UNIVERSITY OF COPENHAGEN

Ancient Genomes Reveal Unexpected Horse Domestication and Management Dynamics

Orlando, Ludovic

Ancient Genomes Reveal Unexpected Horse Domestication and Management Dynamics

Ludovic Orlando

The horse was essential to past human societies but became a recreational animal during the twentieth century as the world became increasingly mechanized. As the author reviews here, recent studies of ancient genomes have revisited the understanding of horse domestication, from the very early stages to the most modern developments. They have uncovered several extinct lineages roaming the far ends of Eurasia some 4000 years ago. They have shown that the domestic horse has been significantly reshaped during the last millennium and experienced a sharp decline in genetic diversity within the last two centuries. At a time when no truly wild horses exist any longer, this calls for enhanced conservation in all endangered populations. These include the Przewalski's horse native to Mongolia, and the many local breeds side-lined by the modern agenda, but yet representing the living heritage of over five millennia of horse breeding.

1. Introduction

Today, horses are mostly known as the motive force behind a multi-billion-dollar-per-year racing industry, or as some animal lovers' four-legged companion of choice. They are generally not part of our daily lives. Yet, not so long ago, these animals played a much more prominent role in our societies and were in fact instrumental at many levels.[1] By providing us with speed, horses revolutionized not only the way we traveled, but also the way our genes, diseases, goods, and languages circulated across the planet, effectively globalizing the world for the first time.[2] With horses and dense networks of postal stations, communication became indeed extremely time efficient. This contributed to the stability of vast empires, such as during the Chinese Tang dynasty (618–907 Common Era, CE),[3] or the Great Mongolian Empire that Genghis Kahn unified across the steppes in the thirteenth century CE, in which the *Arrows Messengers* could make essential information travel up to 300 kilometers a day.[4]

Horses also changed the way we made war in a fundamental way, and represented key military assets for past civilizations.[5] With horses, raids indeed became far more effective, and chariots and cavalry charges could be launched at full speed to impact the enemy lines.[6,7] Horses resisted the development of long-distance railway transport in the nineteenth century as the developing economy of large megacities maintained high demands for short-distance transportation.[8] They, however, surrendered to the invention of the combustion engine and the rise of the car industry in the early twentieth century. After paying the heavy tribute of many million casualties to World War 1 and following an increasingly mechanized weaponry, horses also progressively deserted the battlefield.[9,10]

The impact of the horse on human history was, thus, considerable and there are no key historical war figures, from Alexander the Great to Napoleon, or vast open areas, from the steppes of Central Asia to the great plains of the American West, that can be pictured without a horse.

With the advent of high-throughput DNA sequencing,[11] the genetic diversity of modern animals could be increasingly characterized at whole genome scales. Key parallel technological developments in ancient DNA research have opened for genome time travels across the half to one million year time range.[12–14] The generally poor molecular preservation found in the fossil record, however, first limited genetic work to only a handful of remains. But the discovery of specific osseous material, such as the inner ear bone frequently associated with improved DNA preservation rates,[15] unlocked access to ancient genomes at population scales.

Armed with whole genome sequencing techniques, and extensive collections of well-preserved archaeological horse remains, a number of recent studies have started to chart through space and time the genetic transformations introduced by ancient breeders as they first domesticated and further selected, exchanged, and mixed their horse stocks. With nearly 300 ancient genomes sequenced at above onefold coverage, horses have provided the largest genomic time series characterized to date after humans.[16,17] It led to a number of surprising findings, with deep implications for both common evolutionary models of horse domestication and archaeological scenarios of past human

Prof. L. Orlando
Laboratoire d'Anthropobiologie et d'Imagerie de Synthèse
CNRS UMR 5288
Faculté de Médecine de Purpan
37 allées Jules Guesde, Bâtiment A, 31000 Toulouse, France
E-mail: ludovic.orlando@univ-tlse3.fr
Prof. L. Orlando
The GLOBE Institute
University of Copenhagen
Øster Voldgade 5-7, 1350K Copenhagen, Denmark

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/bies.201900164
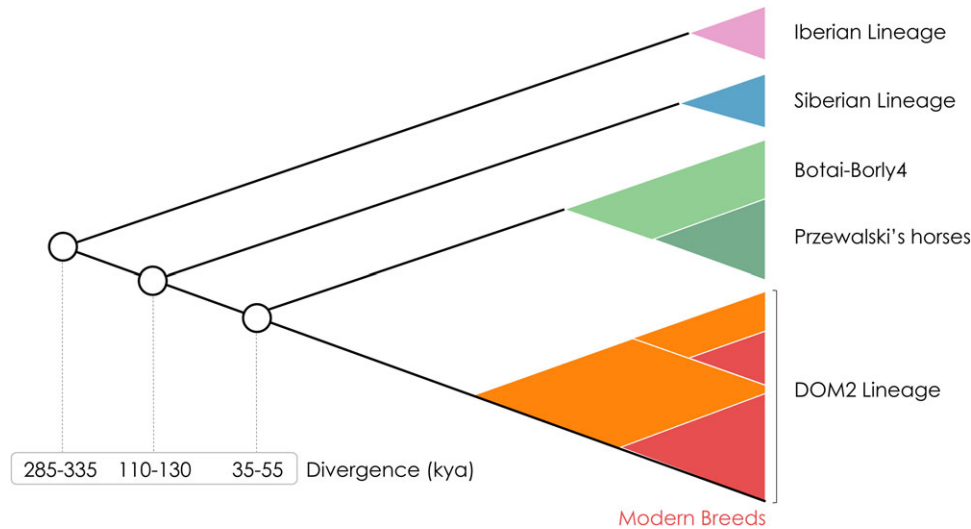
Figure 1. Phylogeny of the main horse lineages identified by genome sequencing. Summarized from the data presented by Schubert et al.,[40] Der Sarkissian et al.,[26] Librado et al.,[41,42] Gaunitz et al.,[24] and Fages et al.[28] The different colors refer to the main horse lineages currently identified, including two extinct lineages in Iberia (pink) and Siberia (blue) as well as two lineages that survived until the present-day in the form of Przewalski's horses (green) and modern domestic horses (red). Ancient animals belonging to the two latter lineages are shown in light green (Botai-Borly) and orange (DOM2), respectively.

migrations. It also revealed that the genetic diversity now found in modern horses only represents a limited fraction of that available to past breeders. This article summarizes the state-of-the-art, indicates areas requiring further research, and emphasizes important implications for conservation. One major lesson, likely true for horses and beyond,[18–20] is that the complex evolutionary history underlying animal domestication cannot be fully reconstructed using patterns of modern DNA variation alone.

2. Do Przewalski's Horses Represent the Last Truly Wild Horses?

The Eneolithic Botai settlements in what is known today as north central Kazakhstan have provided smoking gun archaeological evidence for some of the first stages of horse management and/or domestication.[21] Although this was first debated[22] and does not rule out the possibility of earlier domestication attempts,[23] sites such as Krasnyi Yar and Botai show evidence of horse harnessing, milking and corralling ≈5500 years ago, the earliest evidence to date from the archaeological record.[21,24] Until very recently, Botai horses were naturally considered to be the ancestors of all modern horses. Sequencing their genomes was expected to catch domestication red-handed as humans tamed the animal for the first time, and to characterize the source of the gene pool at the very origin of domestic variation. However, the genome sequences of 20 Botai horses revealed a completely different picture.[24]

Botai horses indeed did not show close genetic affinities to modern domestic breeds. They clustered instead together with the Przewalski's horse, a horse discovered in the late 1870s roaming wild in Mongolia, and considered since as the only truly wild horse living on the planet.[25] In short, the earliest domestic horses known in the archaeological record appeared to be the direct ancestors of the only modern horse that was supposed to

have never been domesticated. It then became obvious that current models of horse evolution required serious rethinking.

Although Botai horses were found related to specimens excavated in the ≈5000-year-old settlement of Borly4 in Kazakhstan, their genomic descent would not be detected in the archaeological record until the first skins of Przewalski's horses prepared in museums in the late nineteenth century were sequenced.[26] This suggests that Botai people domesticated horses, but also that the lineage they formed was not maintained as a domestic stock and returned into the wild sometime after 5000 years ago. Further ancient genomes at the transition between the fourth and the second millennium before common era (BCE) are necessary to map the exact dynamics of this process. For now, it is however noteworthy that analyses of the genomes of Botai people indicated that they were part of a larger steppe cline that also faded away in the early Bronze Age.[27] This supports a model in which both the riders and their horses declined in numbers and were progressively replaced. It is thus not surprising that the genomic trace leading to modern Przewalski's horses was lost after 5000 years ago. This lineage then roamed feral and was not part of any of the bone assemblages investigated so far, all associated with anthropological contexts.

3. Looking for the Ancestors of Modern Domestic Horses

The earliest genomic trace leading to modern domestic horses was found more West than Botai, in fact in the bone remains of a specimen excavated in Hungary and radiocarbon-dated to ≈4100 years ago.[24] Together with an additional 30 horse genomes spanning the last 4000 years, and 80+ others from a follow-up study,[28] this genome belonged to the so-called "DOM2" lineage (**Figure 1**). This lineage led to modern domesticates and split from the Botai-Borly4-Przewalski's horse

lineage ≈35 000–55 000 years ago.[26] Therefore, the time interval stretching between ≈4100 and ≈5000 years ago experienced both the dismiss of the earliest horse domestic lineage and the emergence of the modern one. The exact dynamics underlying this replacement is currently unknown. However, Bayesian skyline reconstructions based on whole mitochondrial genome data and calibrated using the radiocarbon dates of the ancient specimens themselves support an exponential growth of the maternal effective population size at that time,[24] which suggests that horse reproduction was then effectively controlled, and that a full domestication status was, thus, reached.

It is noteworthy that this time period is associated with several waves of large-scale human expansions across the steppes and into Europe,[27,29–31] and ultimately with the development of the spoke-wheeled horse chariots.[1] Both likely had increased horse demands. Further work aimed at mapping the horse genomic diversity in the third millennium BCE should reveal the geographic area(s) and cultural context(s) in which the modern horse lineage emerged and thrived, and whether the Botai-Borly4 ancestry was abandoned upfront or progressively diluted fully into incipient stocks of local wild horses as human populations expanded, through a process called introgressive capture.[32]

4. Did Iberia Represent an Early Domestication Center for Modern Horses?

Since the earliest DOM2 member was identified in eastern Europe, it is essential that the possibility of domestication centers developing outside of the Central Asian steppes is given full attention (e.g., in Anatolia[33,34]; in the Pontic Caspian steppes[23]). One such region, the Iberian Peninsula, located at the western extremity of Eurasia, was recently investigated.[28] There, paleoclimatic niche modeling revealed the continuous persistence of environmental conditions suitable for horses throughout the last ≈45 000 years.[35] Exceptional Upper Paleolithic cave paintings have provided evidence for a long-standing human–horse interaction in the region[36] where the genetic diversity observed at 12 short tandem repeats (STR) loci in modern European breeds also peaks.[37] Altogether, this opened the possibility that the wild horses that had found a climate refugium in Iberia during the Last Ice Age could have been locally domesticated, and have transmitted a substantial fraction of their genes into modern domestic breeds, at least within Iberia.

A handful of genome sequences from Iberian horse specimens dating to ≈4000–4800 years ago, however, supported a different scenario and revealed the existence of a divergent horse lineage in Iberia at that time, which was coined "IBE."[28] This lineage was estimated to have diverged from the ancestors of all other horse lineages at least ≈285 000 years based on modeling that leveraged a multi-population Site Frequency Spectrum. This, and f4-statistics,[38] further indicated only a limited contribution of IBE into the genomes of DOM2 horses, including a Spanish specimen that lived ≈2700 years ago.

With the data at hand, it is not possible to formally reject the possibility that an independent domestication process was started in Iberia. This will require the genetic testing of the IBE ancestry in bone assemblages unequivocally found in a domestication context within the time interval covering ≈4000–

2700 years ago. However, it is clear that the horse lineage native of Iberia has not significantly contributed to the genetic makeup of horse domesticates for at least the last ≈2700 years. The relative excess of genetic diversity found in modern Iberian breeds thus does not reflect the heritage of an early local domestication experience, but rather subsequent breeding strategies maintaining large reproductive stocks and/or effectively incorporating animals from various geographic regions.

5. *The Past is a Foreign Country* (LP Hartley); Massive Diversity Losses Accompanied the Production of Modern Horses

5.1. Several Divergent Horse Lineages Became Extinct During the Last 5000 years

The discovery of a now-extinct but deeply divergent horse lineage in Iberia was not an isolated case. The genomes extracted from bone material found at the other end of the Eurasian range, in the Taymir Peninsula of northern Siberia, revealed that another such lineage, morphologically described as *Equus lenensis*,[39] existed ≈16 000–43 000 years ago.[40] Pairwise-Sequential Markov Coalescent demographic profiles obtained from high-quality genomes[41] indicated a population split from the ancestors of the Botai-Borly4-Przewalski's horse and DOM2 lineages ≈110 000–130 000 years ago, around the last interglacial period (Eemian) (Figure 1). This is consistent with the time estimates obtained using Approximate Bayesian Computation and coalescent simulations under an isolation model,[40] but also when allowing gene-flow into modelling.[28]

We now know that the *E. lenensis* lineage was not endemic to north Siberia, as it was detected also in the foothills of the Altai mountain range in southern Siberia ≈24 000 years ago.[28] It also survived well into the Holocene, at least in Yakutia, where the last known specimen assigned to this lineage was identified and radiocarbon-dated to ≈5200 years ago.[41] All living horses that have been analyzed in this region belong to the DOM2 lineage and do not share more genomic variation with *E. lenensis* than with DOM2 horses from other regions.[41] Therefore, neither of the IBE and the *E. lenensis* lineages seem to have contributed to the genetic makeup of modern horses. It is striking to think that they both existed at the time the Egyptians built the first pyramids, and that they would have never been uncovered without ancient genomic work.

5.2. Additional Ghost Lineages Most Likely Existed But Remain Presently Unsampled

As a matter of fact, it is quite likely that other additional divergent lineages existed as one ≈24 000-year-old *E. lenensis* specimen from northern Siberia was found to carry an extremely divergent mitochondrial haplotype, virtually absent from all other material analyzed so far.[13,42] The same holds true in Iberia, where the Y-chromosome identified in two IBE males fell well outside the variation of all other lineages.[28] Those divergent uniparental haplotypes likely segregated amongst other, yet unidentified ghost populations, and were acquired in *E. lenensis*
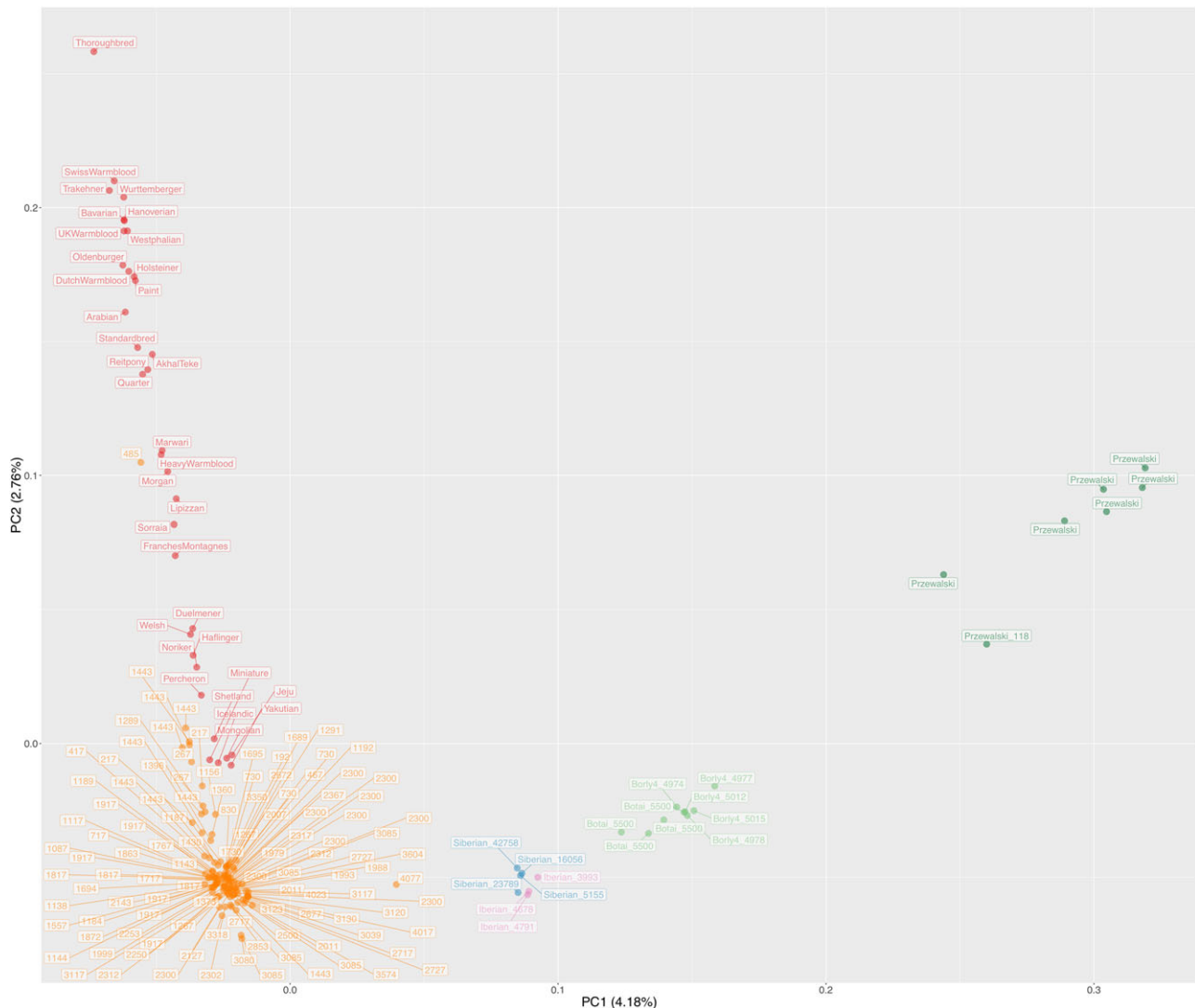
Figure 2. Past and present horse genetic diversity. This principle component analysis is based on the genome sequence presented by Fages and collaborators[28] supplemented by the publicly available data from additional modern horse breeds. The fraction of the genetic variance explained by the first two principle components is shown between parentheses. Ancient DOM2 horses (orange) span the last ≈4100 years. The numbers reported in labels indicate the age the ancient horses represented (years ago). Modern breeds are indicated in red.

and IBE by gene-flow or through incomplete lineage sorting. The temporal and geographic range of the underlying ghost lineages are currently unknown. Nonetheless, together with the IBE and the *E. lenensis* lineages, they provide compelling evidence that the true past diversity of a species cannot be fully comprehended from patterns of modern diversity alone (**Figure 2**). This is true even when considering the last few thousand years, and suggests that more lineages than currently assessed have most likely gone extinct.[43]

5.3. A Few Stallion Lineages Became Increasingly Popular During the Last 2000 years

The declines observed in genetic diversity were not restricted to the most divergent lineages. The DOM2 lineage, which is source

to all modern domesticates, also experienced some severe losses, as first revealed by the Y-chromosome. Modern horses indeed show extremely poor genetic variation at this chromosome, both relative to mtDNA[44] and to Y-chromosome diversity in other domesticates.[45] This was first interpreted as evidence for only a limited number of reproductive males contributing to early horse domestication, in line with stallions being more difficult to tame owing to a more aggressive behavior than females. This model implies that the Y-chromosome diversity entering the domestic horse gene pool was limited from the very early stages of domestication. However, additional haplotypes were identified in ≈2500-year-old Scythian horses[42,46] and in earlier horses from the Iron, Bronze, and Copper ages.[28] Therefore, early domestication stages involved a larger stallion stock than first anticipated.

The reduction of the number of stallion lineages that participated to breeding only started around 2000 years ago,[28,47] a time

1900164 (4 of 8)

when the overall nucleotide diversity of European stallions was approximately halved. This continued during the last thousand years with nucleotide diversity estimates dropping to present-day levels after the Renaissance period. This is in line with modern breeding practices increasingly relying on a few influential stallion lines at that time,[48] and with generally Y-to-autosomal nucleotide diversity ratios closer to zero in modern than in ancient horses.[28] Interestingly, some past equestrian civilizations appear to have selected particular stallion bloodlines for reproduction. This was the case of both the Byzantine and the Great Mongolian Empires. Roman breeders, on the contrary, seem to have been a lot less choosy and to have maintained similar reproductive success across all males.

5.4. Modern Breeding Practices Significantly Eroded the Horse Genetic Diversity

Patterns of autosomal DNA variation have revealed yet another interesting temporal trajectory where the genetic diversity has remained constant for most of the last 3000 years until it significantly dropped in the last ≈250 years.[28] This trend was replicated using two diversity estimates, namely heterozygosity in each individual genome and nucleotide diversity within genomes binned in 250-year time intervals. This suggested that past breeders successfully maintained diverse reproductive pools for millennia, but that in the last few centuries, the introduction of closed stud breeding and the excessive influence of specific bloodlines have considerably reduced the genetic potential of modern horses. This again illustrates the limitations of modern DNA variation, but also how much recent human activities have impacted animal domesticates.

Interestingly, the recent loss of genetic diversity within DOM2 horses was paralleled with an increase in mutational loads.[28,42] These were calculated within protein-coding regions and restricted to homozygous sites to avoid making un-necessary assumptions on the genetic mode of inheritance for each deleterious variant (and their dominance coefficients) and on possible biases introduced by inbreeding.[40] Since inversely correlated to negative selection,[28] these estimates indicated that the reduction in the effective reproductive size of modern studs both eliminated through drift many neutral variants and reduced the efficacy of negative selection in filtering slightly deleterious variants. The inflated amounts of variants with negative fitness impact that are present in the genomes of modern domestic horses has probably reduced their genomic health, and pleads for developing genetically informed reproductive programs to maintain their long-term viability and improve animal welfare.

6. Can We Reconstruct How the Horse Phenotype Was Changed During History?

6.1. Investigating Gene Candidates Can Help Reconstruct Phenotypes That Do Not Fossilize

The availability of a high-quality reference assembly,[49,50] and the development of cost-effective genotyping solutions at the genome-scale[51,52] have opened for the identification of the

genetic basis of key phenotypic traits in horses. This, in turn, has made it possible to predict phenotypes that do not fossilize in ancient animals, on the basis of the genotypes they carried. This includes a whole range of traits, such as coat coloration, capacity for short-distance speed racing, ability to amble, and many others.

It is noteworthy that this information cannot only reveal the traits that past breeders most likely selected, but can also help document past funerary traditions. For instance, the analysis of coat coloration loci in the 13 complete horse skeletons found in the funerary monument of Berel (Kazakhstan) revealed that ≈2500 years ago already, Scythian Pazyryk Iron Age nomads herded the full diversity of horse coat colors present in the region today. Since those horses were specifically killed for the funerals of Pazyryk elite members, the genetic data also showed that sacrifices were not targeted toward particular family groups or coat colors.[42]

Additionally, genotype time series including large numbers of horses have identified shifts in our preference for particular horse characteristics.[53] For instance, spotted coats were extremely popular in Antiquity, but were superseded in the Middle Ages by the reddish-to-brownish uniform chestnut coats.[54]

It is noteworthy that even when sequencing depth is limited and genotypes cannot be called, the sequence data retrieved can still prove useful, as long as the time series investigated include a large number of animals. In this case, the temporal allelic trajectory at any locus can be approximated by randomly sampling one read per specimen, and if neutrality is rejected, derived into quantitative estimates of the time and magnitude of selection.[55,56] Neutral variants can indeed also increase in frequency up to fixation by chance alone, especially when population sizes are small. To support selection, it is thus important to rule out genetic drift as a possible driver, ideally through realistic demographic modeling showing that the allele frequency changes observed at a given locus are exceptional at the scale of the whole genome. Such approaches revealed that selection of variants underpinning key locomotory traits in horses, such as short-distance speed and gaitedness, most likely only started in the last ≈1000 years.[28]

6.2. Scanning the Genome for Footprints of (Positive) Selection

Of course, the genetic basis underlying most horse traits is still unknown. Gene candidate approaches are, thus, necessarily restrictive and cannot uncover the entire range of phenotypic traits that were relevant to past breeders. Other methods can help, however, obtain additional information. For example, the horse genome can be scanned for regions showing maximal changes in allelic frequencies observed between two closely related populations. Such regions are indicative of loci were variants rose in frequency faster than expected, possibly as their carriers were preferentially chosen for breeding the next generation. Determining in which of the two populations selection took place can be done with the addition of an outgroup. This forms the basis of the so-called *Population Branch Statistics* (PBS[57]), but other methods generalize similar principles to infinite numbers of populations in order to account for more complex population histories.[58,59] In order to translate the list of putatively selected loci into possible phenotypes relevant to past breeders,

researchers have commonly leveraged the functional annotations overlaid on the horse genome and identified those over-represented relative to the functional genomic background. The biological processes returned are then only suggestive and require functional validation.

Applying this methodology, Librado and colleagues found, for example, that functional categories pertaining to limb development were over-represented in the genome data of the Pazyryk Iron Age horses sacrificed at Berel.[42] This suggested that Pazyryk Scythians may have selected for a particular morpho-anatomy. Returning to the bone metapods of these ancient horses, and finding morphometric indices characteristic of elevated robustness compared to that seen in other populations lent support to this hypothesis.

A similar approach was applied to a new type of horses showing genetic affinities with Persian Sassanid horses and that entered mainland Europe, Central Asia, and Mongolia where they reshaped the native horse genetic makeup at about the peak of the Arab expansion (between the seventh and ninth centuries CE).[28] The authors have proposed that the breeding success of these horses may pertain to their new morphotypes, as many of the genes located in putatively selected regions following the Arab expansion are involved in the development of the main body plan.

However, the full list of functionally enriched categories returned while scanning genomes for positive selection are often quite difficult to interpret. In the Berel Pazyryk horses, for instance, the neurohypophysis (posterior pituitary) was also identified in addition to limb development. This neuro-endocrine gland produces the oxytocin hormone, which has been shown to trigger milk release, lower stress, and modulate social behaviors,[60] especially maternal bonding, but also bonding between dogs and their owners.[61] Whether the enrichment detected in Pazyryk Scythian horses is indicative of selection in relation to milking and/or behavior facilitating management remains unknown. More importantly, complex traits are only exceptionally determined by single genes and most often involve the small contributions of a large number of genetic variants. In this case, even small but concerted changes in allelic frequencies can result in large phenotypic shifts.[62] Most current methods scanning the genome for individual loci evolving under positive selection will be blind to this type of selection. Therefore, methods designed for detecting polygenic adaptation (e.g., see ref. [63]) may appear more appropriate in the future in order to unveil the true nature of the traits selected by past breeders.

7. Conclusions and Outlook

The first horse genome reference sequence was released in 2009.[49] A decade later, the horse represents the domestic animal with the largest number of ancient genomes sequenced. This resource has not only already helped rewrite important parts of the complex population history underlying horse domestication, but also opened many areas for future research. Arguably, the most surprising discovery thus far pertains to the feral origins of Przewalski's horses.[24] This should by no means undermine ongoing and future conservation efforts toward a population that is still part of the international union for conservation of nature (IUCN) Red List and was once declared extinct in the wild

following reports of the last wild-caught animal and last wild-sighted animal in the late 1950s and 1960s, respectively.[25] The genome sequences obtained from museum specimens, including the holotype, has unambiguously documented a severe loss of their genetic diversity during the twentieth century and the introduction of domestic genetic variants into their genealogy while in captivity.[26] Genetic variants with negative fitness effects now represent a larger fraction of their genome[64], and inbreeding has been galloping, which makes the animal particularly vulnerable. That Botai people herded the ancestors of Przewalski's horses more than 5000 years ago does not change this reality. Conservation efforts must thus not be discontinued but enhanced, especially now that non-invasive[65] and cost-effective genomic tools can positively inform conservation decisions. At the time of global environmental changes and major biodiversity crisis,[66] this should not only apply to horses but to all other endangered species, be they wild, feral, or domesticated.

Conflict of Interest

The author declares no conflict of interest.

[1] P. Kelekna, *The Horse in Human History*, Cambridge University Press, Cambridge/New York **2009**, p. 476.

[2] S. Forrest, *The Age of the Horse, An Equine Journey Through Human History*, Atlantic Monthly Press, New York **2016**, p. 432.

[3] D. Gazagnadou, *La Diffusion des Techniques et les Cultures*, Editions Kimé, Paris **2013**, p. 200.

[4] A. E. Minetti, *Nature* **2003**, *426*, 785.

[5] A. Hyland, *The Horse in the Ancient World*, Sutton Publishing, Gloucestershire **2003**, p. 210.

[6] P. Sidnell, *Warhorse: Cavalry in Ancient Warfare*, Hambledon Continuum, London **2006**, p. 352.

[7] L. A. Di Marco, *War Horse: A History of the Military Horse and Rider*, Westholme Publishing, Yardley **2008**, p. 416.

[8] C. McShane, J. A. Tarr, *The Horse in the City. Living Machines in the Nineteenth Century*, The Johns Hopkins University Press, Baltimore **2007**, p. 280.

[9] E. Baratay, *Bêtes des Tranchées Des Vécus Oubliés*, CNRS éditions, Paris **2013**, p. 300.

[10] C. Milhaud, *1914–1918 L'autre Hécatombe. Enquête sur la Perte de 1 140 000 Chevaux et Mulets*, Belin, Paris **2017**, p. 300.

[11] S. Goodwin, J. D. McPherson, W. R. McCombie, *Nat. Rev. Genet.* **2016**, *17*, 333.

[12] J. Dabney, M. Knapp, I. Gloke, M.T. Gansauge, A. Weihmann, B. Nickel, C. Valdiosera, N. Garcia, S. Pääbo, J. L. Arsuaga, M. Meyer, *Proc. Natl Acad. Sci. USA* **2013**, *110*, 15758.

[13] L. Orlando, A. Ginolhac, G. Zhang, D. Froese, A. Albrechtsen, M. Stiller, M. Schubert, E. Cappellini, B. Petersßen, I. Moltke, P. L. Jonhson, M. Fumagalli, J. T. Vilstrup, M. Raghavan, T. Korneliussen, A. S. Malaspinas, J. Vogt, D. Szklarczyk, C. D. Kelstrup, J. Vinther, A. Dolocan, J. Stenderup, A. M. Velazquez, J. Cahill, M. Rasmussen, Z. Wang, J. Min, G. D. Zazula, A. Seguin-Orlando, C. Mortensen et al., *Nature* **2013**, *499*, 74.

[14] M. Meyer, Q. Fu, A. Aximu-Petri, I. Glocke, B. Nickel, J. L. Arsuaga, I. Martinez, A. Garcia, J. M. de Castro, E. Carbonell, S. Pääbo, *Nature* **2014**, *505*, 403.

[15] R. Pinhasi, D. Fernandes, K. Sirak, M. Novak, S. Connell, S. Alpaslan-Roodenberg, F. Gerritsen, V. Moiseyev, A. Gromov, P. Raczky, A. Anders, M. Pietrusewsky, G. Rollefson, M. Jovanovic, H. Trinhhoang, G. Bar-Oz, M. Oxenham, H. Matsumura, M. Hofreiter, *PLoS One* **2015**, *10*, e0129102.

[16] S. Marciniak, G. H. Perry, *Nat. Rev. Genet.* **2017**, *18*, 659.

[17] K. Brunson, D. Reich, *Trends Genet.* **2019**, *35*, 319.

[18] L. G. Flink, R. Allen, R. Barnett, H. Malmström, J. Peters, J. Eiriksson, L. Andersson, K. Dobney, G. Larson, *Proc. Natl. Acad. Sci. USA* **2015**, *111*, 6184.

[19] L. A. F. Frantz, V. E. Mullin, M. Pionnier-Capitan, O. Lebrasseur, M. Ollivier, A. Perri, A. Linderholm, V. Mattiangeli, M. D. Teadale, E. A. Dimopoulos, A. Tresset, M. Duffraisse, F. McCormick, L. Bartosiewicz, E. Gal, E. A. Nyerges, M. V. Sablin, S. Bréhard, M. Mashkour, A. Balsescu, B. Gillet, S. Hughes, O. Chassaing, C. Hitte, J. D. Vigne, K. Dobney, C. Hänni, D. G. Bradley, G. Larson, *Science* **2016**, *352*, 1228.

[20] L. A. F. Frantz, J. Haile, A. T. Lin, A. Scheu, C. Geörg, N. Benecke, M. Alexander, A. Linderholm, V. E. Mullin, K. G. Daly, V. M. Battista, M. Price, K. J. Gron, P. Alexandri, R. M. Arbogast, B. Arbuckle, A. Balsescu, R. Barnett, L. Bartosiewicz, G. Baryshnikov, C. Bonsall, D. Boric, A. Boroneant, J. Bulatovic, C. Cakirlar, J. M. Carretero, J. Chapman, M. Church, R. Crooijmans, B. De Cupere, et al., *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 17231.

[21] A. Outram, N. A. Stear, R. Bendrey, S. Olsen, A. Kasparov, V. Zaibert, N. Thorpe, R. P. Evershed, *Science* **2009**, *323*, 1332.

[22] N. Benecke, A. von den Driesch, *Prehistoric Steppe Adaptation and the Horse*, McDonald Institute Cambridge, UK **2003**, p. 69.

[23] D. W. Anthony, *The Horse, the Wheel, and Language: How Bronze-Age Riders from the Eurasian Steppes Shaped the Modern World*, Princeton University Press, Princeton and Oxford **2007**, p. 568.

[24] C. Gaunitz, A. Fages, K. Hanghøj, A. Albrechtsen, N. Khan, M. Schubert, A. Seguin-Orlando, I. J. Owens, S. Felkel, O. Bignon-Lau, P. de Barros Damgaard, A. Mittnik, A. F. Mohaseb, H. Davoudi, S. Alquraishi, A. H. Alfarhan, K. A. S. Al-Rasheid, E. Crubézy, N. Benecke, S. Olsen, D. Brown, D. Anthony, K. Massy, V. Pitulko, A. Kasparov, G. Brem, M. Hofreiter, G. Mukhtarova, N. Baimukhanov, L. Lougsas, et al., *Science* **2018**, *360*, 111.

[25] L. Boyd, K. A. Houpt, *Przewalski's Horse: The History and Biology of an Endangered Species (SUNY Series in Endangered Species)*, State University of New York Press, Albany, New York **1994**, p. 313.

[26] C. Der Sarkissian, L. Ermini, M. Schubert, M. A. Yang, P. Librado, M. Fumagalli, H. Jonsson, G. K. Bar-Gal, A. Albrechtsen, F. G. Vieira, B. Petersen, A. Ginolhac, A. Seguin-Orlando, K. Magnussen, A. Fages, C. Gamba, B. Lorente-Galdos, S. Polani, C. Steiner, M. Neuditschko, V. Jagannathan, C. Feh, C. L. Greenblat, A. Ludwog, N. I. Abramson, W. Zimmermann, R. Schafberg, A. Tilhonov, T. Sicheritz-Ponten, E. Willerslev, et al., *Curr. Biol.* **2015**, *25*, 2577.

[27] P. B. Damgaard, R. Martiniano, J. Kamm, J. V. Moreno-Mayar, G. Kroonen, M. Peyrot, G. Barjamovic, S. Rasmussen, C. Zacho, N. Baimukhanov, V. Zaibert, V. Merz, A. Biddanda, I. Merz, V. Loman, V. Evdokimov, E. Usmanova, B. Hemphill, A. Seguin-Orlando, F. E. Yediay, I. Ullah, K. G. Sjögren, K. H. Iversen, J. Choin, C. de la Fuente, M. Ilardo, H. Schroeder, V. Moiseyev, A. Gromov, A. Polyakov, et al., *Science* **2018**, *360*, eaar7711.

[28] A. Fages, K. Hanghøj, N. Khan, C. Gaunitz, A. Seguin-Orlando, M. Leoardi, C. M. Constantz, C. Gamba, K. A. S. Al-Rasheid, S. Albizuri, A. H. Alfarhan, M. Allentoft, S. Alquraishi, D. Anthony, N. Baimukhanov J. H. Barret, J. Bayarsaikhan, N. Benecke, E. Bernaldez-Sanchez, L. Berrocal-Rangel, F. Biglari, S. Boessenkool, B. Boldgiv, G. Brem, D. Brown, J. Burger, E. Crubézy, L. Daugnora, H. Davoudi, P. de Barros Damgaard, et al., *Cell* **2019**, *177*, 1419.

[29] W. Haak, I. Lazaridis, N. Patterson, N. Rohland, S. Mallic, B. Llamas, G. Brandt, S. Nordenfelt, E. Harney, K. Stewardson, Q. Fu, A. Mittnik, E. Banffy, C. Economou, M. Francken, S. Friederich, R. G. Pena, F. Hallgren, V. Khartanovich, A. Khokhlov, M. Kunst, P. Kuznetsov, H. Meller, O. Mochalov, V. Moiseyev, N. Nicklisch, S. L. Pichler, R. Risch, M. A. Roo Guerra, C. Roth, et al., *Nature* **2015**, *522*, 207.

[30] M. Allentoft, M. Sikora, K. G. Sjogren, S. Rasmussen, M. Rasmussen, J. Stenderup, P. B. Damgaard, H. Schroeder, T. Ahlström, L. Vinner, A. S. Malaspinas, A. Margaryan, T. Higham, D. Chivall, N. Lynnerup, L. Harvig, J. Baron, P. Della Casa, P. Dabrowski, P. R. Duffy, A. V. Ebel, A. Epimakhov, K. Frei, M. Furmanek, T. Gralak, A. Gromov, S. Gronkiewicz, G. Grupe, T. Hakdu, R. Jarysz, et al., *Nature* **2015**, *522*, 167.

[31] P. B. Damgaard, N. Marchi, S. Rasmussen, M. Peyrot, G. Renaud, T. Korneliussen, J. V. Moreno-Mayar, M. W. Pedersen, A. Goldbert, E. Usmanova, N. Baimukhanov, V. Loman, L. Hedeager, A. G. Pedersen, K. Nielsen, G. Afanasiev, K. Akmatov, A. Aldashev, A. Alpaslan, G. Baimbetov, V. I. Bazaliiskii, A. Beisenov, B. Boldbaatar, B. Boldgiv, C. Dorzku, S. Ellingvag, D. Erdenebaatar, R. Dajani, E. Dmitriev, V. Evdokimov, et al., *Nature* **2018**, *557*, 369.

[32] G. Larson, D. Q. Fuller, *Ann. Rev. Ecol. Evol. Syst.* **2014**, *45*, 115.

[33] N. Benecke, in *Equids in Time and Space: Papers in Honour of Vera Eisenmann* (Ed: M. Mashkour), Oxbow Books, Oxford **2006**, p. 13.

[34] B. Arbuckle. *Am. Anthropol.* **2012**, *114*, 462.

[35] M. Leonardi, F. Boschin, K. Giampoudakis, R. M. Beyer, M. Krapp, R. Bendrey, R. Sommer, P. Boscato, A. Manica, D. Nogues-Bravo, L. Orlando, *Sci. Adv.* **2018**, *4*, 5589.

[36] N. F. Bicho, A. F. Carvalho, C. Gonzalez-Sainz, J. L. Sanchidrian Torti, V. Villaverde, L. G. Straus, *J. Arc. Met. Theor.* **2007**, *14*, 81.

[37] V. Warmuth, A. Eriksson, M. A. Bower, J. Canon, G. Cothran, O. Distl, M. L. Glowatzki-Mullis, H. Hunt, C. Luis, M. do Mar Oom, I. T. Yupanqui, T. Zabek, A. Manica, *PLoS One* **2011**, *6*, e18194.

[38] N. Patterson, P. Moorjani, Y. Luo, S. Mallick, N. Rohland, Y. Zhan, T. Genschoreck, T. Webster, D. Reich, *Genetics* **2012**, *192*, 1065.

[39] G. G. Boeskorov, O. R. Potapova, A. V. Protopopov, V. V. Plotnikov, E. N. Maschenko, M. V. Shchelchkova, E. A. Petrova, R. Kowalczyk, J. van der Plicht, A. N. Tikhonov, *Mammal Res.* **2018**, *63*, 1.

[40] M. Schubert, H. Jonsson, D. Chang, C. Der Sarkissian, L. Ermini, A. Ginolhac, A. Albrechtsen, I. Dupanloup, A. Foucal, B. Petersen, M. Fumagalli, M. Raghavan, A. Seguin-Orlando, T. S. Korneliussen, A. M. Velzquez, J. Stenderup, C. A. Hoover, C. J. Rubin, A. H. Alfarhan, S. A. Alquraishi, K. A. Al-Rashied, D. E. MacHugh, T. Kalbfleisch, J. N. MacLeod, E. M. Rubin, T. Sicheritz-Ponten, L. Andersson, M. Hofreiter, T. Marques-Bonet, M. T. Gilbert, et al., *Proc. Natl. Acad. Sci. USA* **2014**, *111*, E5661.

[41] P. Librado, C. Der Sarkissian, L. Ermini, M. Schubert, H. Jonsson, A. Albrechtsen, M. Fumagalli, M. A. Yang, C. Gamba, A. Seguin-Orlando,

1900164 (7 of 8)

C. D. Mortensen, B. Petersen, C. A. Hoover, B. Lorente-Galdos, A. Nedoluzhko, E. Boulygina, S. Tsygankova, M. Neuditschko, V. Jagannathan, C. Thèves, A. H. Alfarhan, S. A. Alquraishi, K. A. Al-Rasheid, T. Sicheritz-Ponten, R. Popov, S. Grigoriwv, A. N. Alekseev, E. M. Rubin, M. McCue, S. Rieder, et al., *Proc. Natl. Acad. Sci. USA* **2015**, *112*, E6889.

[42] P. Librado, C. Gamba, C. Gaunitz, C. Der Sarkissian, M. Pruvost, A. Albrechtsen, A. Fages, N. Khan, M. Schubert, V. Jagannathan, S. Serres-Armero, L. F. K. Kuderna, I. S. Povolotskaya, A. Seguin-Orlando, S. Lepetz, M. Neuditschlo, C. Thèves, S. Alquraishi, A. H. Alfrahan, K. Al-Rasheid, S. Rieder, Z. Samahev, H. P. Francfort, N. Benecke, M. Hofreiter, M. Ludwig, C. Keyser, T. Marques-Bonet, B. Ludes, et al., *Science* **2017**, *356*, 442.

[43] L. Orlando, A. Cooper, *Ann. Rev. Ecol. Evol. Syst.* **2014**, *45*, 573.

[44] C. Vila, J. A. Leonard, A. Gotherstrom, S. Marklund, K. Sandberg, K. Liden, R. K. Wayne, H. Ellegren, *Science* **2001**, *291*, 474.

[45] G. Lindgren, N. Backstrom, J. Swinburne, L. Hellborg, A. Einarsson, K. Sandberg, G. Cothran, C. Vila, M. Binns, H. Ellegren, *Nat. Genet.* **2004**, *36*, 335.

[46] S. Lippold, M. Knapp, T. Kuznetsova, J. A. Leonard, N. Benecke, A. Ludwig, M. Rasmussen, A. Cooper, J. Weinstock, E. Willerslev, B. Shapiro, M. Hofreiter, *Nat. Commun.* **2011**, *2*, 450.

[47] S. Wutke, E. Sandoval-Castellanos, M. Benecke, H. J. Döhle, S. Friederich, J. Gonzalez, M. Hofreiter, L. Lougas, O. Magnell, A. S. Malaspinas, A. Morales-Muniz, L. Orlando, M. Reissmann, A. Trinks, A. Ludwig, *Sci. Adv.* **2018**, *4*, eaap9691.

[48] B. Wallner, N. Palmieri, C. Vogl, D. Rigler, E. Bozla, T. Druml, V. Jagannathan, T. Leeb, R. Fries, J. Tetens, G. Thaller, J. Metzger, O. Distl, G. Lindgren, C. J. Rubin, L. Andersson, R. Schaefer, M. McCue, M. Neuditschko, S. Rieder, C. Schlötterer, G. Brem, *Curr. Biol.* **2017**, *27*, 2029.

[49] C. M. Wade, E. Giulotto, S. Sigursson, M. Zoli, S. Gnerre, F. Imsland, T. L. Lear, D. L. Adelson, E. Bailey, R. R. Bellone, H. Blocker, O. Distl, R. C. Edgar, M. Garber, T. Leeb, E. Mauceli, J. N. MacLeod, M. C. Penedo, J. M. Raison, T. Sharpe, J. Vogel, L. Andersson, D. F. Antczak, T. Biagi, M. M. Binns, B. P. Chowdhary, S. J. Coleman, G. Della Valle, S. Fryc, G. Guérin, et al., *Science* **2009**, *326*, 865.

[50] T. S. Kalbfleish, E. S. Rice, M. S. DePriest, B. P. Walenz, M. S. Hestand, J. R. Vermeesch, B. L. O. Connell, I. T. Fiddes, A. O. Vershinina, N. F. Saremi, J. L. Petersen, C. J. Finno, R. R. Bellone, M. E. McCue, S. A. Brooks, E. Bailey, L. Orlando, R. E. Green, D. C. Miller, D. F. Antczak, J. N. MacLeod, *Commun. Biol.* **2018**, *1*, 197.

[51] J. L. Petersen, J. R. Mickelson, E. G. Cothran, L. S. Andersson, J. Axelsson, E. Bailey, D. Bannasch, M. M. Binns, A. S. Borges, P. Brama, A. da Camara Machado, O. Distl, M. Felicetti, L. Fox-Clipsham, K. T. Graves, G. Guérin, B. Haase, T. Hasegawa, K. Hemmann, E. W. Hill, T. Leeb, G. Lindgren, H. Lohi, M. S. Lopes, B. A. McGivney, S. Mikko, N. Orr, M. C. Penedo, R. J. Piercy, M. Raekallio, et al., *PLoS One* **2013**, *8*, e54997.

[52] R. J. Schaefer, M. Schubert, E. Bailey, D. L. Bannasch, E. Barrey, G. K. Bar-Gal, G. Brem, S. A. Brooks, O. Distl, R. Fries, C. J. Finno, V. Gerber, B. Haase, V. Jagannathan, T. Kalbfleisch, T. Leeb, G. Lindgren, M. S. Lopes, N. Mach, A. da Camara Machado, J. N. MacLeod, A. McCoy, J. Metzger, C. Penedo, S. Polani, S. Rieder, I. Tammen, J. Tetens, G. Thaller, A. Verini-Supplizi, et al., *BMC Genomics* **2017**, *18*, 565.

[53] A. Ludwig, M. Reissmann, N. Benecke, R. Bellone, E. Sandoval-Castellanos, M. Ciesla, G. G. Fortes, A. Morales-Muniz, M. Hofreiter, M. Pruvost, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2015**, *370*, 20130386.

[54] S. Wutke, N. Benecke, E. Sandoval-Castellanos, H. J. Döhle, S. Friederich, J. Gonzalez, J. H. Hallsson, M. Hofreiter, L. Lougas, O. Magnell, A. Morales-Muniz, L. Orlando, A. H. Palsdottir, M. Reissmann, A. Ruttay, A. Trinks, A. Ludwig, *Sci. Rep.* **2016**, *6*, 38548.

[55] J. G. Schraiber, S. N. Evans, M. Slatkin, *Genetics* **2016**, *203*, 493.

[56] A. S. Malaspinas. *Mol. Ecol.* **2016**, *25*, 24.

[57] X. Yi, Y. Liang, E. Huerta-Sanchez, X. Jin, Z. X. Cuo, J. E.Pool, X. Xu, H. Jiang, N. Vinckenbosch, T. S. Korneliussen, H. Zheng, T. Liu, W. He, K. Li, R. Luo, X. Nie, H. Wu, M. Zhao, H. Cao, J. Zou, Y. Shan, S. Li, Q. Yang, P. N. Asan, G. Tiang, J. Xu, X. Liu, T. Jiang, R. Wu, G. Zhou, M. Tang, et al., *Science* **2010**, *329*, 75.

[58] P. Librado, L. Orlando, *Mol. Biol. Evol.* **2018**, *35*, 1520.

[59] A. Refoyo-Martinez, R. R. da Fonseca, K. Halldorsdottir, E. Arnason, T. Mailund, F. Racimo, *Genome Res.* **2019**, *29*, 1506.

[60] P. S. Churchland, P. Winkielman, *Horm. Behav.* **2012**, *61*, 392.

[61] M. Nagasawa, S. Mitsui, S. En, N. Ohtani, M. Ohta, Y. Sakuma, T. Onaka, K. Mogi, T. Kikusui, *Science* **2015**, *348*, 333.

[62] J. Pritchard, A. Di Rienzo, *Nat. Rev. Genet.* **2010**, *11*, 665.

[63] F. Racimo, J. J. Berg, J. K. Pickrell, *Genetics* **2018**, *208*, 1565.

[64] P. Librado, L. Orlando, *Genes* **2019**, *10*, E649.

[65] J. L. Metcalf, S. J. Song, J. T. Morton, S. Weiss, A. Seguin-Orlando, F. Joly, C. Feh, P. Taberlet, E. Coissac, A. Amir, E. Willerslev, R. Knight, V. McKenzie, L. Orlando, *Sci. Rep.*, **2017** *7*, 15497.

[66] E. Kolbert, *The Sixth Extinction: An Unnatural History*, Picador USA, New York **2014**, p. 336.

1900164 (8 of 8)