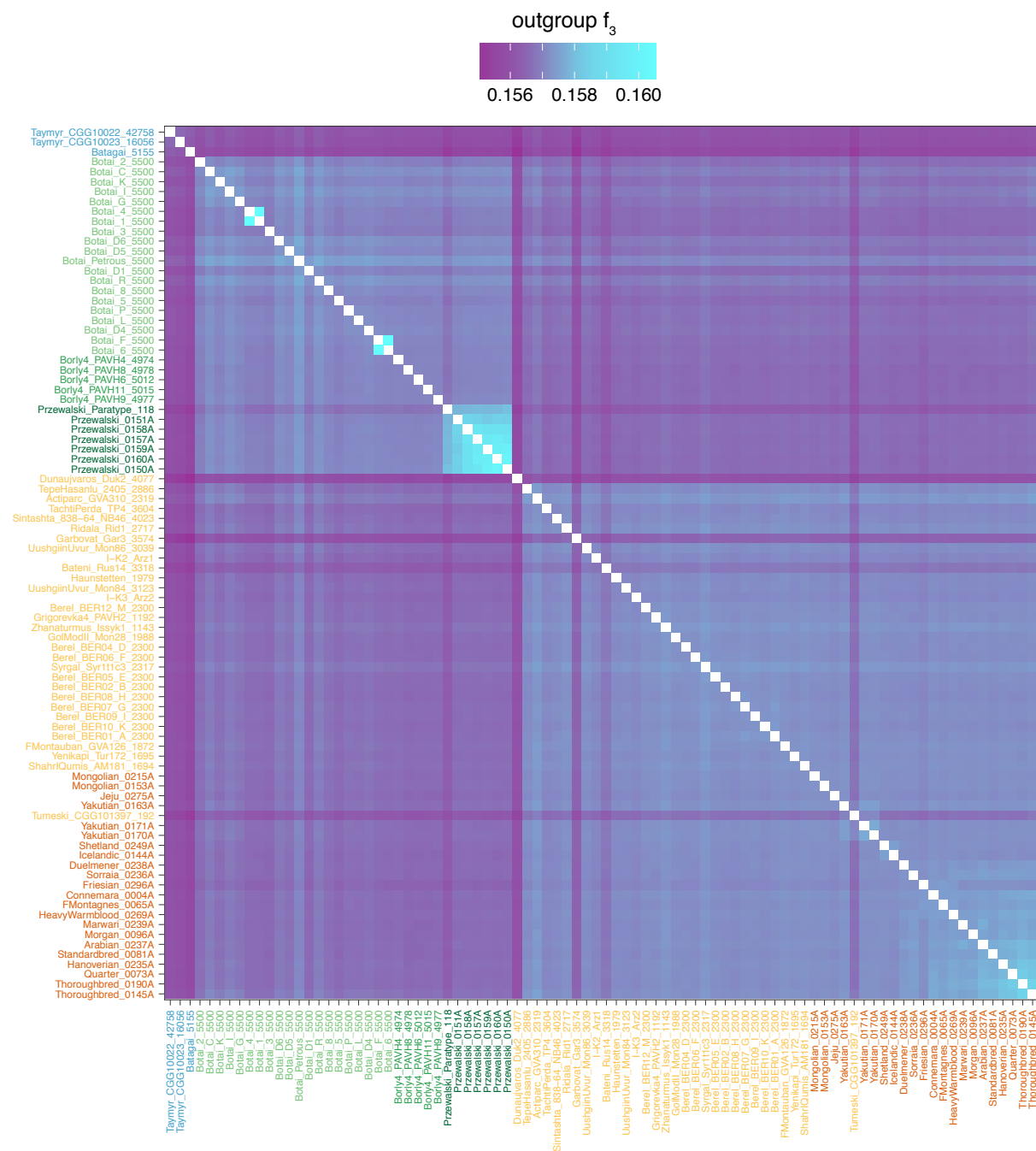
Fig. S21

f_3 statistics in the form (X, Y; Outgroup), where X and Y represents all possible pairwise combinations (n=3,828) of horses analyzed in this study.
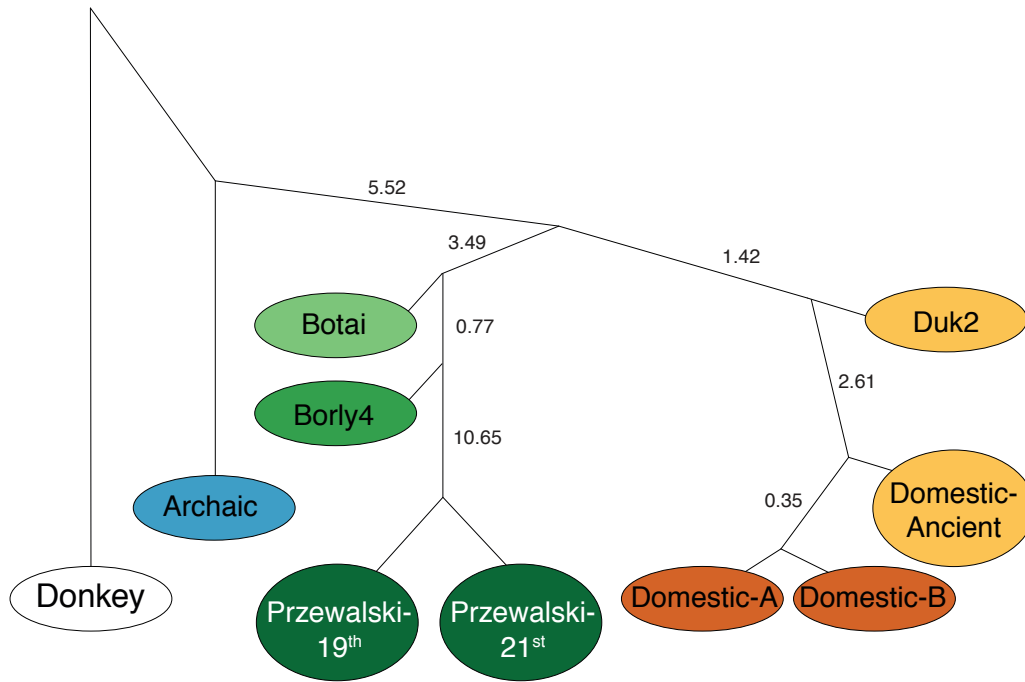
Fig. S22.

Admixture graph best-fitting the 630 permutations of the f_4-statistics. Numbers on edges are in f_2 drift units ($\times 10^{-4}$). Modern domesticates are denoted as Domestic-A and Domestic-B, where Domestic-A includes Yakutian_0163A, Yakutian_0170A, Yakutian_0171A, Mongolian_0153A, Mongolian_0215A, Jeju_0275A and Tumeski_CGG101397_192, while Domestic-B horses consist of all remaining samples belonging to DOM2, except Dunaujvaros_Duk2_4077.
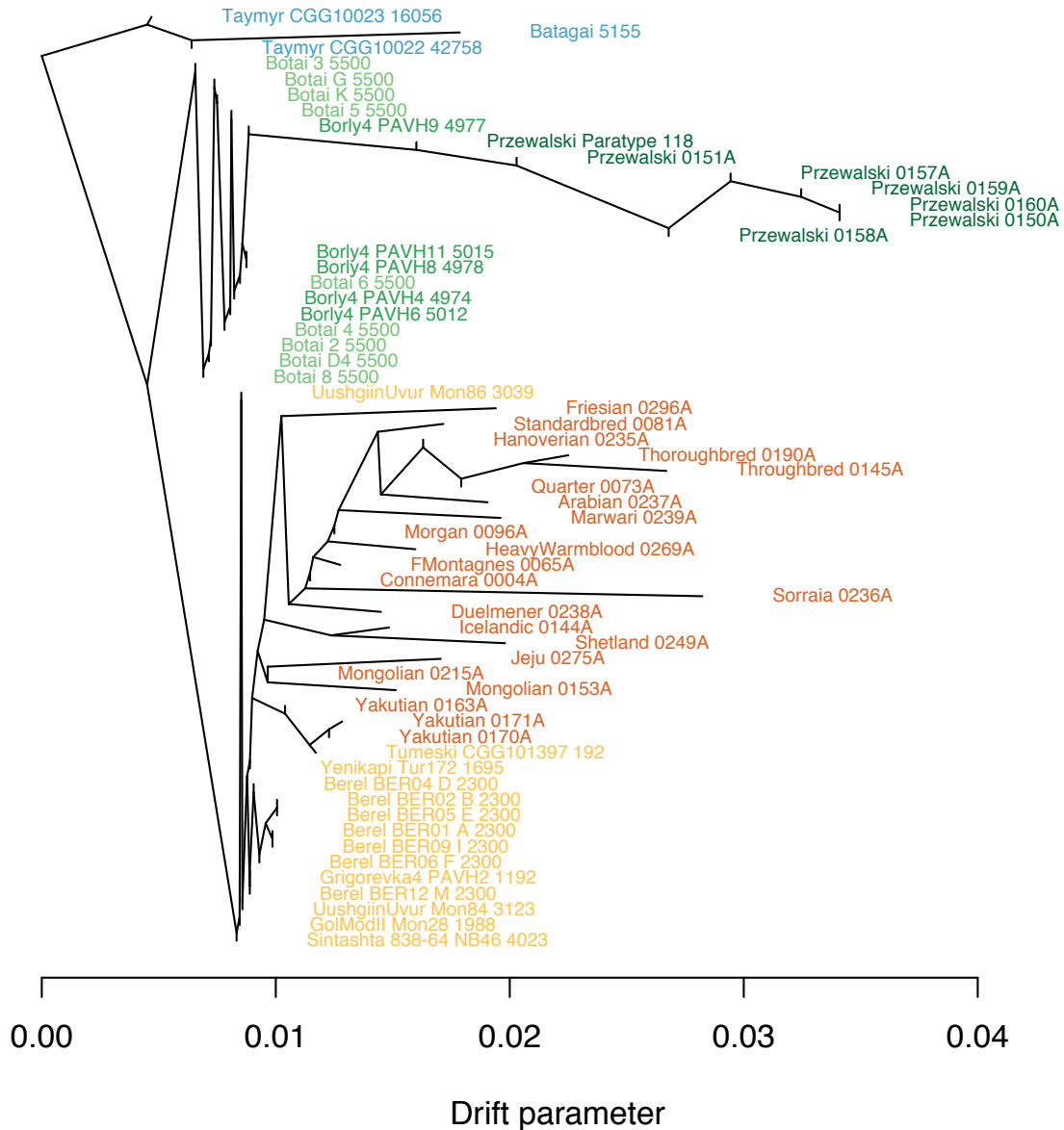
Fig. S23.

TreeMix tree of 60 horse genomes with an average depth-of-coverage greater or equal to 3.0X, representing the major horse populations. For each of the 60 samples, genotypes were randomly sampled according to their posterior probabilities, and subsequently coded in TreeMix input format. The variance explained by adding up to 5 migration edges can be found in Table S5.
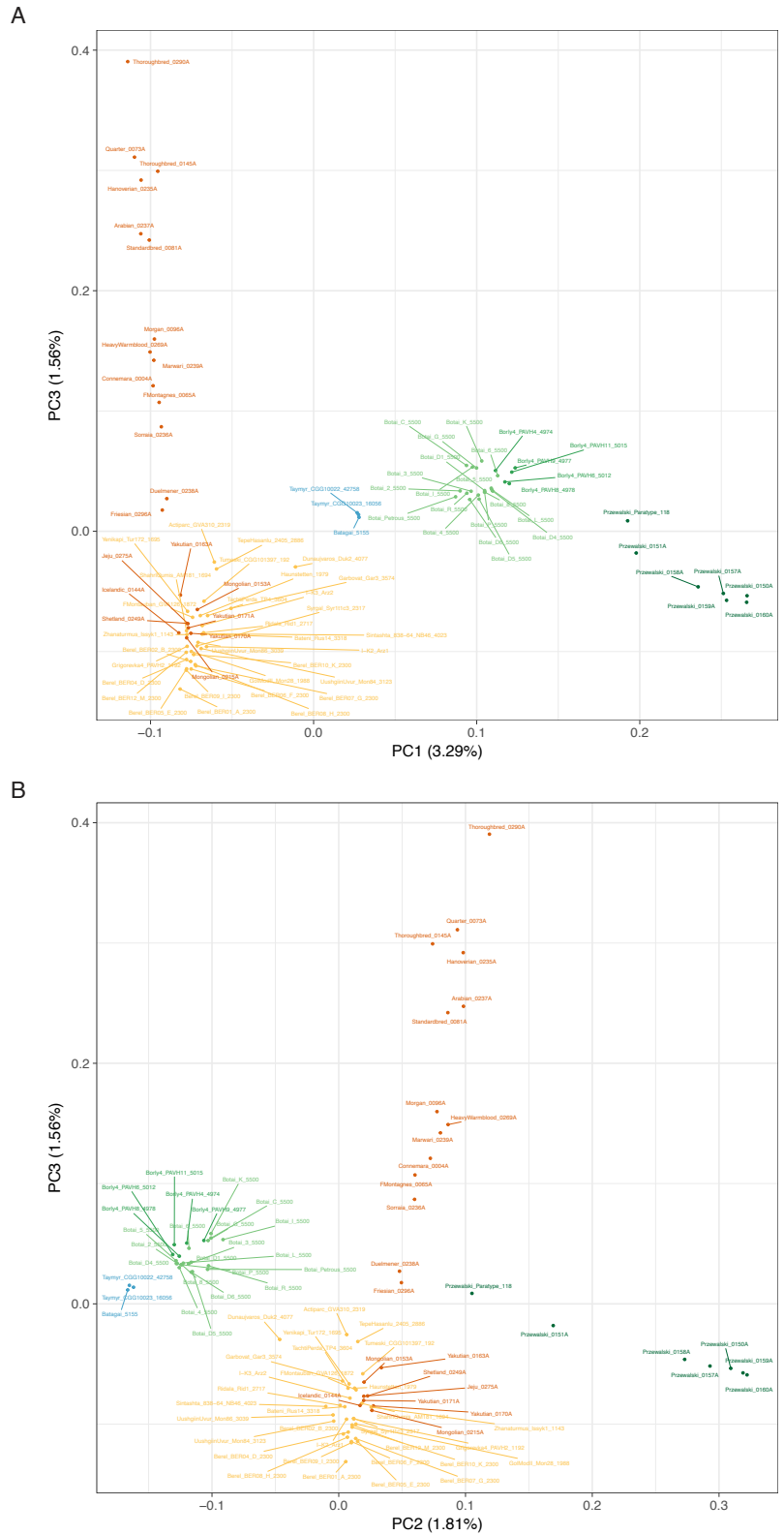
Fig. S24.

Principal Component Analysis (PCA) of the sequence variation present in 88 ancient and modern horse genomes. (A) First two principal components. (B) First and third principal components.

Fig. S25.

D-statistics in the form of (((X, Przewalski), Botai/Borly4), donkey). The tree topology tested is indicated on each individual panel, with reference to the individual genomes considered for populations H_2 and H_3 (in $(((H_1, H_2), H_3),\text{Outgroup})$ topologies). X refers to ancient and modern horses of the second domestic clade (DOM2). The names of the specimens considered for X are reported on the Y-axis. The Przewalski's horse genomes used as population H_2 are (from left to right): Przewalski_0157A, Przewalski_0160A and Przewalski_Paratype_118. The Botai/Borly4 horse genomes used as population H_3 are (from top to bottom): Botai_2_5500, Botai_4_5500, Botai_5_5500, Botai6_5500, Borly4_PAVH8_4978 and Borly4_PAVH9_4977. Analyses have been carried out disregarding transitions. Red dots depict significant tests, with adjusted p-values ≤ 0.05 after correction for multiple testing.

Fig. S26.
D-statistics in the form of (((Borly4, Przewalski), Botai), donkey). The tree topology tested is indicated on each individual panel, with reference to the individual genomes considered for populations H_2 and H_3 (in $(((H_1, H_2), H_3),$Outgroup) topologies). The names of the Borly4 horse genomes considered as part of population H_1 are reported on the Y-axis. The Przewalski's horse genomes used as population H_2 are (from left to right): Przewalski_0157A, Przewalski_0160A and Przewalski_Paratype_118. The Botai horse genomes used as population H_3 are (from top to bottom): Botai_2_5500, Botai_4_5500, Botai_5_5500 and Botai6_5500. Analyses have been carried out disregarding transitions. Black dots depict non-significant tests, with adjusted p-values > 0.05 after correction for multiple testing.

Fig. S27.
D-statistics in the form of (((X, Dunaujvaros_Duk2_4077), Botai/Borly4), donkey). The
tree topology tested is indicated on each individual panel, with reference to the individual
genomes considered for populations H_2 and H_3 (in $(((H_1, H_2), H_3),$Outgroup) topologies). X
refers to modern domestic horses. The names of the specimens X are reported on the Y-axis.
Dunaujvaros_Duk2_4077 is used as population H_2. The Botai/Borly4 horse genomes used as
population H_3 are (from top to bottom): Botai_2_5500, Botai_4_5500, Botai_5_5500,
Botai_6_5500, Borly4_PAVH8_4978 and Borly4_PAVH9_4977. Analyses have been carried
out disregarding transitions. Red dots depict significant tests, with adjusted p-values ≤ 0.05
after correction for multiple testing.

D−statistics

Fig. S28.

D-statistics in the form of (((X, Y), Botai/Borly4), donkey). The tree topology tested is indicated on each individual panel, with reference to the individual genomes considered for populations H_2 and H_3 (in $(((H_1, H_2), H_3), Outgroup)$ topologies). X and Y refer to ancient and modern horses of the second domestic clade (DOM2), respectively. The names of the specimens X are reported on the Y-axis. The horse genomes of subset Y used as population H_2

are: Arabian_0237A, Connemara_0004A, Duelmener_0238A, FMontagnes_0065A, HeavyWarmblood_0269A, Icelandic_0144A, Mongolian_0215A, Quarter_0073A, Standardbred_0081A and Thoroughbred_0290A. The Botai/Borly4 horse genomes used as population H_3 are: Botai_2_5500, Botai_4_5500, Botai_5_5500, Botai6_5500, Borly4_PAVH8_4978 and Borly4_PAVH9_4977. Analyses have been carried out disregarding transitions. Red dots depict significant tests, with adjusted p-values ≤ 0.05 after correction for multiple testing.
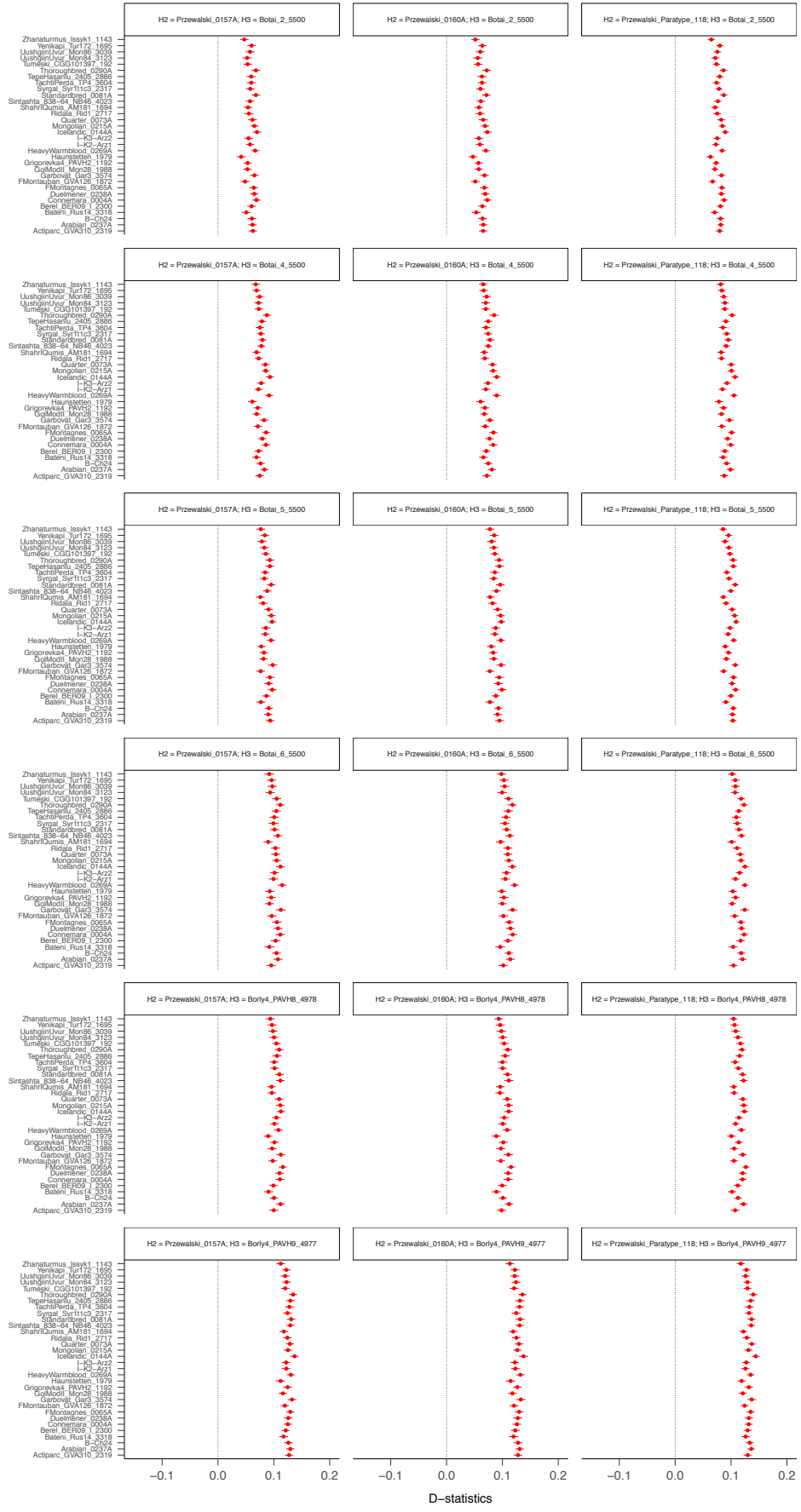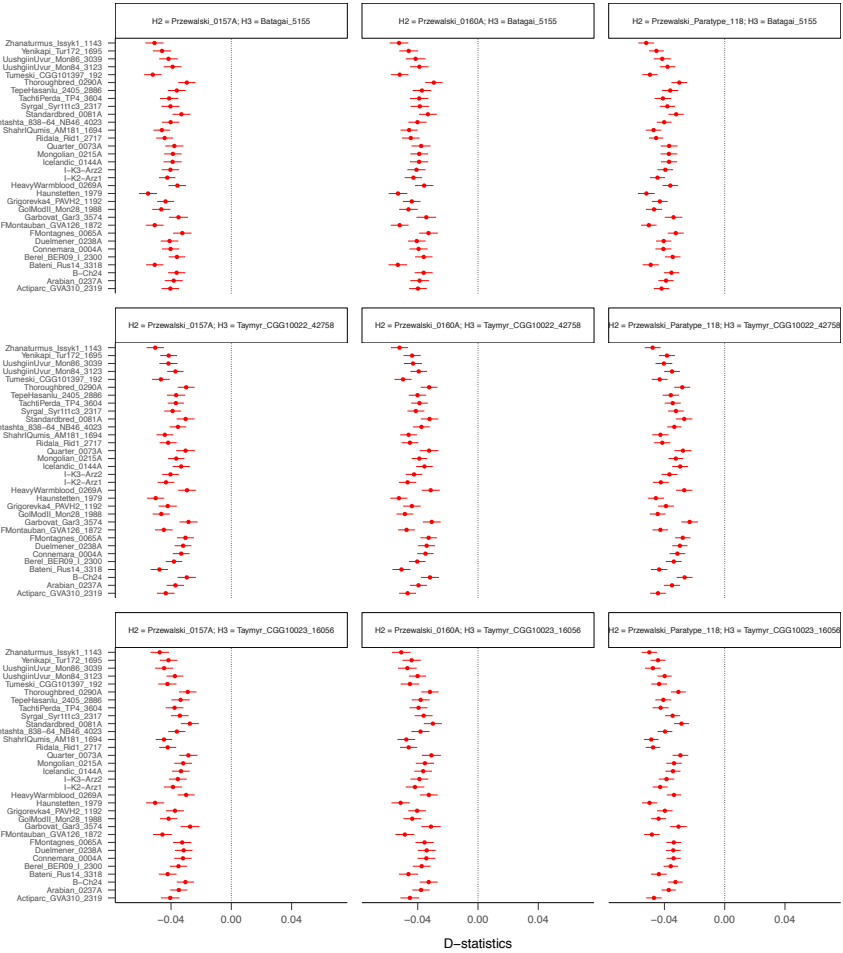
Fig. S29.
D-statistics in the form of (((X, Przewalski), Archaic), donkey). The tree topology tested is indicated on each individual panel, with reference to the individual genomes considered for populations H_2 and H_3 (in $(((H_1, H_2), H_3),$ Outgroup) topologies). X refers to ancient and modern horses of the second domestic clade. (DOM2) The names of the specimens X are reported on the Y-axis. The Przewalski's horse genomes used as population H_2 are (from left to right): Przewalski_0157A, Przewalski_0160A and Przewalski_Paratype_118. The genomes representative of the now-extinct archaic horse lineage, used as population H_3 are (from top to bottom): Batagai_5155, Taymyr_CGG10022_42758 and Taymyr_CGG10023_16056. Analyses have been carried out disregarding transitions. Red dots depict significant tests, with adjusted p-values ≤ 0.05 after correction for multiple testing.
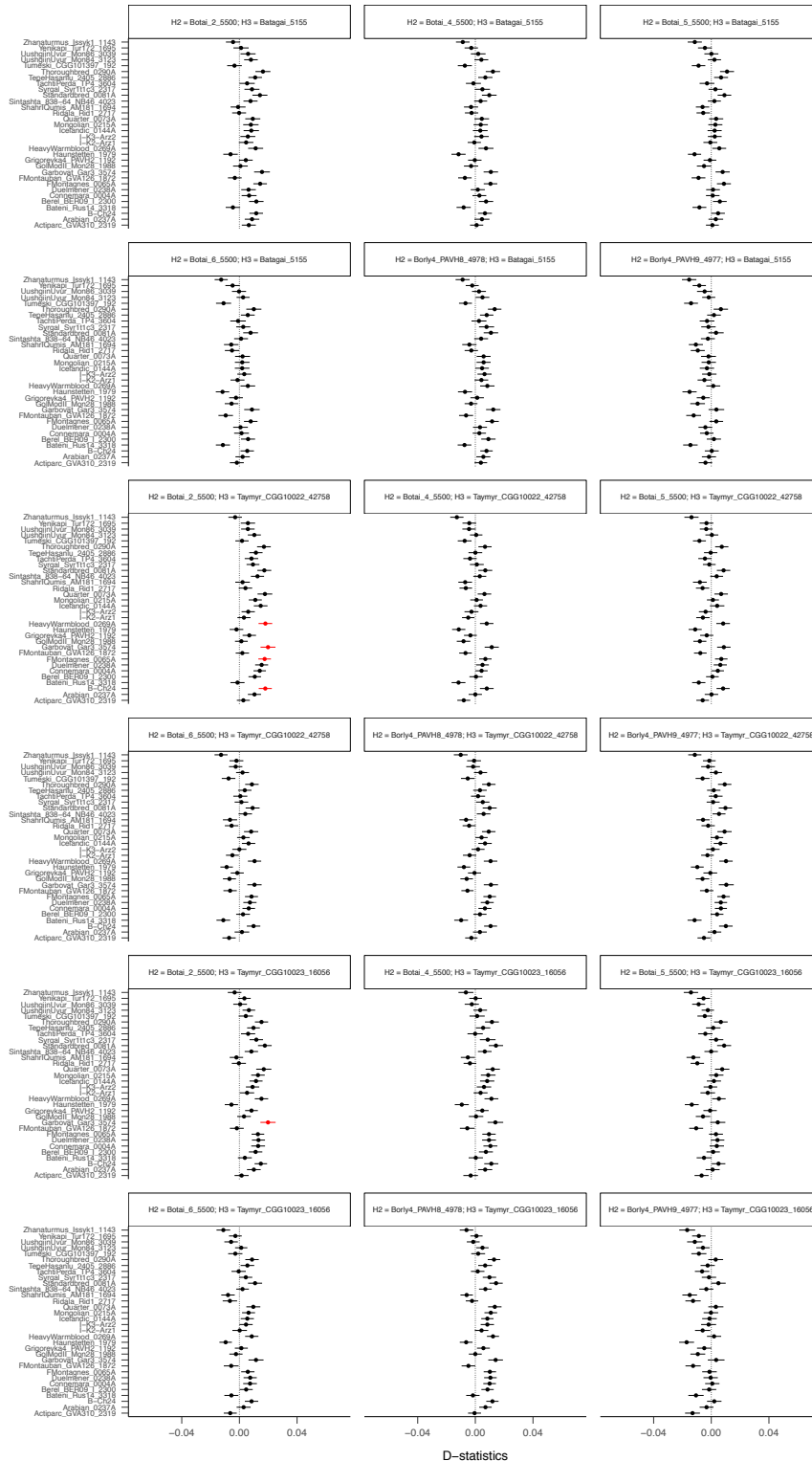
Fig. S30.

D-statistics in the form of (((X, Botai/Borly4), Archaic), donkey). The tree topology tested is indicated on each individual panel, with reference to the individual genomes considered for populations H_2 and H_3 (in ((((H_1, H_2), H_3),Outgroup) topologies). X refers to ancient and modern horses of the second domestic clade (DOM2). The names of the specimens X are reported on the Y-axis. The Botai/Borly4 horse genomes used as population H_2 are: Botai_2_5500,

Botai_4_5500,   Botai_5_5500,   Botai6_5500,   Borly4_PAVH8_4978   and Borly4_PAVH9_4977. The genomes representative of the now-extinct archaic horse lineage, used   as   population H_3 are: Batagai_5155,   Taymyr_CGG10022_42758   and Taymyr_CGG10023_16056. Analyses have been carried out disregarding transitions. Red dots depict significant tests, with adjusted p-values ≤ 0.05 after correction for multiple testing.
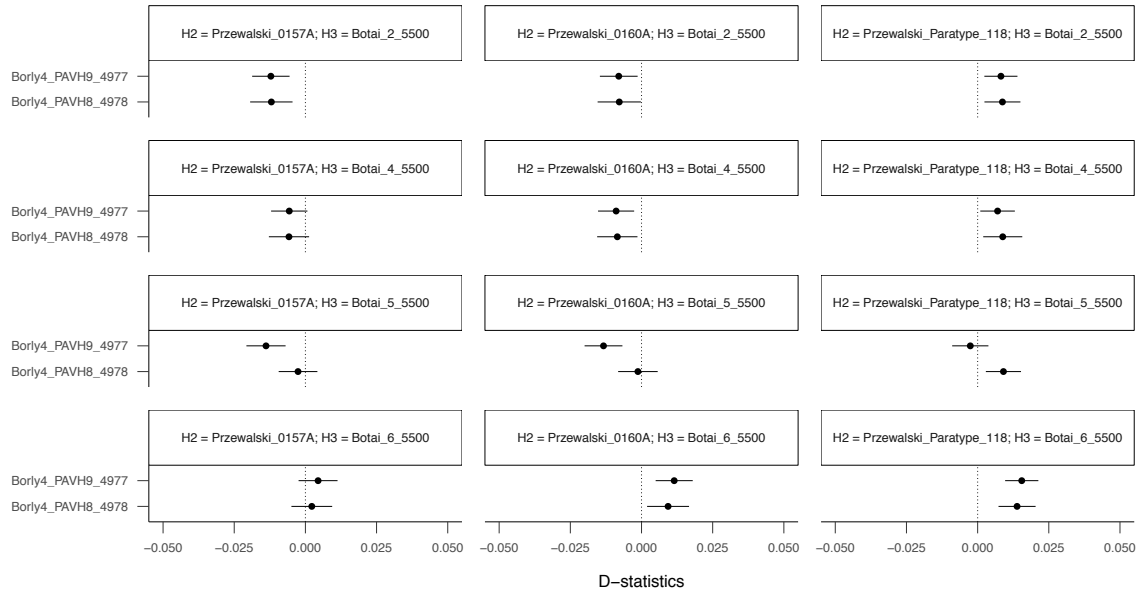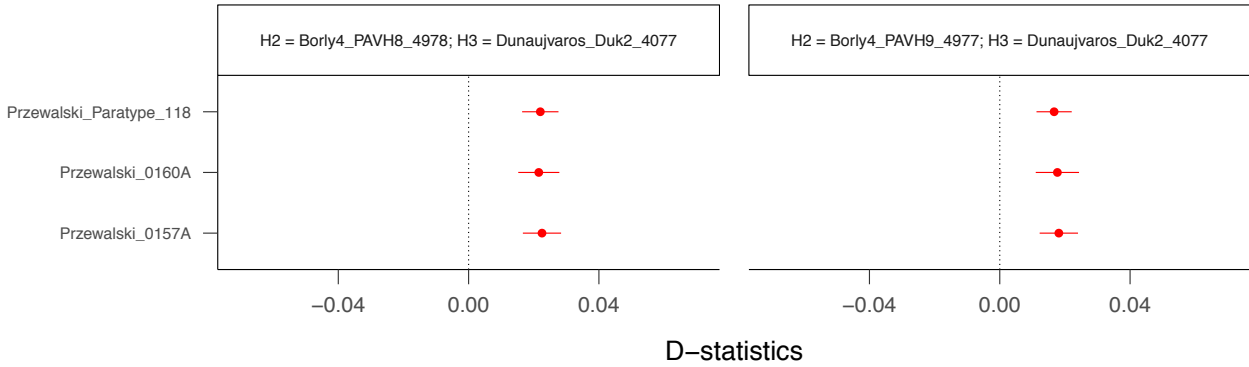
Fig. S31.
D-statistics in the form of (((Przewalski, Borly4), Dunaujvaros_Duk2_4077), donkey). The tree topology tested is indicated on each individual panel, with reference to the individual genomes considered for populations H_2 and H_3 (in (((H_1, H_2), H_3),Outgroup) topologies). The names of the Przewalski's horse genomes considered as part of population H_1 are reported on the Y-axis. The Borly4 horse genomes used as population H_2 are PAVH8_4978 (left) and Borly4_PAVH9_4977 (right). Dunaujvaros_Duk2_4077 is used as population H_3. Analyses have been carried out disregarding transitions. Red dots depict significant tests, with adjusted p-values ≤ 0.05 after correction for multiple testing.

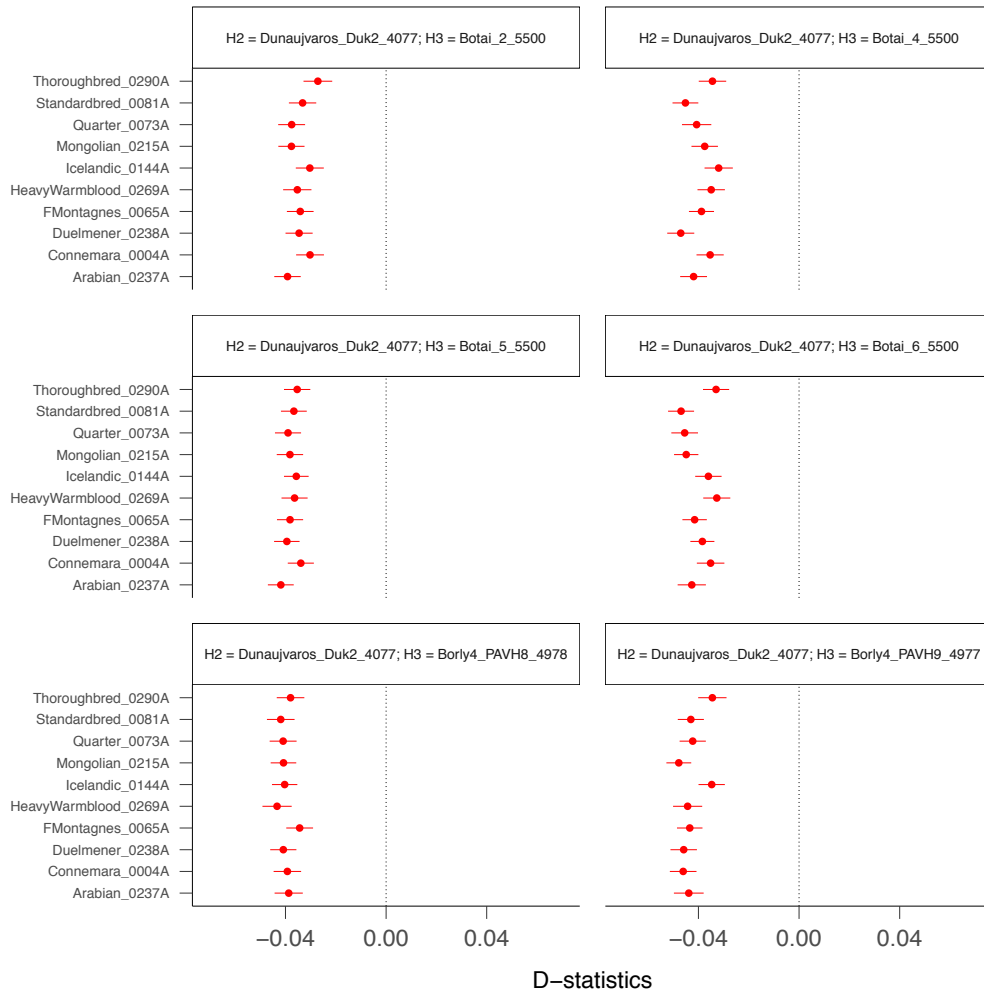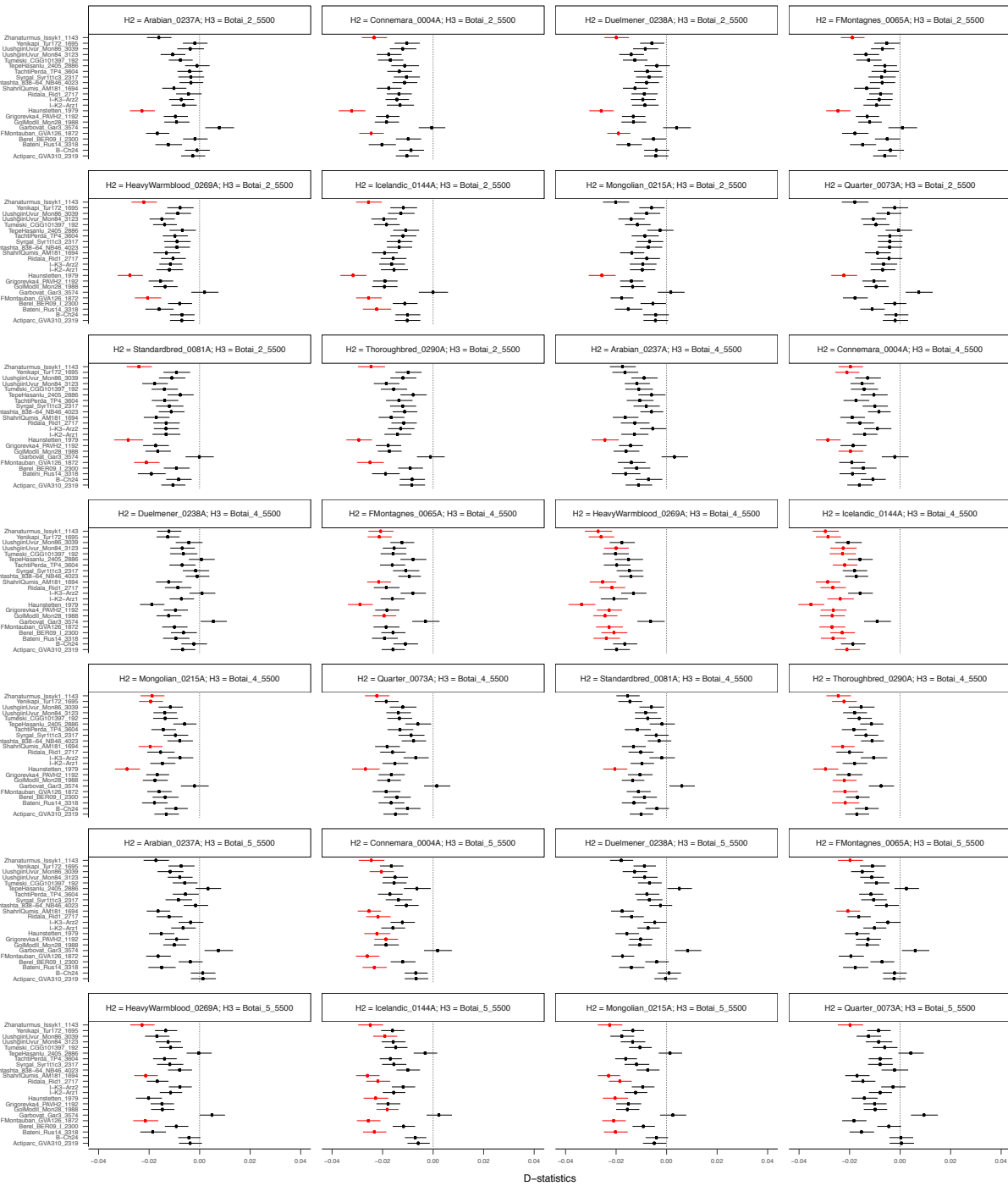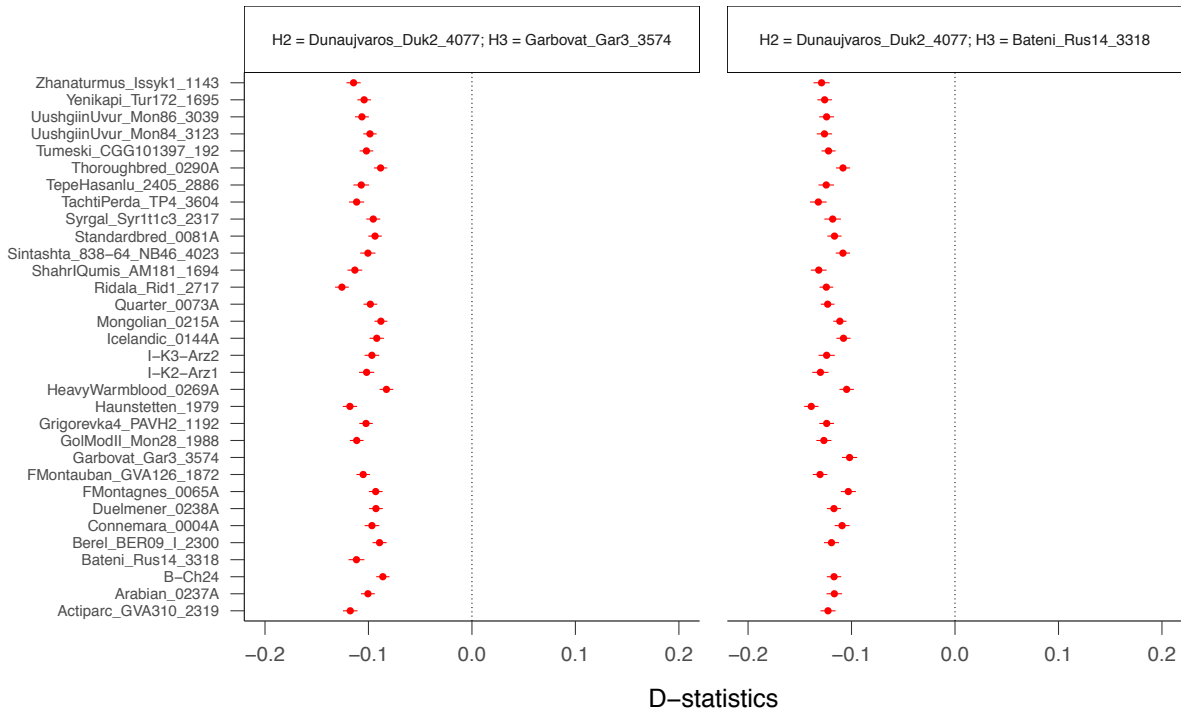**Fig. S32.**

D-statistics in the form of (((X, Dunaujvaros_Duk2_4077), Y), donkey). The tree topology tested is indicated on each individual panel, with reference to the individual genomes considered for populations H_2 and H_3 (in $(((H_1, H_2), H_3),$Outgroup) topologies). X refers to any ancient or modern individual of the second domestic clade. The names of the specimens X are reported on the Y-axis. Y refers to Garbovat_Gar3_3574 or Bateni_Rus14_3318. Analyses have been carried out disregarding transitions. Red dots depict significant tests, with adjusted p-values ≤ 0.05 after correction for multiple testing.
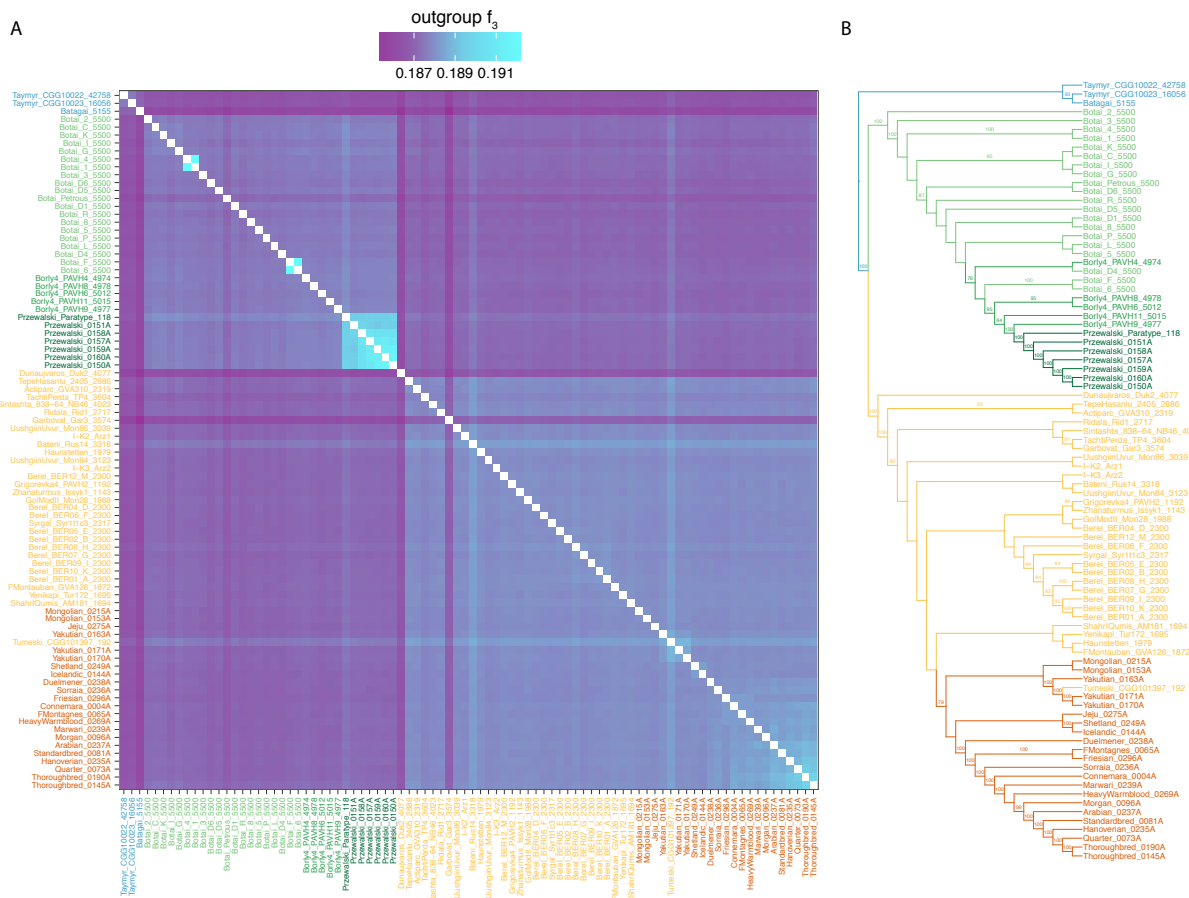
Fig. S33.
No impact of C to A (and G to T) error rates on genomic analyses. (A) f_3 statistics in the form (X, Y; Outgroup). (B) Phylogenetic relationships. For both analyses, 5,041,983 sites excluding C to A (and G to T) errors were considered.
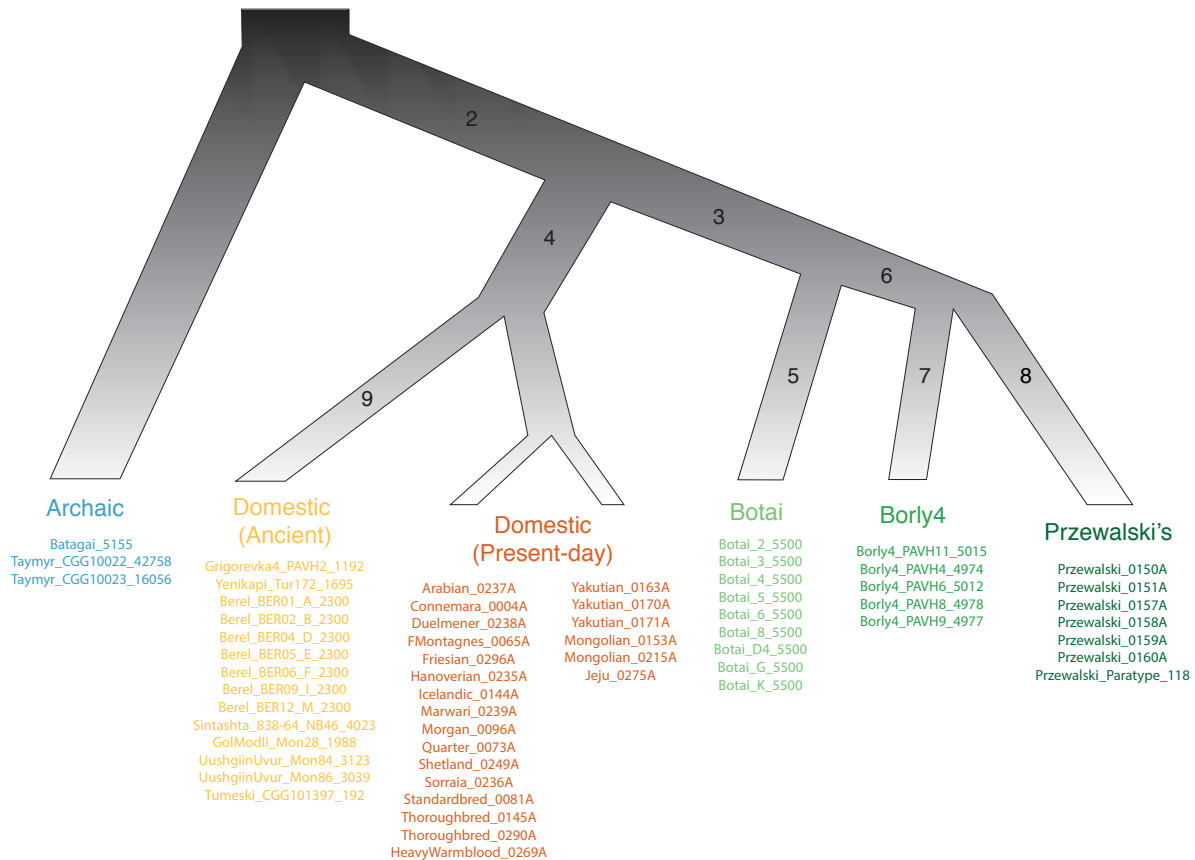
Fig. S34.
Population tree fixed for the LSDnorm analysis, following the population affinities recovered (see Fig. 2, Fig. S23). The investigated branches are enumerated from one to nine, in order to facilitate interpretation of Tables S17 and S18. Lineage 8, for example, represents the branch ancestral to Przewalski's horses, and encompasses the time period when they returned to the wild.

Supplementary Tables

Table S1.

Panel of ancient horses sequenced in this study. Archaeological, cultural and geographical information relating to the 42 ancient horses sequenced in this study, as well as inferred sex, endogenous content, nuclear and mitochondrial coverages.

Sample name	Registration number	Period	Age (years ago)	Culture	Site	Location	Sex	Hits post rescaling & trimming	Endogenous content	Prior to rescaling/trimming mtDNA coverage	nuDNA coverage	After rescaling/trimming mtDNA coverage	nuDNA coverage
Actiparc_GVA310_2319	CGG_1_019860	Late Iron Age	2,319	La Tène culture	Actiparc, Saint-Laurent Blangy	Pas-de-Calais, France	female	138027042	62.7%	119	2.62	119	2.62
Bateni_Rus14_3318	CGG_1_019164	Bronze Age	3,318.5	Karasuk culture	Kokorevo-Novoselovo	Khakassiya Republic, Russia	male	75359217	33.8%	130	1.74	129	1.74
Borly4_PAVH11_5015	CGG_1_018171	Copper Age	5,015.5	Borly culture	Borly 4	Pavlodar, Kazakhstan	female	289786430	71.0%	235	6.63	231	6.53
Borly4_PAVH4_4974	CGG_1_018157	Copper Age	4,974.5	Borly culture	Borly 4	Pavlodar, Kazakhstan	female	309220021	67.6%	465	7.15	465	7.15
Borly4_PAVH6_5012	CGG_1_018161	Copper Age	5,012	Borly culture	Borly 4	Pavlodar, Kazakhstan	female	208715594	74.1%	201	4.56	198	4.54
Borly4_PAVH8_4978	CGG_1_018165	Copper Age	4,978.5	Borly culture	Borly 4	Pavlodar, Kazakhstan	male	426739222	67.9%	269	9.14	269	9.14
Borly4_PAVH9_4977	CGG_1_018167	Copper Age	4,977	Borly culture	Borly 4	Pavlodar, Kazakhstan	male	216933892	70.2%	199	5.27	197	5.27
Botai_1_5500	CGG_1_018173	Copper Age	~5,500	Botai culture	Botai	Aqmola Province, Kazakhstan	male	385911344	71.6%	283	8.58	280	8.48
Botai_2_5500	CGG_1_018174	Copper Age	~5,500	Botai culture	Botai	Aqmola Province, Kazakhstan	male	380314821	74.7%	602	9.29	599	9.29
Botai_3_5500	CGG_1_018175	Copper Age	~5,500	Botai culture	Botai	Aqmola Province, Kazakhstan	male	139650939	70.0%	72.6	3.09	239	3.04
Botai_4_5500	CGG_1_018176	Copper Age	~5,500	Botai culture	Botai	Aqmola Province, Kazakhstan	male	434974462	72.6%	240	9.17	240	9.17
Botai_5_5500	CGG_1_018177	Copper Age	~5,500	Botai culture	Botai	Aqmola Province, Kazakhstan	male	337355963	72.7%	252	8.26	248	8.12
Botai_6_5500	CGG_1_018178	Copper Age	~5,500	Botai culture	Botai	Aqmola Province, Kazakhstan	male	356499791	74.9%	479	8.57	479	8.57
Botai_8_5500	CGG_1_020179	Copper Age	~5,500	Botai culture	Botai	Aqmola Province, Kazakhstan	female	149199894	40.4%	157	3.10	155	3.10
Botai_C_5500	CGG_1_020183	Copper Age	~5,500	Botai culture	Botai	Aqmola Province, Kazakhstan	male	167688216	68.0%	127	2.93	127	2.93
Botai_D1_5500	CGG_1_020204	Copper Age	~5,500	Botai culture	Botai	Aqmola Province, Kazakhstan	male	116742324	57.4%	177	2.76	176	2.76
Botai_D4_5500	CGG_1_020207	Copper Age	~5,500	Botai culture	Botai	Aqmola Province, Kazakhstan	female	139542935	61.0%	240	3.09	237	3.09

Botai_D5_5500	CGG_1_020208	Copper Age	~5,500	Botai culture	Botai	Aqmola Province, Kazakhstan	male	118747823	44.3%	242	2.63	239	2.63
Botai_D6_5500	CGG_1_020209	Copper Age	~5,500	Botai culture	Botai	Aqmola Province, Kazakhstan	female	101863307	58.3%	254	2.27	250	2.27
Botai_F_5500	CGG_1_020186	Copper Age	~5,500	Botai culture	Botai	Aqmola Province, Kazakhstan	male	136256533	62.9%	138	3.05	136	3.05
Botai_G_5500	CGG_1_020187	Copper Age	~5,500	Botai culture	Botai	Aqmola Province, Kazakhstan	male	203404622	71.6%	221	4.41	221	4.41
Botai_I_5500	CGG_1_020189	Copper Age	~5,500	Botai culture	Botai	Aqmola Province, Kazakhstan	male	138810757	73.9%	130	2.54	130	2.54
Botai_K_5500	CGG_1_020191	Copper Age	~5,500	Botai culture	Botai	Aqmola Province, Kazakhstan	male	211362992	69.6%	192	4.49	192	4.49
Botai_LM_5500	CGG_1_020192	Copper Age	~5,500	Botai culture	Botai	Aqmola Province, Kazakhstan	female	118980941	68.2%	171	3.61	169	2.53
Botai_P_5500	CGG_1_020196	Copper Age	~5,500	Botai culture	Botai	Aqmola Province, Kazakhstan	male	122412826	68.6%	113	4.95	113	2.69
Botai_Petrous_5500	CGG_1_020210	Copper Age	~5,500	Botai culture	Botai	Aqmola Province, Kazakhstan	female	84550523	73.2%	56.5	1.62	56.0	1.62
Botai_R_5500	CGG_1_020198	Copper Age	~5,500	Botai culture	Botai	Aqmola Province, Kazakhstan	female	91991287	72.6%	74.8	1.83	74.4	1.82
Dunaujvaros_Duk2_4077	CGG_1_018386	Bronze Age	4,077	Vatya culture	Dunaújváros-Koszider	Hungary	male	44733270	16.8%	60.0	1.35	41.5	1.12
FMontauban_GVA126_1872	CGG_1_019384	Roman period	1,872	Gallo-roman culture	Fresnes les Montauban	Pas-de-Calais, France	male	96959825	72.1%	63.5	2.22	62.5	2.18
Garbovat_Gar3_3574	CGG_1_018389	Bronze Age	3,574		Garbovat	Romania	male	47094944	16.6%	90.8	1.74	55.4	1.11
GolModII_Mon28_1988	CGG_1_018064	Late Iron Age	1,988.5	Xiongnu culture	Gol Mod II	Arkhangai Province, Central Mongolia	male	180301630	73.4%	143	4.10	143	4.10
Grigorevka4_PAVH2_1192	CGG_1_018154	Middle Ages	1,192	Kimek	Gregorievka valley	Gregorievka valley, Kazakhstan	male	220041428	75.7%	114	5.62	112	5.54
Haunstetten_1979	CGG_1_017139		1,979	Roman	Haunstetten	Augsburg, Germany	male	91300913	63.8%	114	2.29	96.5	1.92
Ridala_Rid1_2717	CGG_1_018468	Iron Age	2,600-2,800	Iron Age Estonia	Ridala	Saaremaa Island, Estonia	male	98217919	45.0%	255	4.34	251	2.17
ShahrIQumis_AM181_1694	CGG_1_018580	Historic period	1,694.5	Parthian/Sasanian	Shahr-i-Qumis	Semnan Province, Iran	male	133576425	71.9%	105	2.86	104	2.85
Syrgal_Syr1t1c3_2317	CGG_1_018919	Iron Age	2,317	Pazyryk culture	Syrgal	Bayan-Ulgii Province, Mongolia	male	76728148	61.7%	64.6	2.12	64.6	2.12
TachtiPerda_TP4_3604	CGG_1_018394	Bronze Age	3,604		Tachti Perda	Georgia	male	79743592	21.3%	100	1.62	100	1.62
TepeHasanlu_2405_2886	CGG_1_019997	Iron Age II	2,886.5	Mannaean kingdom	Hasanlu	West Azerbaijan province, Iran	male	55106394	45.3%	162	1.77	162	1.76
UushgiinUvur_Mon84_3123	CGG_1_018120	Bronze Age	3,123.50	Deer stone culture	Uushgiin Uvur Deer Stone	Hövsgöl Province, Northern Mongolia	male	372331275	72.3%	312	8.31	308	8.26
UushgiinUvur_Mon86_3039	CGG_1_018122	Bronze Age	3,039.50	Deer stone culture	Uushgiin Uvur Deer Stone	Hövsgöl Province, Northern Mongolia	female	218239512	72.9%	189	5.12	186	5.10

Yenikapi_Tur172_1695	CGG_1_018738	Byzantine	1,695	Byzantine culture	Yenikapi	Istanbul, Turkey	male	2487599292	71.9%	223	5.78	223	5.78
Zhanaturmus_Issyk1_1143	CGG_1_018577	Middle Ages	1,143	Turk period	Zhanaturmus	Almaty Province, Kazakhstan	male	111164629	77.9%	77.7	2.69	77.7	2.69

Table S2.

Calibrated radiocarbon measurement summary statistics and dating of 20 ancient horses sequenced in this study. Calibration was performed using OxCalOnline with the IntCal13 calibration curve.

Sample name	Registration number	>30kDa collagen yield(%)	Δ15N (‰)	Δ13C(‰)	±	%N	%C	C/N (wt%/wt%)	C/N (atomic)	Fraction Modern	±	Δ14C(‰)	±	14C age(BP)	±	From	To	%	Average	Years ago (from 2017)
ShahrIQumis_AM181_1694	CGG_1_018580	9.3	8.186	-17.974	0.1	16.4	45.6	2.79	3.25	0.8075	0.0018	-192.5	1.8	1715	20	255	390	95.4	322.5	**-1694.5**
Dunaujvaros_Duk2_4077	CGG_1_018386	3	3.999	-20.779	0.1	16.1	44.3	2.75	3.21	0.6324	0.0012	-367.6	1.2	3680	20	-2139	-1981	95.4	-2060	**-4077**
Garbovat_Gar3_3574	CGG_1_018389	6.4	4.283	-21.322	0.1	16.5	45.5	2.76	3.22	0.6654	0.0013	-334.6	1.3	3270	20	-1611	-1503	95.4	-1557	**-3574**
FMontauban_GVA126_1872	CGG_1_019384	1.4	6.605	-22.114	0.1	15.2	43	2.82	3.29	0.7919	0.0015	-208.1	1.5	1875	15	77	213	95.4	145	**-1872**
Actiparc_GVA310_2319	CGG_1_019860	1.5	5.111	-23.122	0.1	14.8	42.2	2.84	3.32	0.7553	0.0015	-244.7	1.5	2255	20	-393	-211	95.4	-302	**-2319**
TepeHasanlu_2405_2886	CCG_1_019992	4.2	5.772	-20.53	0.1	15.8	42.7	2.7	3.15	0.7119	0.0013	-288.1	1.3	2730	15	-908	-831	95.4	-869.5	**-2886.5**
Zhanaturmus_Issyk1_1143	CGG_1_018577	5.5	7.457	-19.943	0.1	15.4	43.4	2.82	3.29	0.8673	0.0016	-132.7	1.6	1145	15	777	971	95.4	874	**-1143**
GolModII_Mon28_1988	CGG_1_018064	7.2	5.856	-19.329	0.1	16.5	45.4	2.75	3.21	0.7828	0.0015	-217.2	1.5	1965	20	-22	79	95.4	28.5	**-1988.5**
UushgiinUvur_Mon84_3123	CGG_1_018120	3.4	6.006	-21.022	0.1	14	43.5	3.11	3.63	0.6962	0.001	-303.8	1	2910	15	-1192	-1021	95.3	-1106.5	**-3123.5**
UushgiinUvur_Mon86_3039	CGG_1_018122	9.6	4.735	-20.139	0.1	16	44.3	2.77	3.23	0.7008	0.0014	-299.2	1.4	2855	20	-1110	-935	95.4	-1022.5	**-3039.5**
Borly4_PAVH11_5015	CGG_1_018171	2.1	7.122	-20.305	0.1	14.9	42.3	2.83	3.31	0.58	0.001	-420	1	4375	15	-3078	-2919	95.4	-2998.5	**-5015.5**
Grigorevka4_PAVH2_1192	CGG_1_018154	7.5	3.916	-20.761	0.1	15.9	43.9	2.77	3.23	0.8604	0.0013	-139.6	1.3	1210	15	768	882	95.4	825	**-1192**
Borly4_PAVH4_4974	CGG_1_018157	2.4	7.275	-19.194	0.1	14.7	41.6	2.84	3.31	0.5824	0.0009	-417.6	0.9	4340	15	-3011	-2904	95.4	-2957.5	**-4974.5**
Borly4_PAVH6_5012	CGG_1_018161	3.1	7.203	-19.401	0.1	15.8	44.3	2.8	3.26	0.5809	0.0012	-419.1	1.2	4365	20	-3078	-2912	95.4	-2995	**-5012**
Borly4_PAVH8_4978	CGG_1_018165	2.5	6.586	-20.107	0.1	14.5	41.7	2.88	3.36	0.5817	0.0009	-418.3	0.9	4350	15	-3013	-2910	95.4	-2961.5	**-4978.5**
Borly4_PAVH9_4977	CGG_1_018167	4.7	4.778	-20.031	0.1	15.2	42.5	2.79	3.26	0.5823	0.0009	-417.7	0.9	4345	15	-3013	-2907	95.4	-2960	**-4977**

Bateni_Rus14_3318	CGG_1_019164	10.7	4.342	-20.191	0.1	16.3	45	2.76	3.23	0.6857	0.0013	-314.3	1.3	3030	20	-1386	-1217	95.4	-1301.5	**-3318.5**
Syrgal_Syr1t1c3_2317	CGG_1_018919	7.1	5.809	-20.374	0.1	15.9	43.6	2.74	3.2	0.7556	0.0014	-244.4	1.4	2250	15	-389	-211	95.5	-300	**-2317**
TachtiPerda_TP4_3604	CGG_1_018394	7.2	4.849	-17.969	0.1	15.8	43.4	2.74	3.2	0.6616	0.0012	-338.4	1.2	3320	15	-1643	-1531	95.4	-1587	**-3604**
Yenikapi_Tur172_1695	CGG_1_018738	3.9	7.024	-19.965	0.1	14.9	42.5	2.86	3.34	0.8077	0.0012	-192.3	1.2	1715	15	257	387	95.4	322	**-1695**

Table S3.

Time to the Most Common Recent Ancestor (tMRCA) of the 211 mitochondrial genomes. The tMRCA posterior distribution is summarized from a total number of 617 million MCMC states in BEAST, disregarding the first 25% as burn-in, and assuming a LogNormal autocorrelated clock model and a Bayesian Skyline demographic model.

	tMRCA (years ago)
Mean	114,850
Standard Deviation	14,126
Median	113,890
2.5% Percentile	87,789
97.5% Percentile	142,550

Table S4.

Genotyping at the *TBX3* locus for the 20 Botai and 5 Borly4 horses. SNP1 (chr8:18,227,267 + 1,067) and SNP2 (chr8:18,226,905) denote two genomic sites associated with the Dun color *(147)*. SNP2 is found 1,067 bp inside a DNA fragment not present in the reference assembly (the insertion site in EquCab2 is chr8:18,227,267). Based on 1,814 horses of known phenotype, Imsland and colleagues (*147)* found three different combination of alleles for these two SNPs. The three haplotypes were "GG" (carrier horses are Dun), "AT" (non-dun1) and "A-deletion" (non-dun2), where "GG" is dominant over "AT" and over "A-deletion". The columns for SNP1 and SNP2 show the number of reads that pass ANGSD quality filters (section S6.3.a) and support each allele variant at those positions. We visually checked for the presence of the deletion in Botai_P_5500, and found multiple reads mapping within its coordinates (chr8:18,227,267 + 1,067). This indicates that Botai_P_5500 did not carry the deletion. Altogether, no horse was found to be non-dun2, in line with the recent origin of this derived haplotype.

Sample	SNP2 (chr8:18,226,905)	SNP1 (chr8:18,227,267 + 1,067)	Haplotype	Most likely phenotype
Botai_8_5500	2G	2G	GG	Dun
Botai_C_5500	2G	--	G?	Dun
Botai_D1_5500	1A; 1G	1G	[A/G]G	Dun
Botai_D4_5500	2G	3G	GG	Dun
Botai_D5_5500	2A	2G	AG	?
Botai_D6_5500	1A; 1G	4G	[A/G]G	Dun
Botai_F_5500	1A; 1G	2G	[A/G]G	Dun
Botai_G_5500	1A; 1G	4G	[A/G]G	Dun
Botai_I_5500	--	2G	?G	Dun
Botai_K_5500	2G	3G; 1T	G[G/T]	Dun
Botai_LM_5500	1G; 1A	3G; 2T	[A/G][G/T]	Dun
Botai_P_5500	--	--	??	Dun or non-dun1
Botai_Petrous_5500	1A; 1G	3G	[A/G]G	Dun
Botai_R_5500	--	3G	?G	Dun
Botai_1_5500	4A; 1G	7G	[A/G]G	Dun
Botai_2_5500	5G	8G	GG	Dun
Botai_3_5500	1A; 1G	2G	[A/G]G	Dun
Botai_4_5500	3A; 1G	7G	[A/G]G	Dun
Botai_5_5500	3A; 1G	8G	[A/G]G	Dun (but heterozygous for the *TRMP1* haplotype associated with CSBN and leopard spotting)
Botai_6_5500	2A; 4G	9G	[A/G]G	Dun
Borly4_PAVH11_5015	2A; 2G	4G	[A/G]G	Dun
Borly4_PAVH4_4974	3A	2G; 4T	A[G/T]	Dun or non-dun1
Borly4_PAVH6_5012	7A	6G	AG	?
Borly4_PAVH8_4978	2A	8G; 3T	[A/G][G/T]	Dun
Borly4_PAVH9_4977	1A; 5G	8T	[A/G]T	Dun or non-dun1

Table S5.
Variance explained for TreeMix models, from 0 to 5 migration edges. Limited improvement is observed for an increasing number of migration edges.

Migration edges	Variance
0	0.9912298
1	0.991649
2	0.9918616
3	0.9921868
4	0.9926172
5	0.9927976

Table S6.
Relatedness between Botai and Borly4 individuals, in percentages. Cells highlighted in red depict horses exhibiting an elevated kinship coefficient. Botai_1_5500 and Botai_F_5500 were consequently excluded from multiple analyses.

	Botai_2_5500	Botai_3_5500	Botai_4_5500	Botai_5_5500	Botai_6_5500	Botai_8_5500	Botai_C_5500	Botai_D1_5500	Botai_D4_5500	Botai_D5_5500	Botai_D6_5500	Botai_F_5500	Botai_G_5500	Botai_I_5500	Botai_K_5500	Botai_LM_5500	Botai_P_5500	Botai_Petrous_5500	Botai_R_5500	Borly4_PAVH6_5012	Borly4_PAVH9_4977	Borly4_PAVH11_5015	Borly4_PAVH4_4974	Borly4_PAVH8_4978
Botai_1_5500	0.00	0.00	**29.31**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Botai_2_5500		0.99	0.00	0.25	0.00	0.31	0.08	0.00	0.00	0.38	0.00	0.00	0.00	0.00	0.06	0.00	0.00	0.00	0.00	0.13	0.44	0.38	0.19	0.55
Botai_3_5500			0.00	0.24	0.00	0.29	0.00	0.16	0.00	0.28	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.36	0.18	0.00	0.24	0.16	0.00	0.29
Botai_4_5500				0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Botai_5_5500					0.00	0.04	0.00	1.60	0.79	0.37	0.65	0.00	0.00	0.00	0.00	0.64	0.55	0.06	0.00	0.00	0.00	0.49	0.00	0.25
Botai_6_5500						0.00	0.00	0.00	0.00	0.00	0.00	**28.13**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Botai_8_5500							0.00	0.40	0.00	0.40	0.13	0.00	0.00	0.00	0.00	0.14	0.52	0.34	0.00	0.00	0.00	0.07	0.00	0.47
Botai_C_5500								0.00	0.00	0.00	0.00	0.00	0.69	0.67	0.60	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.19	0.00
Botai_D1_5500									0.91	0.59	0.40	0.00	0.00	0.00	0.00	0.01	0.52	0.61	0.24	0.00	0.00	0.65	0.00	0.00
Botai_D4_5500										0.00	0.48	0.00	0.00	0.00	0.00	0.31	0.78	0.29	0.00	0.00	0.00	0.00	0.65	0.00
Botai_D5_5500											0.50	0.00	0.00	0.00	0.00	0.00	0.42	0.47	0.00	0.00	0.00	0.46	0.00	0.00
Botai_D6_5500												0.00	0.00	0.00	0.00	0.00	0.30	0.50	0.24	0.00	0.00	0.00	0.00	0.23
Botai_F_5500													0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Botai_G_5500														0.63	1.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.27	0.00
Botai_I_5500															0.88	0.00	0.00	0.00	0.00	0.00	0.19	0.00	0.00	0.00
Botai_K_5500																0.00	0.00	0.00	0.00	0.05	0.00	0.35	0.00	0.00
Botai_LM_5500																	0.20	0.00	0.09	0.00	0.00	0.00	0.00	0.00
Botai_P_5500																		0.75	0.19	0.00	0.00	0.00	0.72	0.00
Botai_Petrous_5500																			0.21	0.00	0.00	0.00	0.06	0.00
Botai_R_5500																				0.00	0.00	0.00	0.00	0.00
Borly4_PAVH6_5012																					0.08	0.04	0.19	0.92
Borly4_PAVH9_4977																						0.77	0.36	0.18
Borly4_PAVH11_5015																							0.10	0.59
Borly4_PAVH4_4974																								0.57

Table S7.

The 107 genes overlapping the significant LSDnorm regions in the branch ancestral to Przewalski's horses. Elevated LSDnorm values reveal local gene trees where Botai, Borly4 and ancient and modern domesticates have a very similar haplotype, and different to that observed within Przewalski's horses.

Gene stable ID	Gene name	Gene description
ENSECAG00000000700	EPRS	glutamyl-prolyl-tRNA synthetase [Source:HGNC Symbol;Acc:HGNC:3418]
ENSECAG00000001069	CCRL2	C-C motif chemokine receptor like 2 [Source:HGNC Symbol;Acc:HGNC:1612]
ENSECAG00000001515		
ENSECAG00000001873		
ENSECAG00000002155	NUP93	nucleoporin 93 [Source:HGNC Symbol;Acc:HGNC:28958]
ENSECAG00000002380	RAB35	RAB35, member RAS oncogene family [Source:HGNC Symbol;Acc:HGNC:9774]
ENSECAG00000002617		
ENSECAG00000003006		
ENSECAG00000003445	C20orf96	chromosome 20 open reading frame 96 [Source:HGNC Symbol;Acc:HGNC:16227]
ENSECAG00000004135	AMPH	amphiphysin [Source:HGNC Symbol;Acc:HGNC:471]
ENSECAG00000004220	SLC6A11	solute carrier family 6 member 11 [Source:HGNC Symbol;Acc:HGNC:11044]
ENSECAG00000005166	MGAT2	mannosyl (alpha-1,6-)-glycoprotein beta-1,2-N-acetylglucosaminyltransferase [Source:HGNC Symbol;Acc:HGNC:7045]
ENSECAG00000005173		
ENSECAG00000005646		
ENSECAG00000005719	PRKD1	protein kinase D1 [Source:HGNC Symbol;Acc:HGNC:9407]
ENSECAG00000006257	SNTG2	syntrophin gamma 2 [Source:HGNC Symbol;Acc:HGNC:13741]
ENSECAG00000006851		
ENSECAG00000006925	WDR53	WD repeat domain 53 [Source:HGNC Symbol;Acc:HGNC:28786]
ENSECAG00000006983	RAPGEF2	Rap guanine nucleotide exchange factor 2 [Source:HGNC Symbol;Acc:HGNC:16854]
ENSECAG00000007104	NISCH	nischarin [Source:HGNC Symbol;Acc:HGNC:18006]
ENSECAG00000007183	SEMA3E	semaphorin 3E [Source:HGNC Symbol;Acc:HGNC:10727]
ENSECAG00000007493	ACAN	aggrecan [Source:HGNC Symbol;Acc:HGNC:319]
ENSECAG00000008545	ECM1	extracellular matrix protein 1 [Source:HGNC Symbol;Acc:HGNC:3153]
ENSECAG00000008743	AGBL3	ATP/GTP binding protein like 3 [Source:HGNC Symbol;Acc:HGNC:27981]
ENSECAG00000008753		
ENSECAG00000008754	NAA25	N(alpha)-acetyltransferase 25, NatB auxiliary subunit [Source:HGNC Symbol;Acc:HGNC:25783]
ENSECAG00000008839		
ENSECAG00000009059		
ENSECAG00000009147	CTCFL	CCCTC-binding factor like [Source:HGNC Symbol;Acc:HGNC:16234]
ENSECAG00000009450	DDC	dopa decarboxylase [Source:HGNC Symbol;Acc:HGNC:2719]

ENSECAG00000009692		
ENSECAG00000009755	EPHB3	EPH receptor B3 [Source:HGNC Symbol;Acc:HGNC:3394]
ENSECAG00000009891	ARMC5	armadillo repeat containing 5 [Source:HGNC Symbol;Acc:HGNC:25781]
ENSECAG00000010284	PKNOX1	PBX/knotted 1 homeobox 1 [Source:HGNC Symbol;Acc:HGNC:9022]
ENSECAG00000010901	CTNND2	catenin delta 2 [Source:HGNC Symbol;Acc:HGNC:2516]
ENSECAG00000010919	MYO9B	myosin IXB [Source:HGNC Symbol;Acc:HGNC:7609]
ENSECAG00000010989	SH3YL1	SH3 and SYLF domain containing 1 [Source:HGNC Symbol;Acc:HGNC:29546]
ENSECAG00000011559	RHOBTB1	Rho related BTB domain containing 1 [Source:HGNC Symbol;Acc:HGNC:18738]
ENSECAG00000011853		
ENSECAG00000011964	IL27RA	interleukin 27 receptor subunit alpha [Source:HGNC Symbol;Acc:HGNC:17290]
ENSECAG00000012031		
ENSECAG00000012343	PNPLA1	patatin like phospholipase domain containing 1 [Source:HGNC Symbol;Acc:HGNC:21246]
ENSECAG00000012896	CDC37	cell division cycle 37 [Source:HGNC Symbol;Acc:HGNC:1735]
ENSECAG00000012982		
ENSECAG00000013022	ACTC1	actin, alpha, cardiac muscle 1 [Source:HGNC Symbol;Acc:HGNC:143]
ENSECAG00000013280	AADACL3	arylacetamide deacetylase like 3 [Source:HGNC Symbol;Acc:HGNC:32037]
ENSECAG00000014059	ECT2	epithelial cell transforming 2 [Source:HGNC Symbol;Acc:HGNC:3155]
ENSECAG00000014071	SLC30A7	solute carrier family 30 member 7 [Source:HGNC Symbol;Acc:HGNC:19306]
ENSECAG00000014276	TARS2	threonyl-tRNA synthetase 2, mitochondrial (putative) [Source:HGNC Symbol;Acc:HGNC:30740]
ENSECAG00000014897	RLN3	relaxin 3 [Source:HGNC Symbol;Acc:HGNC:17135]
ENSECAG00000015035	TM2D1	TM2 domain containing 1 [Source:HGNC Symbol;Acc:HGNC:24142]
ENSECAG00000015123		
ENSECAG00000015387		
ENSECAG00000015452	MCM8	minichromosome maintenance 8 homologous recombination repair factor [Source:HGNC Symbol;Acc:HGNC:16147]
ENSECAG00000015892	CACNA1D	calcium voltage-gated channel subunit alpha1 D [Source:HGNC Symbol;Acc:HGNC:1391]
ENSECAG00000016269		
ENSECAG00000016493	PHF12	PHD finger protein 12 [Source:HGNC Symbol;Acc:HGNC:20816]
ENSECAG00000016513	NLGN2	neuroligin 2 [Source:HGNC Symbol;Acc:HGNC:14290]
ENSECAG00000016794	ANKRD13C	ankyrin repeat domain 13C [Source:HGNC Symbol;Acc:HGNC:25374]
ENSECAG00000017329	CPB2	carboxypeptidase B2 [Source:HGNC Symbol;Acc:HGNC:2300]
ENSECAG00000017387	SPEM1	spermatid maturation 1 [Source:HGNC Symbol;Acc:HGNC:32429]
ENSECAG00000017640	ILF3	interleukin enhancer binding factor 3 [Source:HGNC Symbol;Acc:HGNC:6038]
ENSECAG00000017827	ANTXR2	anthrax toxin receptor 2 [Source:HGNC Symbol;Acc:HGNC:21732]
ENSECAG00000017928	NDUFA13	NADH:ubiquinone oxidoreductase subunit A13 [Source:HGNC Symbol;Acc:HGNC:17194]
ENSECAG00000018536	USE1	unconventional SNARE in the ER 1 [Source:HGNC Symbol;Acc:HGNC:30882]

ENSECAG00000018617	RPRD2	regulation of nuclear pre-mRNA domain containing 2 [Source:HGNC Symbol;Acc:HGNC:29039]
ENSECAG00000018867	LINS1	lines homolog 1 [Source:HGNC Symbol;Acc:HGNC:30922]
ENSECAG00000018924	PTK2B	protein tyrosine kinase 2 beta [Source:HGNC Symbol;Acc:HGNC:9612]
ENSECAG00000019081	ST7L	suppression of tumorigenicity 7 like [Source:HGNC Symbol;Acc:HGNC:18441]
ENSECAG00000019425	ALKBH8	alkB homolog 8, tRNA methyltransferase [Source:HGNC Symbol;Acc:HGNC:25189]
ENSECAG00000019533	FGFR3	fibroblast growth factor receptor 3 [Source:HGNC Symbol;Acc:HGNC:3690]
ENSECAG00000019594	DENND2C	DENN domain containing 2C [Source:HGNC Symbol;Acc:HGNC:24748]
ENSECAG00000019893	RIIAD1	regulatory subunit of type II PKA R-subunit (RIIa) domain containing 1 [Source:HGNC Symbol;Acc:HGNC:26686]
ENSECAG00000019950	RALGAPA2	Ral GTPase activating protein catalytic alpha subunit 2 [Source:HGNC Symbol;Acc:HGNC:16207]
ENSECAG00000020111	IQGAP3	IQ motif containing GTPase activating protein 3 [Source:HGNC Symbol;Acc:HGNC:20669]
ENSECAG00000020127	THEM4	thioesterase superfamily member 4 [Source:HGNC Symbol;Acc:HGNC:17947]
ENSECAG00000020313	SLC16A10	solute carrier family 16 member 10 [Source:HGNC Symbol;Acc:HGNC:17027]
ENSECAG00000020343	SKAP2	src kinase associated phosphoprotein 2 [Source:HGNC Symbol;Acc:HGNC:15687]
ENSECAG00000020357	SETDB2	SET domain bifurcated 2 [Source:HGNC Symbol;Acc:HGNC:20263]
ENSECAG00000020452	TNNC1	troponin C1, slow skeletal and cardiac type [Source:HGNC Symbol;Acc:HGNC:11943]
ENSECAG00000020618		
ENSECAG00000020647	PLCB2	phospholipase C beta 2 [Source:HGNC Symbol;Acc:HGNC:9055]
ENSECAG00000020651	CCAR2	cell cycle and apoptosis regulator 2 [Source:HGNC Symbol;Acc:HGNC:23360]
ENSECAG00000020775	C9orf85	chromosome 9 open reading frame 85 [Source:HGNC Symbol;Acc:HGNC:28784]
ENSECAG00000020807	THEM5	thioesterase superfamily member 5 [Source:HGNC Symbol;Acc:HGNC:26755]
ENSECAG00000020838	USP47	ubiquitin specific peptidase 47 [Source:HGNC Symbol;Acc:HGNC:20076]
ENSECAG00000021082	ZBTB46	zinc finger and BTB domain containing 46 [Source:HGNC Symbol;Acc:HGNC:16094]
ENSECAG00000021208	RC3H1	ring finger and CCCH-type domains 1 [Source:HGNC Symbol;Acc:HGNC:29434]
ENSECAG00000021302	POC1A	POC1 centriolar protein A [Source:HGNC Symbol;Acc:HGNC:24488]
ENSECAG00000021308		
ENSECAG00000021490		
ENSECAG00000021622	WARS2	tryptophanyl tRNA synthetase 2, mitochondrial [Source:HGNC Symbol;Acc:HGNC:12730]
ENSECAG00000022194	ANKRD13A	ankyrin repeat domain 13A [Source:HGNC Symbol;Acc:HGNC:21268]
ENSECAG00000022244	CYP26A1	cytochrome P450 family 26 subfamily A member 1 [Source:HGNC Symbol;Acc:HGNC:2603]
ENSECAG00000022609	CHRNA2	cholinergic receptor nicotinic alpha 2 subunit [Source:HGNC Symbol;Acc:HGNC:1956]
ENSECAG00000023164		
ENSECAG00000023338	CSTF3	cleavage stimulation factor subunit 3 [Source:HGNC Symbol;Acc:HGNC:2485]
ENSECAG00000023411		
ENSECAG00000023488	MED13L	mediator complex subunit 13 like [Source:HGNC Symbol;Acc:HGNC:22962]
ENSECAG00000024031	KIAA0232	KIAA0232 [Source:HGNC Symbol;Acc:HGNC:28992]

ENSECAG00000024050	SCLT1	sodium channel and clathrin linker 1 [Source:HGNC Symbol;Acc:HGNC:26406]
ENSECAG00000024081	BCL2L15	BCL2 like 15 [Source:HGNC Symbol;Acc:HGNC:33624]
ENSECAG00000024377	RXRG	retinoid X receptor gamma [Source:HGNC Symbol;Acc:HGNC:10479]
ENSECAG00000024560	PTPN11	protein tyrosine phosphatase, non-receptor type 11 [Source:HGNC Symbol;Acc:HGNC:9644]
ENSECAG00000024564	PEBP1	phosphatidylethanolamine binding protein 1 [Source:HGNC Symbol;Acc:HGNC:8630]
ENSECAG00000024616		
ENSECAG00000025033	RAMP3	receptor activity modifying protein 3 [Source:HGNC Symbol;Acc:HGNC:9845]

Table S8.
Ancient comparative panel. ID, temporal and cultural background associated with each of the 18 previously published ancient individuals included in this study, as well as inferred sex, accession number and study of publication, nuclear and mitochondrial coverages prior and after trimming and rescaling.

Table S9.
Comparative panel of modern domestic horses. ID, breed, sex, accession number and study of publication, nuclear and mitochondrial coverages associated with each of the 28 modern horses included in this study.

Table S10.
Raw morphological measurements of 375 horse specimens, for the third metacarpal.
Measurement #1: Greatest length, #2: Lateral length, #3: Breadth at midshaft, #4: Depth at midshaft, #5: Proximal breadth, #6: Proximal articular depth, #7: Diameter of articular facet for Os carpale III, #8: Diameter of anterior facet for Os carpale IV, #8': Diameter of posterior facet for Os carpale IV, #9: Diameter of facet for Os carpal II, #10: Distal supra-articular breadth, #11: Distal articular breadth, #12: Depth of sagittal keel, #13: Smallest depth of medial condyle, #14: Greatest depth of medial condyle.

Table S11.
Laboratory Procedures carried out on each sample sequenced in this study. (A) Summary information on DNA extractions, library construction, and amplification. (B) Independent libraries built with list of external index primers and internal indexed adapters selected as well as sequencing platform.

Table S12.
Sequencing statistics per library for each sample sequenced in this study. nuDNA and mtDNA refer to nuclear DNA and mitochondrial DNA, respectively. All values have been calculated based on sequencing output prior to rescaling.

Table S13.
Error rates for each substitution type. Reads showing mapping quality scores strictly inferior to 30 were disregarded. Bases showing quality scores strictly inferior to 20 were disregarded.

Table S14.
Error rates for each substitution type. Reads showing mapping quality scores strictly inferior to 30 were disregarded. Bases showing quality scores strictly inferior to 30 were disregarded.

Table S15.
Panel of modern and ancient horses included in mitochondrial analyses. Sample name, population, age, accession or registration number and reference associated with each individual.

Table S16.
Statistical *p*-values for the test of direct ancestry. Shaded cells highlight significance, where the horse in the row belonged to a population that was directly ancestral to the population of the horse in the column

Table S17.
WebGestalt analysis for genes overlapping significant LSDnorm regions. The 'branch' column indicates the population investigated, as labeled in Fig. S34. Branch 8, for example, corresponds to the branch ancestral to Przewalski's horses. Only categories with an adjusted *p*-value < 0.1 are shown.

Table S18. Developmental stages and anatomical body parts where genes overlapping significant LSDnorm windows are transcribed, as revealed by Bgee (*173*). The 'branch' column indicates the population investigated, as labeled in Fig. S34. Branch 4, for example, corresponds to the branch ancestral to Botai and Borly4 horses. No over-represented categories were found in lineages 8 and 9. Only transcriptional categories with an adjusted *p*-value < 0.1 are shown.