



**HAL**  
open science

## Genomics and the Evolutionary History of Equids

Pablo Librado, Ludovic Orlando

► **To cite this version:**

Pablo Librado, Ludovic Orlando. Genomics and the Evolutionary History of Equids. Annual Review of Animal Biosciences, 2021, 9 (1), 10.1146/annurev-animal-061220-023118 . hal-03030307

**HAL Id: hal-03030307**

**<https://hal.science/hal-03030307v1>**

Submitted on 30 Nov 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Genomics and the Evolutionary History of Equids

Pablo Librado and Ludovic Orlando

Laboratoire d'Anthropobiologie Moléculaire et d'Imagerie de Synthèse, CNRS UMR 5288,  
Université Paul Sabatier, Toulouse 31000, France; email: [plibradosanz@gmail.com](mailto:plibradosanz@gmail.com)

## Keywords

equid, horse, evolution, donkey, ancient DNA, population genomics

## Abstract

The equid family contains only one single extant genus, *Equus*, including seven living species grouped into horses on the one hand and zebras and asses on the other. In contrast, the equine fossil record shows that an extraordinarily richer diversity existed in the past and provides multiple examples of a highly dynamic evolution punctuated by several waves of explosive radiations and extinctions, cross-continental migrations, and local adaptations. In recent years, genomic technologies have provided new analytical solutions that have enhanced our understanding of equine evolution, including the species radiation within *Equus*; the extinction dynamics of several lineages; and the domestication history of two individual species, the horse and the donkey. Here, we provide an overview of these recent developments and suggest areas for further research.

## FROM PALEONTOLOGY TO PALEOGENOMICS

Although many other genera roamed the Old and the New World in the past, only one single genus of equine species, *Equus*, currently exists. It is subdivided into three main subgenera, most likely representing a total of seven species (1). *Equus* subg. *Equus* groups together wild and domestic horses, whereas *Equus* subg. *Asinus* is composed of wild and domestic donkeys, including the Tibetan kiang and hemiones (or onagers) from Asia. Finally, *Equus* subg.

*Hippotrigis* includes only African species, namely, the mountain zebra, Grevy's zebra, and plains zebra. All of the more than 42 genera currently ascribed to the equid family are now extinct and thus have been defined based on the paleontological record, according to their morpho-anatomical characteristics and their geographic and temporal distribution ranges. In the last two decades, ancient DNA data have aided in revisiting classical and often contentious taxonomic categories, most often resulting in the collapse of lineages that were previously considered to form their own species.

The equid family gained its popularity among paleontologists as early as the late nineteenth century, when the famous anthropologist Thomas H. Huxley, also known to be one of the fiercest defenders of Darwin's theory of evolution, became fascinated by the collection of equine fossils excavated by Othniel C. Marsh. This series documented a striking evolutionary sequence accompanying the transition from small, multi-toed, and browsing animals to larger, single-toed grazers. Although it was originally misinterpreted as a mere linear succession into more advantageous morphological forms (2), this transformation became one flagship example of gradual macroevolutionary change, which has been part of every evolutionary biology textbook since its early characterization, when the theory of natural selection was challenged (3). Despite considerable progress within the field of paleontology, such as the development of novel scanning and imaging techniques (4), several difficulties have limited our understanding of equine evolution, especially at short, microevolutionary timescales. These limitations include the fragmentary nature of the fossil record, the extensive morphological plasticity found within particular groups, and pervasive instances of morphological convergence, all of which preclude robust delineation of the underlying species and/or taxa, and consequently of their phylogenetic relationships (5).

The sequencing of DNA molecules present in fossil remains has partially alleviated some of these limitations, providing an additional source of phylogenetic information alongside morpho-anatomical variation. As a matter of fact, the field of ancient DNA research started approximately one century after the original work of Huxley and Marsh, when Allan C. Wilson's team retrieved a 224-bp-long fragment of mitochondrial DNA from a museum specimen of the now-extinct quagga zebra (*Equus quagga quagga*). Albeit short, the sequence indicated close genetic affinities to the sympatric mountain zebras (*Equus zebra hartmannae*) (6), a phylogenetic placement that was further confirmed by the sequencing of their complete nuclear genomes

approximately 30 years later (7). In addition to clarifying one branch of the equine evolutionary tree, the original quagga study showed for the first time that DNA molecules can persist through time, despite the rapid accumulation of postmortem chemical damage (8). Ever since, ancient DNA has proved that extinction is not an absolute barrier for genetic investigations. Owing to technical challenges, nevertheless, early work prioritized the sequencing of high-copy, and hence more abundant, mitochondrial DNA markers from specimens dating to the last 10,000–100,000 years (9). Since the mid-2000s, our capacity to retrieve ancient DNA has largely improved, following a series of technical developments that have been thoroughly described elsewhere (10, 11), culminating with the advent of high-throughput sequencing. As a result, the characterization of complete ancient genomes has become increasingly affordable and is accomplished almost routinely in dedicated clean-lab facilities. Although the vast majority of ancient genomes sequenced thus far date to the last 1,000–10,000 years, much older specimens can also be analyzed, including from the Middle Pleistocene. The current record for the oldest genome was obtained from a horse metapodial that was preserved in the dry and cold Yukon permafrost for 560,000–780,000 years. The sequence helped calibrate an age of 4.0–4.5 million years for the age of the most recent common ancestor of the genus *Equus* (12).

## **EQUINE SPECIATION AND EXTINCTIONS OVER TIME**

Equids originated approximately ~55 Mya in North America from a group of small browsing species collectively known as hyracotheres (Figure 1a). This group was long considered a single genus, but cladistic reevaluation of fossil collections ruled out the monophyly of the *Hyracotherium* genus and even excluded the type species *Hyracotherium leporinum* from the family (13). *Sifrhippus sandrae* is now considered to be the most basal equid species. Beyond their origin, the fossil record suggests that equids experienced relatively limited morphological change during the Eocene (55–34 Mya) as well as during the Oligocene (34–23 Mya), until the mid-Miocene Climatic Optimum (mMCO) represented an evolutionary turning point approximately 18–15 Mya (14). At that time, temperatures rose, transforming terrestrial biomes from woodlands to mixed woodland–savanna habitats and opening a wide range of vacant ecological niches, which triggered one of the most explosive radiations ever characterized, giving rise to a minimum of 19 species (14). Many of the anatomical features that have become highly characteristic of extant equine species emerged during the following million years. These

include larger body sizes, digit fusion, and hypsodonty (i.e., the development of high-crowned molars, in relation to diet shifting to more abrasive, silica-rich grasses; [Figure 1a](#)).

[Figure 1](#) Deep evolutionary history of equids. (a) Important morpho-anatomical changes have punctuated equine evolution. (b) Phylogenetic tree replacing those extant and extinct members of the *Equus* genus discussed in the main text. Abbreviations: IBE, extinct lineage of horses, native from Iberia; NWSL, New World stilt-legged horses.

Whether the mMCO radiation was adaptive or not remains, however, contentious. For example, microwear abrasion signatures are undoubtedly related to habitat and dietary turnovers ([15](#)). Yet, tooth damage marks typical of mixed-grazing diets are present millions of years before hypsodonty emerged ([16](#), [17](#)). Likewise, monodactyly does not appear immediately optimal to withstand heavier body sizes and likely required additional conformational and locomotory changes ([18](#)). That extinct equid species often exhibited similar anatomies, with a range of intraspecific plasticity largely overlapping the magnitude of between-species divergence, further contradicts the theoretical outcome of adaptive radiations. Such scenarios, involving speciation pulses driven by morphological adaptations, are predicted to yield well-differentiated anatomical clades, in contrast to the morphological plasticity observed in the equid fossil record, especially regarding tooth shape ([19](#)). Combined, these studies suggest that the most iconic morphological characteristics of present-day equids likely emerged only as secondary adaptations and that the radiation process was not necessarily immediately adaptive ([20](#)).

As the rates of speciation and extinction remained balanced thereafter, the levels of species richness remained steady during the 10 million years following the mMCO radiation ([14](#)). The fossil record provides evidence that small, horselike creatures, retaining two potentially vestigial outer toes, were the most abundant equine tribe during that period ([21](#)). Known as the hipparionines, they successfully expanded far beyond their native northern American range and entered the Old World ~11.5 Mya. The success of this group was considerable, and they clearly outnumbered other taxonomical tribes considered potentially ancestral to present-day *Equus* members, such as the dinohippines ([22](#)). However, the continuous decline of temperatures since the mMCO ended ([23](#)) converted mixed woodland–savannas into dry and cold grasslands by the beginning of the Pliocene, some 5.3 Mya, and hipparionines experienced a dramatic drop in biodiversity. A few genera continued to exist in North America, but the last likely became extinct during the late Early Pleistocene, some 2 Mya ([24](#)).

## LATE-PLEISTOCENE EXTINCT EQUIDS FROM NORTH AMERICA

Besides hipparionines, other equine lineages subsisted in North America throughout the Pliocene. One such example is provided by the so-called New World stilt-legged (NWSL) horses, originally defined by the slender characteristics of their foot bones, which are reminiscent of modern Asiatic wild asses (25). Other equids, in contrast, thrived outside their native ranges. This is, for example, the case of hippidiforms and their characteristic genus, *Hippidion*, which entered into South America after the Panama isthmus opened approximately ~3 Mya (26). They are easily recognizable by their long and domed nasal bones and survived as relict populations in Patagonia until at least ~10 kya (27). The period spanning 13 and 8 kya underwent a massive episode of megafaunal extinction (known as the Late Pleistocene Megafauna Extinction) that erased other iconic species from America, including mammoths and the saber-toothed cat (28, 29). Equids in general were not spared; fossil and molecular evidence support their total extinction in the Americas at that time.

The survival of *Hippidions* and NWSL well into the Late Pleistocene, a time period compatible with DNA preservation, also explains why their fossil remains have been the focus of many ancient DNA studies and controversies. Surprisingly, initial phylogenetic work placed *Hippidion* closer to horses than to other living members of the *Equus* genus (30, 31), contradicting paleontological expectations of a minimum 10-million-year divergence. As Orlando and colleagues pointed out in their original work (30), this discrepancy could be solved by assuming that the fossil remains used for genetic analyses were misidentified as belonging to *Hippidion*. Out of the many *Equus* paleontological species that spread into South America during the last million years (32), *Equus (Amerhippus) neogenus* was contemporaneous and sympatric to *Hippidion*. Owing to the additional presence of convergent morphological characteristics, it was first suggested as a potential source of the misidentified material. This hypothesis was later demonstrated based on mitochondrial DNA evidence, despite morphological arguments suggesting otherwise (33, 34). Further sequencing of full mitochondrial genomes, from unambiguous *Hippidion* fossils, placed this lineage as external to the *Equus* crown and indicated divergence dates of 5.6–6.5 Mya (35). Similar work on northern NWSL horses initially supported earlier studies based on short, partial mitochondrial DNA fragments, which had erroneously placed the group as sister to horses (36). Only thanks to complete mitochondrial genomes from a larger number of specimens, as well as partial nuclear genomes, were NWSL

horses finally revealed as forming their own new genus, *Haringtonhippus*, currently consisting of a single species, *Haringtonhippus francisci* (37).

## EURASIA

*Equus* members had already spread across Beringia into the Old World prior to their extinction in northern America, similar to what hipparionines had done one million years earlier. However, little is known about the exact timing of such dispersal, except that it likely involved a minimum of two independent waves. The earliest took place ~2.6 Mya and is embodied within the so-called Elephant–*Equus* dispersal event, which brought a whole new range of herbivores into the Eurasian steppes. These included the so-called stenorine-like equids, which were the ancestors of all living equine species except the horse (38, 39). The second migration of *Equus* members out of America is thought to have occurred during the end-Villafranchian event, ~0.9 to 1 Mya, during an episode that represented a major turnover in the Eurasian faunal landscape (40). This second migration brought the true horses and their *Equus* ancestors (collectively, the caballines), which rapidly spread across Eurasia and largely replaced stenorine populations everywhere except Africa (41). Here again, ancient DNA helped refine the suite of evolutionary events that gave rise to the emergence, dispersal, and extinction of two species for which paleontological models could rely only on fragmentary morphological data (42).

The first such species is the so-called hydruntine, *Equus hydruntinus*, which was distributed from Western Europe to the Middle East and is first documented within the 350,000-year-old fossil record of the Les Trois Frères cave in southern France. Sparse fossil occurrences have suggested possible survival in scattered areas such as the Iberian Peninsula, southeastern Europe, Turkey, and Iran until the Iron Age (43). However, the species may have continued to exist until the Middle Ages, as it could correspond to the enigmatic zebro often described in historical documents to populate Iberia until the sixteenth century (44). The complex mixture of anatomical traits exhibited by *Equus hydruntinus*, combining ass-like and zebra-like characters (45), makes their identification particularly difficult in the absence of complete bone assemblages and its phylogenetic placement only suggestive. Mitochondrial genetic data, however, consistently supported its close affinities to wild Asian asses, as best represented by present-day onagers (*Equus hemionus*) (46, 47). More specifically, the genetic proximity to certain onager subspecies, such as Indian khurs or Iranian kulans, indicates that *E. hydruntinus*

was most likely a subpopulation of onagers that differentiated from southwest Asia into Europe (48). Ancient DNA analyses of an alleged zebro specimen, in turn, rejected genetic affinities to onagers and revealed a close relationship to donkeys, challenging the hypothesis of hydruntine survival beyond the Iron Age (47).

Another species for which our knowledge has much benefited from ancient DNA analyses is *Equus ovodovi*. This species belongs to a lineage of *Equus* that was first considered to have disappeared during the Middle Pleistocene (0.77–0.13 Mya) (49). However, complete mitochondrial genomes revealed their presence during and until the end of the Late Pleistocene in Russia (49) and northeastern China (50), respectively. Its exact phylogenetic placement remains, however, contentious, with some ancient DNA studies supporting closer affinities to zebras than to asses (50, 51) but others suggesting basal positioning within stenonines (36, 37, 47). Such phylogenetic instability may be due to the fact that the stenonine radiation was explosive and took place very rapidly, as suggested from the complete nuclear genomes of their living relatives (7). Of note, the only study exploiting partial nuclear information, from a limited number of markers, supported a basal placement within stenonines, as represented in **Figure 1b**. Yet, more data are necessary before the true phylogenetic affinities of *E. ovodovi* can be determined confidently.

## **AFRICA**

In contrast to other continents, the equid dispersal into and across Africa remains highly uncertain, as best illustrated by the evolution of the African wild ass. Its temporal origin, delineated by its split from Asian wild asses, is ambiguous. According to the distribution of genetic variation along individual genomes (53), African and Asian wild asses had parallel demographic trajectories until ~1.4 Mya but divergent dynamics thereafter (7, 53). Phylogenetic inference, however, estimated that African and Asian wild asses split approximately ~0.5 Mya based on genome-wide data (7), or ~2.3 Mya, leveraging variation at the mitochondrial DNA (mtDNA) control region (54). Its geographical origin remains unclear as well. One widely accepted view is that they emerged in Asia, with some scholars postulating subsequent entry into Africa ~0.5 Mya (55), although others have described ~1.8-Mya-old ass-like fossil remains at Aïn Hanech, Algeria (56). The other possibility assumes an African origin, followed by a more recent recolonization of Asia. The latter partly relies on the turnover that occurred in Eurasia by



the end-Villafranchian dispersal event, the consequences of which could have erased wild asses from Eurasia. An African origin is possible based on the earliest African fossil, which is stratigraphically dated to be 2.3 million years old (39), supporting the presence of ancient stenorhines in that continent well before the earliest genomic estimate of speciation into asses and zebras 2 Mya (7). This fossil was identified as belonging to an Olduvai zebra (*Equus oldowayensis*), which intriguingly resembled the extant Grevy's zebra (*Equus grevy*) more than the wild ass (57), despite the fact that its 2.33-million-year age largely precedes the split of both lineages.

Regardless of their deep origins, the subsequent evolutionary history of equids in Africa remains equally uncertain. The warm temperatures present across the vast majority of the continent provide less favorable environmental conditions for DNA preservation. Ancient DNA studies have been successful only on the most recent time periods, shedding some light on the evolutionary history of two extinct equines native from Africa. The first species is known as the giant Cape zebra (*Equus capensis*) and was described in the early 1900s from a partial mandible recovered near Cape Town. Weighing up to 400 kg and standing two meters high, *E. capensis* was potentially widespread across Africa in the past and survived at least until the Late Pleistocene (58, 59). Traditionally synonymized to be the same species as *E. oldowayensis* (59), *E. capensis* bore resemblance to the extant Grevy's zebra as well. For now, mitochondrial DNA data from the only giant Cape zebra specimen characterized contradicted such a view, supporting closer affinities to plain zebras (47). Future work including more specimens and/or nuclear genomic data is needed to justify a full revision of its current taxonomical definition.

In contrast, the placement of the extinct quagga as a subpopulation of plain zebras is well supported by both a complete genome sequence (7) and multiple partial mtDNA sequences (60). Comparison of the extinct quagga genome to restriction site-associated DNA sequencing SNP data from living plains zebras (61) confirmed quagga as an additional group alongside the nine extant pockets of molecular diversity described for that species. Of note, these subpopulations poorly mirror the conservation units defined according to morphological criteria, in line with earlier work based on microsatellite and mitochondrial control region data (62), which suggests nonoptimal management strategies in the conservation of plains zebras.

## DOMESTICATION

With horses and donkeys, the *Equus* genus provides an exceptional opportunity for documenting dissimilarities and parallelisms in two independent domestication processes. In contrast to other livestock that provide meat, milk, and wool, horses and donkeys have been used mostly for their transportation capacity, which considerably enhanced the mobility of pastoralist societies. Donkeys indeed were crucial as working and pack animals and revolutionized long-range overland trade between northern Africa and western Asia, and horse riding provided the first opportunity for fast, long-distance transportation. Given their far-reaching importance in human history, both species have received much scholarly attention, including from ancient DNA researchers in the last few years.

## **THE DONKEY**

It is still unknown how, where, and when donkeys were domesticated. The most commonly accepted view builds on genetic, ethnographic, and archaeological evidence. It assumes that the desertification of the Sahara region that followed the African Humid Period during the mid-Holocene triggered the domestication of the donkey as nomad cattle herders sought for pack animals better adapted to harsh, arid environments and able to travel long distances and carry heavy loads (63). Wild asses from northeastern Africa satisfied all such requirements. Alternative models have suggested the Levant, present-day Yemen, and the region of southwestern Asia as potential domestication centers (64, 65). Regardless of the exact location, there is little genetic doubt that the origins of the domestic donkey lie in African wild ass ancestors (*Equus africanus asinus*) (66), disregarding hemiones and kiangs (*Equus kiang*) as possible parental species (67).

Debate about which among the three subpopulations of African wild asses known to have existed in north and east Africa (68) gave rise to the domestic donkey remains contentious. These include the Atlas (Algerian) wild ass, which is depicted in Roman mosaics with striped legs and a shoulder cross and likely ranged from present-day Morocco and possibly to Tunisia and beyond (69). Characterized by longer ears and absence of stripes on its legs, the Nubian wild ass also represents a candidate source of ancestry for donkey domestication. Similar to what happened to their Atlas cousins, Nubian wild asses have been cornered into the Nubian Desert of northeastern Sudan owing to human encroachment and habitat competition with domestic stocks. None has been seen in the wild since the 1970s, although a remnant population, possibly feral or

admixed with domesticates, is claimed to survive in the Gabal Elba nature reserve, currently under Egyptian sovereignty. The Nubian wild ass is thus considered critically endangered and most likely extinct (70). Finally, the Somali wild ass from Kenya, Ethiopia, and the Afars is the last of the three possible candidate sources of domestication. Their conservation status is also critically endangered, as not even 200 individuals may still exist in the wild (71).

The earliest undisputed archaeological evidence of donkey domestication is provided by 10 adult skeletons excavated from the pharaonic complex of Abydos, in Egypt, and dated to 5 kya (72). The geographic location and specific morphometrical similarities posit such remains as Nubian asses at early domestication stages, which had not yet developed the full morpho-anatomical features of domestic animals. The presence of clear arthropathies leaves no doubt that these animals were used for heavy load bearing and were, hence, domesticated.

Together with the likely-long process needed to fully evolve domestic anatomical features, that multiple possible wild ancestral population sources co-existed until recently has further limited current models of donkey domestication. Yet, genetic studies have attempted to explore patterns of mitochondrial genetic variation to gain insights into this process. These revealed the presence of two major mtDNA haplogroups within living donkeys, hereafter referred to as Clades 1 and 2 (66). With an estimated divergence date of at least 100 kya, these two maternal lineages have been interpreted to represent two populations of wild asses contributing to contemporary domestic matriline (73). Clade 1 shows the closest affinity to the Nubian wild ass and includes multiple haplotypes segregating at small-to-moderate frequencies. This haplotype structure lent support to the possible restocking from wild jenny progenitors into the domestic pool and to speculation that donkey herders may have maintained crossbreeding with wild specimens in an attempt to augment the strength and hardiness of their animals. In contrast, Clade 2 members group together within two major haplotypes (57). They were first assumed to have originated from the Somali wild ass, but the genetic distance calculated at the mitochondrial control region later ruled out this hypothesis (73). Alternative sources of Clade 2 ancestry involve now-extinct relatives of the Somali wild ass, an unknown population from Yemen, or even the Atlas wild ass (56), although the latter option was originally considered as least likely (68).

Demographic variations within Clade 1 and Clade 2 have been reconstructed using Bayesian skyline plots based on an extensive worldwide panel of donkey mitochondrial control regions

(74). This supported distinctive evolutionary trajectories for Clade 1 and 2 members, in which the effective population of Clade 1 almost doubled ~8 kya, well prior to the earliest archaeological evidence of donkey domestication in Egypt. Whether this reflects a natural demographic response to environmental change, the signature of a surprisingly young domestication success, or calibration issues of the mitochondrial DNA clock, remains unknown. Mitochondrial variation also revealed some level of geographic structure within Africa, in which donkeys native from the Gulf of Guinea up to Senegal generally clustered within Clade 1, whereas those sampled from the east coast of Africa and Maghreb, but also from Iberia and Italy, belonged to Clade 2. The rest of Eurasia shows balanced proportions between both clades, possibly reflecting that both maternal lineages spread out of Africa via the Arabian Peninsula once domesticated. The Arabian Peninsula could in fact represent a hybridization zone, owing to the high levels of genetic variability measured there, but it may also have served as another domestication center, as suggested by the substantial amounts of private diversity found at 15 autosomal microsatellites among the populations of Yemen and Oman (75).

Combined with archaeological data, genetic diversity points toward two potential and non-mutually exclusive centers of donkey domestication. The first is known as the Egyptian hypothesis and places domestication within the Nile valley 5–6 kya. The second is known as the pastoralist hypothesis and advocates for an earlier domestication in northeastern Africa, following Saharan hyper-aridification during the mid-Holocene. This hypothesis is gaining popularity and includes the African Horn and the Arabian Peninsula, which were historically connected by the Bab-el-Mandeb Strait, as possible domestication centers. In favor of this hypothesis are the greatest levels of nuclear private diversity found in this region (75), the central position of Ethiopian mtDNA haplotypes (76), and the presence of the largest collection of archaeological donkeys (or wild asses) known to date in present-day Yemen, despite its archaeological context and morphological description are inconclusive regarding their possible domestication status (77). To conclude, our understanding of the process underlying donkey domestication is very preliminary, and further archaeological work and characterization of the genomic diversity present within both modern and ancient donkeys are necessary.

## **THE HORSE**

The end of the Ice Age, geologically dated to 11.5 kya, marked the beginning of a substantially

warmer epoch known as the Holocene (78). Northern ice caps and glaciers gradually melted, and elevated temperatures increased the forest density in Eurasia, fragmenting the preferred steppe-like habitats of horses (79). Whether European wild horses then adapted to postglacial forests has been the subject of intense scientific debate, with ecological niche modeling (80) reporting quite dissimilar habitat preferences for horses in Asia and Europe and thus supporting theories of local adaptation (81). The signatures of selection for a mutation within the *Agouti* gene, tightly associated with black coat coloration, have also lent support to adaptive theories in which dark camouflage could confer greater survival rates within forests essentially made of willow trees (80, 82). The rarity of horse remains, illustrated by both declining demographic trajectories reconstructed from patterns of genetic variation (12, 83) and far fewer fossil occurrences for horses than for other forest-adapted mammals, such as deer and wild boars, argues against horse populations thriving in forest environments at the time (84). Either way, the scattered population landscape found in Eurasia during the first half of the Holocene constrains the potential geographical centers of domestication to a limited number of regions providing suitable ecological conditions for horse survival, such as the Iberian Peninsula, Anatolia, and the Pontic and Central Asian steppes (85).

### **The Central Asian Steppes: Botai**

The earliest archaeological evidence of horse domestication points to Central Asia ~5.5 kya (86). Here, human settlements associated with the so-called Botai culture included more than one hundred pit houses and vast faunal assemblages consisting almost exclusively of horses. Evidence of harnessing, milking, and corralling, all suggestive of human-controlled management, were identified (83). Botai horses were thus considered to be the most likely ancestors of all domestic horses living on the planet today. Sequencing of 20 of their genomes revealed, however, a completely different picture, in which neither Botai horses nor five additional horses excavated from the ~5,000-year-old archaeological site of Borly, located approximately 450 miles away from Botai, showed significant ancestry with modern domestic horses (83). They were found to be the direct ancestors of Przewalski's horses, which were previously considered the last wild horses on the planet. After Borly, the evolutionary trajectory of this lineage remains enigmatic during almost five millennia, until Przewalski's horses were rediscovered only 150 years ago as a free-range population and were assumed to be wild. Przewalski's horses almost became extinct during the twentieth century and survived from a

captive stock of 12–15 founders only thanks to huge conservation and reintroduction efforts. Ancient genomics helped clarify their true evolutionary origins and revealed that they represent the only living legacy of a now-extinct human population that first domesticated the horse ~5.5 kya (87). Paleogenomics has also shown that the genome of modern domestic horses descended from a distinct genetic lineage (hereafter referred to as DOM2), which diverged from that leading to Botai horses ~35–45 kya (83, 88). Where and when this second lineage of horses was domesticated are currently unknown.

### **The Iberian Peninsula**

Iberia was long proposed as one possible domestication center for horses, mostly building on the suitable paleoclimatic conditions and extensive biological and cultural archaeological material associated to horses in the region, including cave paintings (89). In the last 15 years, genetic studies have aimed to clarify the role that Iberia played during horse domestication. First, the present genetic diversity was found to be especially elevated in native Iberian DOM2 populations for 12 microsatellite loci (85). Two mitochondrial haplogroups, C and D1, were also found to be exclusive to modern breeds native from Iberia, suggesting a local domestication event (90, 91). Ancient DNA data, however, dismissed haplogroup D1 as a domestication marker, because it seems to have spread in Iberia only during Medieval times (90).

The genome sequence of ancient Iberian horses, excavated from Bell Beaker settlements, revealed that a yet-unknown and extremely divergent lineage existed in the region during the third millennium BCE (this now-extinct lineage, native from Iberia, is hereafter referred to as IBE; **Figure 2a**) (92). Even though it remained undetected on the basis of morphological variation, IBE exhibited elevated levels of genetic differentiation, which supports a history of isolation within Iberia during the Last Glacial Maximum. Their genomes were found to additionally carry ancestry from an even more divergent and yet-uncharacterized ghost lineage, including the Y chromosomes of the only two IBE stallions ever characterized. The geographic and temporal distributions of this ghost lineage remain unknown and invite further research. For now, the exclusive levels of genetic drift found in the four sequenced IBE genomes reject IBE as the ancestors of modern domesticates. However, according to  $f_G$ -statistics (93), they contributed ~1–12% ancestry to various modern domestic breeds and ancient Iberian domestic specimens dated to the Iron Age (92). Whether the IBE ancestry entered into the DOM2 lineage following a natural admixture event, alongside the Bell Beaker phenomenon expanding from Iberia across

Western Europe, or under the influence of other important cultural and/or civilization changes remains to be investigated.

**Figure 2** Horse domestication. (a) Polymorphisms within the *Equus* mitochondrial control region, as recovered by Lira and colleagues (88), genotyped here for a few ancient genomes for comparison purposes. The yellow square highlights shared mutations between IBE[\*\*AU: Spell out?\*\*) and Portuguese Lusitano horses within mitochondrial haplogroup C, absent in other breeds or ancient populations. (b) Genome-wide selection scan conducted by Fages and colleagues (90) along chromosome ECA11 (90). The bottom subpanel zooms into the *HOXB* gene cluster, showing the strongest selection signal between *HOXB13* and *TLL6*. Mutations in the latter gene are associated with the Joubert syndrome causing alterations in locomotory coordination.

### Other Potential Genomic Origins

As patterns of genomic variation has revealed that Botai and IBE horses do not represent the most dominant ancestry source within modern domestic horses, further research turned to other existing lineages as possible candidates. One such candidate pertains to a lineage that diverged from DOM2 ~110 kya, and that was first identified in Late Pleistocene wild populations of the subarctic Taymir Peninsula in Siberia (94). This was tentatively assigned to *Equus lenensis*, a species described based on the morphological analysis of a foal specimen that remained mummified in the Siberian permafrost (95). Additional genome sequences demonstrated that the lineage survived well into the Holocene in northeast Siberia and at least until 5.1 kya (94, 96), a time contemporary to Botai settlements (86). Although D-statistics (97) point to an excess of shared drift between *E. lenensis* and DOM2 relative to Przewalski's horses (94), this excess is lost when replacing Przewalski's horses by their Botai ancestors (83). Therefore, the most parsimonious scenario is that *E. lenensis* did not contribute to DOM2 but that the Botai–Borly lineage experienced a dilution of *E. lenensis* alleles (relative to DOM2) until it was rediscovered in the form of Przewalski's horses. This dilution was possibly due to introgression from an even more divergent ghost lineage of horses that remains uncharacterized for now.

### The Historical Construction of Modern Horses

Regardless of the evolutionary roots of the DOM2 ancestry, which remain to be discovered, the domestic horse that we know today was not engineered in an overnight process but gradually shaped following millennia of breeding. Current evidence supports nonsymmetric roles for mares

and stallions in this process, as males appear overrepresented in the archaeological record since the early Bronze Age, some ~3.9 kya (98). The considerable mitochondrial diversity found in modern breeds also indicates that multiple matrilineages were incorporated into the domestic pool in the course of history (99), whereas their extremely reduced Y-chromosomal diversity confirms that only a handful of influential Oriental sire lines dominated domestic reproduction during the last 850–1,300 years, most especially during the last three centuries (100). This Oriental influence is so pronounced that only those populations that were established earlier, outside the westernized industry of horse breeding, were found to carry more divergent Y-chromosomal haplotypes. This is, for example, the case in Yakutian horses, a breed of Mongolian origin and adapted to survive the coldest winters in the northern hemisphere (96). Scandinavian breeds, such as Icelandic and Shetland horses, were also found to retain additional sire lines (101).

In addition to gender-biased preferences, other criteria contributed to shape the modern horse genome. Past breeders have preferentially selected certain horse phenotypes, a practice that distorted local gene genealogies and left molecular footprints of selection along their genomes. The most common selection targets found in the genome of modern horses have been reviewed extensively (99, 102, 103). We thus decided to summarize the information gained by adding the temporal dimension provided by ancient DNA time series. In their seminal study, Schubert and colleagues (94) compared the genomes of 2 wild *E. lenensis* and 6 domestic breeds to identify 125 gene candidates for selection. Their respective functions in behavior and nervous, circulatory, and musculoskeletal systems suggested a possible role during domestication, although adaptation to natural environments instead of human exposure may also underpin the signatures detected, especially as the two lineages diverged ~110 kya. The development of LSD (stands for Levels of exclusively Shared Differences), a new statistical method allowing the detection of ancient sweeps on any predefined branch of a population tree (104), helped further narrow the list of gene candidates that have been selected after the divergence of the DOM2 and Przewalski's lineages, potentially in relation with domestication (105). Interestingly, “abnormal associative learning” was one of the Gene Ontology categories found to be significantly overrepresented, suggesting that past breeders may have rewarded enhanced learning capabilities during early domestication stages. Enrichment for neural crest was also significant, in line with the neural-crest theory positing an important role of these stem cells in the development of the traits commonly found across domestic animals and collectively forming the so-called



domestication syndrome (106; though see 107 for a critical overview).

*KITLG*, which is associated with neural crest development but also with the lack of wild-type Dun pelage, provides one example of such a gene under selection (108). Early selection at *KITLG* concurs with the explosion of coat-coloration patterns detected at other loci (109), which is traditionally considered a hallmark of domestication (110, 111). Besides coat-coloration shifts, locomotion and morpho-anatomy appear to have been recurrent selection targets during domestication (92). For example, variants involved in racing performance (112) or gaitedness (113) show parallel trajectories over time, first remaining at low-to-moderate frequencies until becoming increasingly popular within the last millennium. Selection signatures at key loci for limb and morphological development were detected in the horses of Iron Age Scythian nomads (105), but also in more recent samples, coinciding with the growing influence of Oriental sire lines (92). Interestingly, the latter involves two independent *HOX* gene clusters, one of which encompasses the *HOXB* complex (Figure 2b), which has been also found as selection candidate on different breeds, such as the German warmbloods and breeds native to the Balkans, Austria, and Hungary (115).

## **BREED CONSERVATION**

With the industrial revolution, the economies of developed countries have turned increasingly toward higher productivity through task specialization. This also applies to domestic animals, which became increasingly specialized and bred in independent bloodlines aimed to serve specific tasks. Additionally, the development of the first motor vehicles provided an appealing alternative to horses and donkeys for transportation, and many local equine populations suffered dramatic declines during the twentieth century. Combined, the development of a breed industry and the overall deprecation of horses in our modern economies have had a considerably negative impact on their effective population sizes, and the conservation of livestock breeds has become an increasingly important political concern.

In this context, the UN Food and Agriculture Organization (FAO) launched the Domestic Animal Diversity Information System (DAD-IS; <http://www.fao.org/dad-is/data/en/>), a collaborative database defining conservation priorities according to the number of reproductive males and females available in each individual breed (116). Although still perfectible, this tool and FAOSTAT reveal that European countries have lost 80% of their donkey stocks in the last

50 years, a proportion reaching up to 97% in countries such as Greece and Italy ([117](#)). The situation does not appear to be much better for horses, with 200 breeds listed as threatened in DAD-IS and 80 listed as already extinct.

From a genetic standpoint, considering the sole number of reproductive males and females may not be sufficient to assess the risk of breed extinction. This is because new mutations are rare, and it will take hundreds of generations before significant amounts of variability can be recovered from de novo mutations, even if the census population grows rapidly ([Figure 3a](#)). Direct measures of genetic diversity would, thus, certainly help define robust conservation strategies. In donkeys, such measures are based mostly on mitochondrial markers and autosomal microsatellites, but these data sets and their cross-compatibility remain limited. Although China seems to retain considerable ancestral variation ([122](#)), the number and location of other possible hot and cold spots of genetic diversity remain to be further assessed. More genetic studies have been carried out for horses, including at the genome scale. Full-genome ancient DNA time series have revealed a ~16% sudden drop in heterozygosity during the last 200 years ([90](#)). This drop was mirrored by a 4% burst in mutational loads, a measure proportional to the total number of deleterious variants present in a given genome at the homozygous state ([81](#), [90](#), [103](#)). Both the diversity drop and the increase in mutation loads are believed to result from the combined effects of population fragmentation in studs and overall demographic declines, which hamper the efficacy of negative selection to filter out deleterious mutations ([123](#)). These effects are particularly prevalent in coldblooded working horses, and to a lower extent in hotblooded and especially warmblood horses, as would be expected if mutational loads were driven by the deprecation of working breeds and close stud practices ([124](#)).

[Figure 3](#) Genetic load and conservation. (a) Representation of the relatedness, mean heterozygosity, fitness [as calculated by SLiM 3 ([131](#))], and mutational loads for a population experiencing a demographic bottleneck, including the recovery phase. Forward simulations of 10-Mb DNA fragments were conducted considering equal proportions of neutral and deleterious mutations (approximately mimicking protein-coding sites). For harmful alleles, selection coefficients were drawn from an exponential with rate parameter equal to 0.01, whereas the dominance coefficient was fixed to 0.2. Note that mutational loads consider only homozygous sites, while fitness estimates account for partially recessive variants. (b) Per-site mutational loads were calculated using samples and methodology as in References [103](#) and [124](#) from highly constrained and confidently called homozygous positions.

Within such timescales, mutational loads are most likely explained by the interplay between inbreeding fitness depression and genetic purging. The first mechanism accounts for the deleterious effects of recessive mutations when phenotypically expressed at homozygous sites resulting from inbreeding (124). Once recessive alleles are expressed phenotypically, in homozygous state, the effects of natural selection can however begin, and harmful alleles can be purged, thus tending to alleviate mutational loads. Forward simulations presented in **Figure 3a** show how mutational loads follow inbreeding increments right after a population decline but also that their trajectories become decoupled after a few hundred generations, as runs of homozygosity can be enriched for deleterious mutations, but also depleted once purging is effective, regardless of inbreeding levels.

The explanation above serves to illustrate the controversy associated with the impact of purging in Thoroughbreds, one of the most influential and controlled breeds on the planet. Their mean genome fraction within runs of homozygosity can be estimated to have gone from approximately 21.8% to 22.75% in 15 years. This has been misinterpreted as incompatible with substantial purging (125), despite pedigree-based measures of purging correlated positively with racing performance (126) but not foaling rate (127). The latter findings confirmed that pressure to eliminate deleterious variants primordially focused on adult athletic achievements, as expected for this breed. This expectation is confirmed with direct measurements of mutational loads, which increased by an average ~4% genome wide but decreased by 12.5% at firmly constrained zero-fold degenerate positions, i.e., those positions always causing amino-acid replacements (124) (**Figure 3b**). The load decrease at these positions reflects efficient negative selection, partially purging strongly deleterious mutations. Genome-guided purging programs, therefore, might constitute an efficient alternative toward more sustainable breeding strategies, even for the most industrial horse studs.

## **CONCLUDING REMARKS AND FUTURE WORK**

This review summarizes current knowledge about equine evolution based on paleontological and genomic data. The considerable variation in the equine karyotype, which ranges from  $2n = 32$  to  $2n = 66$  (7), is not addressed here but will likely constitute an important research topic in the near future, as solutions for assembling genomes de novo become increasingly affordable. In the meantime, multiple other areas require further research. For example, although our

understanding of equine evolutionary history has made much progress concerning the Late Pleistocene and Holocene periods, earlier periods have been largely overlooked. DNA decay will limit future genetic investigations beyond 0.5–1 Mya, but with time records close to 2.0 Mya ([128](#), [129](#)), paleoproteomic approaches applied to dental enamel, an extremely dense and abundant material in the fossil record, bear enormous potential.

The donkey domestication process represents another area in which major discoveries could be made, especially given the paucity of current genetic data sets, mostly limited to a few autosomal microsatellites ([73](#)), mtDNA ([66](#)), and the Y chromosome ([130](#)). The recent release of a new de novo assembly for the species, including scaffolds of subchromosomal size ([53](#)), will be instrumental for mapping patterns of genetic variation through space and time and identifying convergent genetic pathways that may have been selected independently during both horse and donkey domestication. However, DNA preservation in Africa, southwest Asia, and the Arabian Peninsula may limit our capacity to retrieve extensive whole-genome time series for this species, given current technologies. Target-enrichment approaches aimed at recovering sequence variation at preselected informative neutral and functional loci will thus likely represent key technological developments to help reconstruct the process of donkey domestication.

Despite recent progress, many aspects of the horse domestication process remain unclear, not least of which is mapping of the origins of the major source of DOM2 ancestry. In this regard, the Pontic–Caspian steppes and their connection with the Central Asian steppes, Eastern Europe and Anatolia remain important regions with extensive archaeological records suggestive of long-term equestrian traditions and intensified human-horse relationships. Following these hypotheses, a recent paleogenetic study revealed that Anatolia witnessed an apparent extensive mare turnover 4.2 kya ([132](#)). This is interpreted to result from the arrival of horses that replaced local, wild populations and that were domesticated outside the favorable climate and abundant biodiversity present in the Fertile Crescent 8-10 kya, where some of the most important livestock species were domesticated, including cattle, pigs, goats and sheep ([133-134](#)). Establishing whether full or partial replacement was involved will require further confirmation at the genomic level. However, this illustrates how humans commenced to manage equid populations not only later than other species, but also outside major geographic centers of domestication, and would be echoed in donkeys as well, as these were most likely domesticated in the Africa horn. This may reflect domestication processes motivated by different purposes, tightly associated with

increasing needs for enhanced mobility. Further emphasizing the peculiarities of the horse domestication process, both sexual markers show disparate levels of diversity, unparalleled in other livestock species, and involving extensive mtDNA haplotype variation coalescing ~150 kya, as well as the extremely limited Y-chromosomal diversity established during the last centuries (135). Precisely the incorporation of a narrow fraction of reproductive stallions in close-studs and the industrialization of modern human societies constitute important threats for the long-term viability of multiple horse breeds. Facilitating access to genomic technologies, especially to nongovernmental associations and in least-developed regions, appears of utmost importance to help identify conservation priorities and design the best possible conservation strategies.

## **DISCLOSURE STATEMENT**

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

## **LITERATURE CITED**

1. Moehlman PD, ed. 2002. *Equids: Zebras, Asses and Horses: Status Survey and Conservation Action Plan*. Gland, Switz.: Int. Union Conserv. Nat.
2. Marsh OC. 1879. Polydactyl horses, recent and extinct. *Am. J. Sci. Ser. 3* 17(102):499–505
3. MacFadden BJ. 1994. *Fossil Horses: Systematics, Paleobiology, and Evolution of the Family Equidae*. Cambridge, UK: Cambridge Univ. Press
4. Cucchi T, Mohaseb A, Peigné S, Debue K, Orlando L, Mashkour M. 2017. Detecting taxonomic and phylogenetic signals in equid cheek teeth: towards new palaeontological and archaeological proxies. *R. Soc. Open Sci.* 4(4):160997
5. Vershinina AO, Kapp JD, Baryshnikov GF, Shapiro B. 2019. The case of an arctic wild ass highlights the utility of ancient DNA for validating problematic identifications in museum collections. *Mol. Ecol. Resourc.* In press
6. Higuchi R, Bowman B, Freiberger M, Ryder OA, Wilson AC. 1984. DNA sequences from the quagga, an extinct member of the horse family. *Nature* 312(5991):282–84
7. Jónsson H, Schubert M, Seguin-Orlando A, Ginolhac A, Petersen L, et al. 2014. Speciation

- with gene flow in equids despite extensive chromosomal plasticity. *PNAS* 111(52):18655–60
8. Sawyer S, Krause J, Guschanski K, Savolainen V, Pääbo S. 2012. Temporal patterns of nucleotide misincorporations and DNA fragmentation in ancient DNA. *PLOS ONE* 7(3):e34131
  9. Hofreiter M, Serre D, Poinar HN, Kuch M, Pääbo S. 2001. Ancient DNA. *Nat. Rev. Genet.* 2(5):353–59
  10. Leonardi M, Librado P, Der Sarkissian C, Schubert M, Alfarhan AH, et al. 2017. Evolutionary patterns and processes: lessons from ancient DNA. *Syst. Biol.* 66(1):e1–29
  11. Orlando L, Gilbert MTP, Willerslev E. 2015. Reconstructing ancient genomes and epigenomes. *Nat. Rev. Genet.* 16(7):395–408
  12. Orlando L, Ginolhac A, Zhang G, Froese D, Albrechtsen A, et al. 2013. Recalibrating *Equus* evolution using the genome sequence of an early Middle Pleistocene horse. *Nature* 499(7456):74–78
  13. Froehlich DJ. 2002. Quo vadis eohippus? The systematics and taxonomy of the early Eocene equids (Perissodactyla). *Zool. J. Linn. Soc.* 134(2):141–256
  14. MacFadden BJ, Hulbert RC. 1988. Explosive speciation at the base of the adaptive radiation of Miocene grazing horses. *Nature* 336(6198):466–68
  15. Muhlbachler MC, Rivals F, Solounias N, Semprebon GM. 2011. Dietary change and evolution of horses in North America. *Science* 331(6021):1178–81
  16. Semprebon GM, Rivals F, Solounias N, Hulbert RC. 2016. Paleodietary reconstruction of fossil horses from the Eocene through Pleistocene of North America. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 442:110–27
  17. Strömberg CAE. 2006. Evolution of hypsodonty in equids: testing a hypothesis of adaptation. *Paleobiology* 32(2):236–58
  18. Janis CM, Bernor RL. 2019. The evolution of equid monodactyly: a review including a new hypothesis. *Front. Ecol. Evol.* 7:119
  19. Cantalapiedra JL, Prado JL, Fernández MH, Alberdi MT. 2017. Decoupled ecomorphological evolution and diversification in Neogene-Quaternary horses. *Science* 355(6325):627–30
  20. Parker AK, McHorse BK, Pierce SE. 2018. Niche modeling reveals lack of broad-scale habitat partitioning in extinct horses of North America. *Palaeogeogr. Palaeoclimatol.*

*Palaeoecol.* 511:103–18

21. Hulbert RC. 1993. Taxonomic evolution in North American Neogene horses (subfamily Equinae): the rise and fall of an adaptive radiation. *Paleobiology* 19(2):216–34
22. Macfadden BJ. 1997. Pleistocene horses from Tarija, Bolivia, and validity of the genus †*Onohippidium* (Mammalia: Equidae). *J. Vertebr. Paleontol.* 17(1):199–218
23. Flower BP, Kennett JP. 1994. The middle Miocene climatic transition: East Antarctic ice sheet development, deep ocean circulation and global carbon cycling. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 108(3):537–55
24. MacFadden BJ. 2005. Fossil horses—evidence for evolution. *Science* 307(5716):1728–30
25. Lundelius EL, Stevens MS. 1970. *Equus francisci* Hay, a small stilt-legged horse, Middle pleistocene of Texas. *J. Paleontol.* 44(1):148–53
26. O’Dea A, Lessios HA, Coates AG, Eytan RI, Restrepo-Moreno SA, et al. 2016. Formation of the isthmus of Panama. *Sci. Adv.* 2(8):e1600883
27. Prado JL, Martinez-Maza C, Alberdi MT. 2015. Megafauna extinction in South America: a new chronology for the Argentine Pampas. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 425:41–49
28. Lorenzen ED, Nogués-Bravo D, Orlando L, Weinstock J, Binladen J, et al. 2011. Species-specific responses of Late Quaternary megafauna to climate and humans. *Nature* 479(7373):359–64
29. Barnosky AD, Koch PL, Feranec RS, Wing SL, Shabel AB. 2004. Assessing the causes of late Pleistocene extinctions on the continents. *Science* 306(5693):70–75
30. Orlando L, Eisenmann V, Reynier F, Sondaar P, Hänni C. 2003. Morphological convergence in *Hippidion* and *Equus* (*Amerhippus*) South American equids elucidated by ancient DNA analysis. *J. Mol. Evol.* 57(1):S29–40
31. Weinstock J, Willerslev E, Sher A, Tong W, Ho SYW, et al. 2005. Evolution, systematics, and phylogeography of Pleistocene horses in the New World: a molecular perspective. *PLOS Biol.* 3(8):e241
32. MacFadden BJ. 2013. Dispersal of Pleistocene *Equus* (family Equidae) into South America and calibration of GABI 3 based on evidence from Tarija, Bolivia. *PLOS ONE* 8(3):e59277
33. Alberdi MT, Prado JL, Prieto A. 2005. Considerations on the paper “Morphological convergence in *Hippidion* and *Equus* (*Amerhippus*) South American equids elucidated by

- ancient DNA analysis,” by Ludovic Orlando, Véra Eisenmann, Frédéric Reynier, Paul Sondaar, Catherine Hänni. *J. Mol. Evol.* 61(1):145–47
34. Orlando L, Male D, Alberdi MT, Prado JL, Prieto A, et al. 2008. Ancient DNA clarifies the evolutionary history of American late Pleistocene equids. *J. Mol. Evol.* 66(5):533–38
  35. Der Sarkissian C, Vilstrup JT, Schubert M, Seguin-Orlando A, Eme D, et al. 2015. Mitochondrial genomes reveal the extinct *Hippidion* as an outgroup to all living equids. *Biol. Lett.* 11(3):20141058
  36. Vilstrup JT, Seguin-Orlando A, Stiller M, Ginolhac A, Raghavan M, et al. 2013. Mitochondrial phylogenomics of modern and ancient equids. *PLOS ONE* 8(2):e55950
  37. Heintzman PD, Zazula GD, MacPhee RD, Scott E, Cahill JA, et al. 2017. A new genus of horse from Pleistocene North America. *eLife* 6:e29944
  38. Lindsay EH, Opdyke ND, Johnson NM. 1980. Pliocene dispersal of the horse *Equus* and late Cenozoic mammalian dispersal events. *Nature* 287(5778):135–38
  39. Rook L, Bernor RL, Avilla LS, Cirilli O, Flynn L, et al. 2019. Mammal biochronology (land mammal ages) around the world from late Miocene to middle Pleistocene and major events in horse evolutionary history. *Front. Ecol. Evol.* 7:278
  40. Azzaroli A. 1983. Quaternary mammals and the “end-Villafranchian” dispersal event—a turning point in the history of Eurasia. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 44(1):117–39
  41. Forsten A. 1988. Middle Pleistocene replacement of stenorhinid horses by caballoid horses—ecological implications. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 65(1):23–33
  42. Forsten A. 1988. The small caballoid horse of the upper Pleistocene and Holocene. *J. Anim. Breed. Genet.* 105(1–6):161–76
  43. Cress JJ, Turvey ST. 2014. Holocene extinction dynamics of *Equus hydruntinus*, a late-surviving European megafaunal mammal. *Quat. Sci. Rev.* 91:16–29
  44. Nores C, Muñoz AM, Rodríguez LL, Bennett EA, Geigl E-M. 2015. The Iberian zebro: What kind of a beast was it? *Anthropozoologica* 50(1):21–32
  45. Burke A, Eisenmann V, Ambler GK. 2003. The systematic position of *Equus hydruntinus*, an extinct species of Pleistocene equid. *Quat. Res.* 59(3):459–69
  46. Orlando L, Mashkour M, Burke A, Douady CJ, Eisenmann V, Hänni C. 2006. Geographic distribution of an extinct equid (*Equus hydruntinus*: Mammalia, Equidae) revealed by



- morphological and genetical analyses of fossils. *Mol. Ecol.* 15(8):2083–93
47. Orlando L, Metcalf JL, Alberdi MT, Telles-Antunes M, Bonjean D, et al. 2009. Revising the recent evolutionary history of equids using ancient DNA. *PNAS* 106(51):21754–59
  48. Bennett EA, Champlot S, Peters J, Arbuckle BS, Guimaraes S, et al. 2017. Taming the late Quaternary phylogeography of the Eurasiatic wild ass through ancient and modern DNA. *PLOS ONE* 12(4):e0174216
  49. Eisenmann V, Sergej V. 2011. Unexpected finding of a new *Equus* species (Mammalia, Perissodactyla) belonging to a supposedly extinct subgenus in late Pleistocene deposits of Khakassia (Southwestern Siberia). *Geodiversitas* 33(3):519–30
  50. Yuan J-X, Hou X-D, Barlow A, Preick M, Taron UH, et al. 2019. Molecular identification of late and terminal Pleistocene *Equus ovodovi* from northeastern China. *PLOS ONE* 14(5):e0216883
  51. Druzhkova AS, Makunin AI, Vorobieva NV, Vasiliev SK, Ovodov ND, et al. 2017. Complete mitochondrial genome of an extinct *Equus (Sussemionus) ovodovi* specimen from Denisova cave (Altai, Russia). *Mitochondrial DNA B* 2(1):79–81
  52. Li H, Durbin R. 2011. Inference of human population history from individual whole-genome sequences. *Nature* 475(7357):493–96
  53. Renaud G, Petersen B, Seguin-Orlando A, Bertelsen MF, Waller A, et al. 2018. Improved de novo genomic assembly for the domestic donkey. *Sci. Adv.* 4(4):eaaq0392
  54. Rosenbom S, Costa V, Chen S, Khalatbari L, Yusefi GH, et al. 2015. Reassessing the evolutionary history of ass-like equids: insights from patterns of genetic variation in contemporary extant populations. *Mol. Phylogenet. Evol.* 85:88–96
  55. Prothero DR, Schoch RM. 2002. *Horns, Tusks, and Flippers: The Evolution of Hoofed Mammals*. Baltimore, MD: Johns Hopkins Univ. Press.
  56. Sam Y. 2020. African origins of modern asses as seen from paleontology and DNA: What about the Atlas wild ass? *Geobios* 58:73–84
  57. Bernor RL, Cirilli O, Jukar AM, Potts R, Buskianidze M, Rook L. 2019. Evolution of early *Equus* in Italy, Georgia, the Indian Subcontinent, East Africa, and the origins of African zebras. *Front. Ecol. Evol.* 7:166
  58. Churcher CS. 2006. Distribution and history of the Cape zebra (*Equus capensis*) in the Quaternary of Africa. *Trans. R. Soc. South Afr.* 61(2):89–95

59. Churcher CS. 2014. A vacant niche? The curious distributions of African Perissodactyla. *Trans. R. Soc. South Afr.* 69(1):1–8
60. Leonard JA, Rohland N, Glaberman S, Fleischer RC, Caccone A, Hofreiter M. 2005. A rapid loss of stripes: the evolutionary history of the extinct quagga. *Biol. Lett.* 1(3):291–95
61. Pedersen C-ET, Albrechtsen A, Etter PD, Johnson EA, Orlando L, et al. 2018. A southern African origin and cryptic structure in the highly mobile plains zebra. *Nat. Ecol. Evol.* 2(3):491–98
62. Lorenzen ED, Arctander P, Siegismund HR. 2008. High variation and very low differentiation in wide ranging plains zebra (*Equus quagga*): insights from mtDNA and microsatellites. *Mol. Ecol.* 17(12):2812–24
63. Shanahan TM, McKay NP, Hughen KA, Overpeck JT, Otto-Bliesner B, et al. 2015. The time-transgressive termination of the African Humid Period. *Nat. Geosci.* 8(2):140–44
64. Clutton-Brock J. 1992. *Horse Power: A History of the Horse and Donkey in Human Societies*. Cambridge, MA: Harvard Univ. Press. 1st ed.
65. Vilà C, Leonard J, Beja-Pereira A. 2006. Genetic documentation of horse and donkey domestication. In *Documenting Domestication: New Genetic and Archaeological Paradigms*, ed. MA Zeder, D Bradley, E Emshwiller, BD Smith, pp. 342–53. Berkeley: Univ. Calif. Press
66. Beja-Pereira A, England PR, Ferrand N, Jordan S, Bakhiet AO, et al. 2004. African origins of the domestic donkey. *Science* 304(5678):1781–81
67. Han L, Zhu S, Ning C, Cai D, Wang K, et al. 2014. Ancient DNA provides new insight into the maternal lineages and domestication of Chinese donkeys. *BMC Evol. Biol.* 14(1):246
68. Van Bemmelen ACV. 1972. Some remarks on the African wild ass. *Zool. Meded.* 47(21):261–72
69. Moehlman P, Kebede F, Yohannes H. 2014. *African wild ass*. Spec., IUCN Red List Threat. Species, Int. Union Conserv. Nat., Gland, Switz.  
<https://www.iucnredlist.org/species/7949/45170994> **【\*\*AU: Edit OK?】**
70. Kimura B, Marshall FB, Chen S, Rosenbom S, Moehlman PD, et al. 2011. Ancient DNA from Nubian and Somali wild ass provides insights into donkey ancestry and domestication. *Proc. R. Soc. B* 278(1702):50–57
71. Kimura B, Marshall F, Beja-Pereira A, Mulligan C. 2013. Donkey domestication. *Afr.*

*Archaeol. Rev.* 30(1):83–95

72. Ma X-Y, Ning T, Adeola AC, Li J, Esmailizadeh A, et al. 2020. Potential dual expansion of domesticated donkeys revealed by worldwide analysis on mitochondrial sequences. *Zool. Res.* 41(1):51–60
73. Rosenbom S, Costa V, Al- Araithi N, Kefena E, Abdel- Moneim AS, et al. 2015. Genetic diversity of donkey populations from the putative centers of domestication. *Anim. Genet.* 46(1):30–36
74. Kefena E, Dessie T, Tegegne A, Beja-Pereira A, Yusuf Kurtu M, et al. 2014. Genetic diversity and matrilineal genetic signature of native Ethiopian donkeys (*Equus asinus*) inferred from mitochondrial DNA sequence polymorphism. *Livest. Sci.* 167:73–79
75. Cattani M, Bökönyi S. 2002. Ash-Shumah: an Early Holocene settlement of desert hunters and mangrove foragers in the Yemeni Tihamah. In *Essays on the Late Prehistory of the Arabian Peninsula*, Vol. 92, ed. S Cleuziou, M Tosi, J Zarins, pp. 3–52. Rome: Inst. Ital. Afr. L'Orient.
76. Walker M, Johnsen S, Rasmussen SO, Popp T, Steffensen J-P, et al. 2009. Formal definition and dating of the GSSP (Global Stratotype Section and Point) for the base of the Holocene using the Greenland NGRIP ice core, and selected auxiliary records. *J. Quat. Sci.* 24(1):3–17
77. Giesecke T, Brewer S, Finsinger W, Leydet M, Bradshaw RHW. 2017. Patterns and dynamics of European vegetation change over the last 15,000 years. *J. Biogeogr.* 44(7):1441–56
78. Sandoval-Castellanos E, Wutke S, Gonzalez-Salazar C, Ludwig A. 2017. Coat colour adaptation of post-glacial horses to increasing forest vegetation. *Nat. Ecol. Evol.* 1(12):1816–19
79. Leonardi M, Boschin F, Giampoudakis K, Beyer RM, Krapp M, et al. 2018. Late Quaternary horses in Eurasia in the face of climate and vegetation change. *Sci. Adv.* 4(7):eaar5589
80. Rieder S, Taourit S, Mariat D, Langlois B, Guérin G. 2001. Mutations in the agouti (*ASIP*), the extension (*MC1R*), and the brown (*TYRP1*) loci and their association to coat color phenotypes in horses (*Equus caballus*). *Mamm. Genome* 12(6):450–55
81. Gaunitz C, Fages A, Hanghøj K, Albrechtsen A, Khan N, et al. 2018. Ancient genomes revisit the ancestry of domestic and Przewalski's horses. *Science* 360(6384):111–14
82. Sommer RS, Hegge C, Schmölcke U. 2018. Lack of support for adaptation of post-glacial

- horses to woodlands. *Nat. Ecol. Evol.* 2(4):582–83
83. Warmuth V, Eriksson A, Bower MA, Cañon J, Cothran G, et al. 2011. European domestic horses originated in two Holocene refugia. *PLOS ONE* 6(3):e18194
84. Outram AK, Stear NA, Bendrey R, Olsen S, Kasparov A, et al. 2009. The earliest horse harnessing and milking. *Science* 323(5919):1332–35
85. de Barros Damgaard P, Martiniano R, Kamm J, Moreno-Mayar JV, Kroonen G, et al. 2018. The first horse herders and the impact of early Bronze Age steppe expansions into Asia. *Science* 360(6396):eaar7711
86. Der Sarkissian C, Ermini L, Schubert M, Yang MA, Librado P, et al. 2015. Evolutionary genomics and conservation of the endangered Przewalski's horse. *Curr. Biol.* 25(19):2577–83
87. Bernáldez Sánchez E, García-Viñas E. 2019. The equids represented in cave art and current horses: a proposal to determine morphological differences and similarities. *Anthropozoologica* 54(1):1–12
88. Lira J, Linderholm A, Olaria C, Durling MB, Gilbert MTP, et al. 2010. Ancient DNA reveals traces of Iberian Neolithic and Bronze Age lineages in modern Iberian horses. *Mol. Ecol.* 19(1):64–78
89. Jansen T, Forster P, Levine MA, Oelke H, Hurles M, et al. 2002. Mitochondrial DNA and the origins of the domestic horse. *PNAS* 99(16):10905–10
90. Fages A, Hanghøj K, Khan N, Gaunitz C, Seguin-Orlando A, et al. 2019. Tracking five millennia of horse management with extensive ancient genome time series. *Cell* 177(6):1419–1435.e31
91. Martin SH, Davey JW, Jiggins CD. 2015. Evaluating the use of ABBA-BABA statistics to locate introgressed loci. *Mol. Biol. Evol.* 32(1):244–57
92. Schubert M, Jónsson H, Chang D, Sarkissian CD, Ermini L, et al. 2014. Prehistoric genomes reveal the genetic foundation and cost of horse domestication. *PNAS* 111(52):E5661–69
93. Boeskorov GG, Potapova OR, Protopopov AV, Plotnikov VV, Maschenko EN, et al. 2018. A study of a frozen mummy of a wild horse from the Holocene of Yakutia, East Siberia, Russia. *Mamm. Res.* 63(3):307–14
94. Librado P, Sarkissian CD, Ermini L, Schubert M, Jónsson H, et al. 2015. Tracking the origins of Yakutian horses and the genetic basis for their fast adaptation to subarctic environments.

*PNAS* 112(50):E6889–97

95. Patterson N, Moorjani P, Luo Y, Mallick S, Rohland N, et al. 2012. Ancient admixture in human history. *Genetics* 192(3):1065–93
96. Fages A, Seguin-Orlando A, Germonpré M, Orlando L. 2020. Horse males became over-represented in archaeological assemblages during the Bronze Age. *J. Archaeol. Sci. Rep.* 31:102364
97. Librado P, Fages A, Gaunitz C, Leonardi M, Wagner S, et al. 2016. The evolutionary origin and genetic makeup of domestic horses. *Genetics* 204(2):423–34
98. Wallner B, Vogl C, Shukla P, Burgstaller JP, Druml T, Brem G. 2013. Identification of genetic variation on the horse Y chromosome and the tracing of male founder lineages in modern breeds. *PLOS ONE* 8(4):e60015
99. Felkel S, Vogl C, Rigler D, Dobretsberger V, Chowdhary BP, et al. 2019. The horse Y chromosome as an informative marker for tracing sire lines. *Sci. Rep.* 9(1):1–12
100. Raudsepp T, Finno CJ, Bellone RR, Petersen JL. 2019. Ten years of the horse reference genome: insights into equine biology, domestication and population dynamics in the post-genome era. *Anim. Genet.* 50(6):569–97
101. Liu X, Ma Y, Jiang L. 2017. Genomic regions under selection for important traits in domestic horse breeds. *Front. Agric. Sci. Eng.* 4(3):289–94
102. Librado P, Orlando L. 2018. Detecting signatures of positive selection along defined branches of a population tree using LSD. *Mol. Biol. Evol.* 35(6):1520–35
103. Librado P, Gamba C, Gaunitz C, Sarkissian CD, Pruvost M, et al. 2017. Ancient genomic changes associated with domestication of the horse. *Science* 356(6336):442–45
104. Wilkins AS, Wrangham RW, Fitch WT. 2014. The “domestication syndrome” in mammals: a unified explanation based on neural crest cell behavior and genetics. *Genetics* 197(3):795–808
105. Lord KA, Larson G, Coppinger RP, Karlsson EK. 2020. The history of farm foxes undermines the animal domestication syndrome. *Trends Ecol. Evol.* 35(2):125–36
106. Imsland F, McGowan K, Rubin C-J, Henegar C, Sundström E, et al. 2016. Regulatory mutations in *TBX3* disrupt asymmetric hair pigmentation that underlies Dun camouflage color in horses. *Nat. Genet.* 48(2):152–58
107. Ludwig A, Pruvost M, Reissmann M, Benecke N, Brockmann GA, et al. 2009. Coat color

- variation at the beginning of horse domestication. *Science* 324(5926):485
108. Ludwig A, Reissmann M, Benecke N, Bellone R, Sandoval-Castellanos E, et al. 2015. Twenty-five thousand years of fluctuating selection on leopard complex spotting and congenital night blindness in horses. *Philos. Trans. R. Soc. B* 370(1660):20130386
109. Wutke S, Benecke N, Sandoval-Castellanos E, Döhle H-J, Friederich S, et al. 2016. Spotted phenotypes in horses lost attractiveness in the Middle Ages. *Sci. Rep.* 6(1):1–9
110. Petersen JL, Mickelson JR, Rendahl AK, Valberg SJ, Andersson LS, et al. 2013. Genome-wide analysis reveals selection for important traits in domestic horse breeds. *PLOS Genet.* 9(1):e1003211
111. Andersson LS, Larhammar M, Memic F, Wootz H, Schwochow D, et al. 2012. Mutations in DMRT3 affect locomotion in horses and spinal circuit function in mice. *Nature* 488(7413):642–46
112. Nolte W, Thaller G, Kuehn C. 2019. Selection signatures in four German warmblood horse breeds: tracing breeding history in the modern sport horse. *PLOS ONE* 14(4):e0215913
113. Grilz-Seger G, Neuditschko M, Ricard A, Velie B, Lindgren G, et al. 2019. Genome-wide homozygosity patterns and evidence for selection in a set of European and near Eastern horse breeds. *Genes* 10(7):491
114. Deleted in proof [**\*\*AU: Per house style, general URLs (<http://www.fao.org/dad-is/data/en/>) are listed in text rather than lit. cited. Deleted here and added there, okay?**]
115. Camillo F, Rota A, Biagini L, Tesi M, Fanelli D, Panzani D. 2018. The current situation and trend of donkey industry in Europe. *J. Equine Vet. Sci.* 65:44–49
116. Ivankovic A, Kavar T, Caput P, Mioc B, Pavic V, Dovc P. 2002. Genetic diversity of three donkey populations in the Croatian coastal region. *Anim. Genet.* 33(3):169–77
117. Bordonaro S, Guastella AM, Criscione A, Zuccaro A, Marletta D. 2012. Genetic diversity and variability in endangered Pantesco and two other Sicilian donkey breeds assessed by microsatellite markers. *Sci. World J.* 2012:648427
122. Zhang RF, Xie WM, Zhang T, Lei CZ. 2016. High polymorphism at microsatellite loci in the Chinese donkey. *Genet. Mol. Res.* 15(2):gmr8291
123. Charlesworth B. 2009. Effective population size and patterns of molecular evolution and variation. *Nat. Rev. Genet.* 10(3):195–205
124. Orlando L, Librado P. 2019. Origin and evolution of deleterious mutations in horses. *Genes*

10(9):649

125. McGivney BA, Han H, Corduff LR, Katz LM, Tozaki T, et al. 2020. Genomic inbreeding trends, influential sire lines and selection in the global thoroughbred horse population. *Sci. Rep.* 10(1):1–12
126. Todd ET, Ho SYW, Thomson PC, Ang RA, Velie BD, Hamilton NA. 2018. Founder-specific inbreeding depression affects racing performance in thoroughbred horses. *Sci. Rep.* 8(1):6167
127. Todd ET, Hamilton NA, Velie BD, Thomson PC. 2020. The effects of inbreeding on covering success, gestation length and foal sex ratio in Australian thoroughbred horses. *BMC Genet.* 21:41
128. Welker F, Ramos-Madrigal J, Kuhlwilm M, Liao W, Gutenbrunner P, et al. 2019. Enamel proteome shows that *Gigantopithecus* was an early diverging pongine. *Nature* 576(7786):262–65
129. Cappellini E, Welker F, Pandolfi L, Ramos-Madrigal J, Samodova D, et al. 2019. Early Pleistocene enamel proteome from Dmanisi resolves *Stephanorhinus* phylogeny. *Nature* 574(7776):103–7
130. Han H, Chen N, Jordana J, Li C, Sun T, et al. 2017. Genetic diversity and paternal origin of domestic donkeys. *Anim. Genet.* 48(6):708–11
131. Haller BC, Messer PW. 2019. SLiM 3: forward genetic simulations beyond the Wright-Fisher model. *Mol. Biol. Evol.* 36(3):632–37
132. Guimaraes S, Arbuckle BS, Peters J, Adcock SE, Buitenhuis H, et al. 2020. Ancient DNA shows domestic horses were introduced in the southern Caucasus and Anatolia during the Bronze Age. *Science Advances.* 6(38):eabb0030
133. McHugo GP, Dover MJ, MacHugh DE. 2019. Unlocking the origins and biology of domestic animals using ancient DNA and paleogenomics. *BMC Biology.* 17(1):98
134. Frantz LAF, Bradley DG, Larson G, Orlando L. 2020. Animal domestication in the era of ancient genomics. *Nature Reviews Genetics.* 21(8):449–60

135. Orlando L. 2020. The Evolutionary and Historical Foundation of the Modern Horse: Lessons from Ancient Genomics. *Annual Review of Genetics*. 54(1):-