

TITLE: Tracking five millennia of horse management with extensive ancient genome time-series

AUTHORS:

Antoine Fages^{1,2,+}, Kristian Hanghøj^{1,2,+}, Naveed Khan^{2,3+}, Charleen Gaunitz², Andaine Seguin-Orlando^{1,2}, Michela Leonardi^{2,4}, Christian McCrory Constantz^{2,5}, Cristina Gamba², Khaled A. S. Al-Rasheid⁶, Silvia Albizuri⁷, Ahmed H. Alfarhan⁶, Morten Allentoft², Saleh Alquraishi⁶, David Anthony⁸, Nurbol Baimukhanov⁹, James H. Barrett¹⁰, Jamsranjav Bayarsaikhan¹¹, Norbert Benecke¹², Eloísa Bernáldez-Sánchez¹³, Luis Berrocal-Rangel¹⁴, Fereidoun Biglari¹⁵, Sanne Boessenkool¹⁶, Bazartseren Boldgiv¹⁷, Gottfried Brem¹⁸, Dorcas Brown⁸, Joachim Burger¹⁹, Eric Crubézy¹, Linas Daugnora²⁰, Hossein Davoudi^{21,22}, Peter de Barros Damgaard², María de los Ángeles de Chorro y de Villa-Ceballos²³, Sabine Deschler-Erb²⁴, Cleia Detry²⁵, Nadine Dill²⁴, Maria do Mar Oom²⁶, Anna Dohr^{27,28,29}, Sturla Ellingvåg³⁰, Diimaajav Erdenebaatar³¹, Homa Fathi^{21,32}, Sabine Felkel¹⁸, Carlos Fernández-Rodríguez³³, Esteban García-Viñas³⁴, Mietje Germonpré³⁵, José D. Granado²⁴, Jón H. Hallsson³⁶, Helmut Hemmer¹⁹, Michael Hofreiter³⁷, Aleksei Kasparov³⁸, Mutalib Khasanov³⁹, Roya Khazaeli^{21,32}, Pavel Kosintsev⁴⁰, Kristian Kristiansen⁴¹, Tabaldiev Kubatbek⁴², Lukas Kuderna⁴³, Pavel Kuznetsov⁴⁴, Haeedeh Laleh^{21,32}, Jennifer A. Leonard⁴⁵, Johanna Lhuillier⁴⁶, Corina Liesau von Lettow-Vorbeck⁴⁷, Andrey Logvin⁴⁸, Lembi Lõugas⁴⁹, Arne Ludwig^{50,51}, Cristina Luis^{52,53,54}, Ana Margarida-Arruda²⁶, Tomas Marques-Bonet^{43,55,56,57}, Raquel Matoso Silva⁵⁴, Victor Merz⁵⁸, Enkhbayar Mijiddorj³¹, Bryan K. Miller⁵⁹, Oleg Mochalov⁴⁴, Fatemeh A. Mohaseb^{21,32,60}, Arturo Morales⁶¹, Ariadna Nieto-Espinet⁶², Heidi Nistelberger¹⁶, Vedat Onar⁶³, Albína H. Pálsdóttir^{16,36}, Vladimir Pitulko³⁸, Konstantin Pitskhelauri⁶⁴, Mélanie Pruvost⁶⁵, Petra Rajic Sikanjic⁶⁶, Anita Rapan Papeša⁶⁷, Natalia Roslyakova⁴⁴, Alireza Sardari⁶⁸, Eberhard Sauer⁶⁹, Renate Schafberg⁷⁰, Amelie Scheu¹⁹, Jörg

Schibler²⁴, Angela Schlumbaum²⁴, Nathalie Serrand^{61,71}, Aitor Serres-Armero⁴³, Beth Shapiro⁷², Shiva Sheikhi Seno^{21,32,61}, Irina Shevnina⁴⁸, Sonia Shidrang⁷³, John Southon⁷⁴, Bastiaan Star¹⁶, Naomi Sykes^{75,76}, Kamal Taheri⁷⁷, William Taylor⁷⁸, Wolf-Rüdiger Teegen^{27,28}, Tajana Trbojević Vukičević⁷⁹, Simon Trixl²⁹, Dashzeveg Tumen⁸⁰, Sainbileg Undrakhbold¹⁷, Emma Usmanova⁸¹, Ali Vahdati⁶⁵, Silvia Valenzuela-Lamas^{59, 9}, Catarina Viegas²⁵, Barbara Wallner¹⁸, Jaco Weinstock⁸², Victor Zaibert⁸³, Benoit Clavel⁶⁰, Sébastien Lepetz⁶⁰, Marjan Mashkour^{21,32,60}, Agnar Helgason⁸⁴, Kári Stefánsson⁸⁴, Eric Barrey⁸⁵, Eske Willerslev², Alan K. Outram⁷⁵, Pablo Librado^{1,2}, Ludovic Orlando^{1,2,*}.

AFFILIATIONS:

¹Laboratoire d'Anthropobiologie Moléculaire et d'Imagerie de Synthèse, CNRS UMR 5288, Université de Toulouse, Université Paul Sabatier, 31000 Toulouse, France

²Lundbeck Foundation GeoGenetics Center, University of Copenhagen, Øster Voldgade 5-7, 1350K Copenhagen, Denmark

³Department of Biotechnology, Abdul Wali Khan University, Mardan, Pakistan

⁴Evolutionary Ecology Group, Department of Zoology, University of Cambridge. Cambridge CB2 3EJ, UK

⁵Institute for Immunity, Transplantation and Infection, Stanford University, Stanford, CA 94305, USA

⁶Zoology Department, College of Science, King Saud University, Riyadh 11451, Saudi Arabia

⁷Seminari d'Estudis i Recerques Prehistòriques, HAR2017-87695-P, Universitat de Barcelona, Barcelona, Spain

⁸Anthropology Department, Hartwick College 1, Oneonta, Ny 13820 USA

⁹Shejire DNA project, 050046 Almaty, Kazakhstan

¹⁰McDonald Institute for Archaeological Research, Department of Archaeology, University of Cambridge, Cambridge, CB2 3ER, UK

¹¹National Museum of Mongolia, Ulaanbaatar 210646, Mongolia

¹²Deutsches Archäologisches Institut (DAI), 14195 Berlin, Germany

¹³Laboratorio de Paleontología y Paleobiología, Instituto Andaluz del Patrimonio Histórico, Sevilla, Spain

¹⁴Departamento de Prehistoria y Arqueología, Universidad Autónoma de Madrid, Madrid, Spain

¹⁵Department of Paleolithic, National Museum of Iran, 1136918111, Tehran, Iran

¹⁶Centre for Ecological and Evolutionary Synthesis (CEES), Department of Biosciences, University of Oslo, Postbox 1066, Blindern, 0316 Oslo, Norway

¹⁷Ecology Group, Department of Biology, School of Arts and Sciences, National University of Mongolia, Ulaanbaatar 14201, Mongolia

¹⁸Institute of Animal Breeding and Genetics, Department of Biomedical Sciences, Veterinary University of Vienna, 1210 Vienna, Austria

¹⁹ Palaeogenetics Group, Institute of Organismic and Molecular Evolution (iOME), Johannes Gutenberg-University Mainz, 55099 Mainz, Germany

²⁰Osteological material research laboratory, Klaipėda university, Klaipėda 92294, Lithuania

²¹Department of Osteology, National Museum of Iran, 1136918111, Tehran, Iran

²²Department of Archaeology, Faculty of Humanities, Tarbiat Modares University, Tehran, Iran

²³Centro de Biología Molecular Severo Ochoa (CSIC-UAM), E-28049, Madrid, Spain

²⁴Integrative prähistorische und naturwissenschaftliche Archäologie (IPNA), CH-4055 Basel, Switzerland

- ²⁵Uniarq, Centro de Arqueologia da Universidade de Lisboa, Faculdade de Letras da Universidade de Lisboa, 1600-214 Lisboa, Portugal
- ²⁶CE3C – Centre for Ecology, Evolution and Environmental Changes, Faculdade de Ciências, Universidade de Lisboa, 1749-016 Lisboa, Portugal
- ²⁷Institute for Pre- and Protohistoric Archaeology and Archaeology of the Roman Provinces, Ludwig-Maximilians-University Munich, 80539 München, Germany
- ²⁸ArchaeoBioCenter, Ludwig-Maximilians-University Munich, 80539 München, Germany
- ²⁹Institute of Palaeoanatomy, Domestication Research and History of Veterinary Medicine, Ludwig-Maximilians-University Munich, 80539, München, Germany
- ³⁰Explico Foundation, 6900 Florø, Norway
- ³¹Department of Archaeology, Ulaanbaatar State University, Ulaanbaatar 51, Mongolia
- ³²Archaeozoology section, Bioarchaeology Laboratory of the Central Laboratory, University of Tehran, Tehran CP1417634934, Iran
- ³³Departamento de Historia, Facultad de Filosofía y Letras, Universidad de León, León, Spain
- ³⁴Departamento de Sistemas Físicos, Químicos y Naturales, Universidad Pablo de Olavide, 41013 Sevilla, Spain
- ³⁵Operational Direction, Earth and History of Life, Royal Belgian Institute of Natural Sciences, 1000, Brussels, Belgium
- ³⁶Faculty of Agricultural and Environmental Sciences, The Agricultural University of Iceland, Keldnaholti - Árleyni 22, 112 Reykjavík, Iceland
- ³⁷University of Potsdam, Faculty of Mathematics and Natural Sciences, Institute for Biochemistry and Biology, 14476 Potsdam, Germany
- ³⁸Institute for the History of Material Culture, Russian Academy of Sciences, St. Petersburg 191186, Russia

- ³⁹Archaeology Institute of Samarkand, Uzbekistan
- ⁴⁰Institute of Plant and Animal Ecology, Urals Branch of the Russian Academy of Sciences, Ekaterinburg 620144, Russia
- ⁴¹Department of Historical Studies, University of Gothenburg, Sweden
- ⁴²Department of History, Kyrgyz-Turkish Manas University, Bishkek, Kyrgyzstan
- ⁴³Institut de Biologia Evolutiva, (CSIC-Universitat Pompeu Fabra), PRBB, Barcelona, Catalonia 08003, Spain
- ⁴⁴Samara State University of Social Science and Education, Samara, Russia
- ⁴⁵Conservation and Evolutionary Genetics Group, Estación Biológica de Doñana (EBD-CSIC), 41092 Sevilla, Spain
- ⁴⁶Laboratoire Archéorient, UMR 5133, Maison de l'Orient et de la Méditerranée, 69365 Lyon cedex 7, France
- ⁴⁷Departamento de Prehistoria y Arqueología, Universidad Autónoma de Madrid, Madrid, Spain
- ⁴⁸Laboratory for Archaeological Research, Faculty of History and Law, Kostanay State University, Kostanay, Kazakhstan
- ⁴⁹Archaeological Research Collection, Tallinn University, 10130 Tallinn, Estonia
- ⁵⁰Department of Evolutionary Genetics, Leibniz Institute for Zoo and Wildlife Research, Berlin 10315, Germany
- ⁵¹Faculty of Life Sciences, Albrecht Daniel Thaer-Institute, Humboldt University Berlin, 10115 Berlin, Germany
- ⁵²Museu Nacional de História Natural e da Ciência, Universidade de Lisboa, Portugal
- ⁵³Centro Interuniversitário de História das Ciências e da Tecnologia (CIUHCT), Faculdade de Ciências, Universidade de Lisboa, Lisboa, Portugal
- ⁵⁴Instituto Universitário de Lisboa (ISCTE-IUL), CIES-IUL, Lisboa, Portugal

⁵⁵Catalan Institution of Research and Advanced Studies (ICREA), 08010, Barcelona, Spain

⁵⁶CNAG-CRG, Centre for Genomic Regulation (CRG), Barcelona Institute of Science and Technology (BIST), 08028 Barcelona, Spain

⁵⁷Institut Català de Paleontologia Miquel Crusafont, Universitat Autònoma de Barcelona, Edifici ICTA-ICP, c/ Columnes s/n, 08193, Cerdanyola del Vallès, Barcelona, Spain

⁵⁸S.Toraighyrov Pavlodar State University, Joint Research Center for Archeological Studies, 637000 Pavlodar, Kazakhstan

⁵⁹University of Oxford, Faculty of History, George Street, Oxford OX1 2RL, United Kingdom

⁶⁰Centre National de la Recherche Scientifique, Muséum National d'Histoire Naturelle, Archéozoologie, Archéobotanique, Sociétés, Pratiques et Environnements (UMR 7209), 75005 Paris, France

⁶¹Laboratory of Archaeozoology, Dept. Biología, Universidad Autónoma de Madrid, Madrid, Spain

⁶²Archaeology of Social Dynamics Group (ADS), Institució Milà i Fontanals - Consejo Superior de Investigaciones Científicas (IMF-CSIC), 08001 Barcelona, Spain

⁶³Osteoarchaeology Practice and Research Center & Department of Anatomy, Faculty of Veterinary Medicine, Istanbul University-Cerrahpaşa

⁶⁴Ivane Javakhishvili Tbilisi State University, Tbilisi, 0179 Tbilisi, 1 Chavchavadze ave., Georgia

⁶⁵Université de Bordeaux, CNRS, UMR 5199-PACEA, 33615 Pessac Cedex, France

⁶⁶Institute for Anthropological Research, Gajeva 32, 10000 Zagreb, Croatia

⁶⁷Vinkovci Municipal Museum, 32 100 Vinkovci, Croatia

⁶⁸Iranian Center for Archaeological Research (ICAR), Iranian Cultural Heritage, Handicrafts, and Tourism Organization (ICHHTO)

⁶⁹School of History, Classics and Archaeology, The University of Edinburgh, Edinburgh EH8 9AG, UK

⁷⁰Central Natural Science Collections (ZNS), Martin Luther University Halle-Wittenberg, Domplatz 4, 06108 Halle (Saale), Germany

⁷¹INRAP Guadeloupe, Centre de recherches archéologiques, UMR 7209 CNRS/MNHN, 97113 Gourbeyre

⁷²Department of Ecology and Evolutionary Biology and Howard Hughes Medical Institute, University of California, Santa Cruz, CA 95060, USA

⁷³Saeedi Institute for Advanced Studies, University of Kashan, Kashan 87317-51167, Iran

⁷⁴Department Earth System Science, University of California Irvine, Irvine, CA 92697, USA

⁷⁵Department of Archaeology, University of Nottingham, Nottingham NG7 2RD, UK

⁷⁶Department of Archaeology, University of Exeter, Exeter EX4 4QE, UK

⁷⁷Kermanshah Regional Water Authority, 67145-1466, Kermanshah, Iran

⁷⁸Max Planck Institute for the Science of Human History, 07745 Jena, Germany

⁷⁹Department of Anatomy, Histology and Embryology, Faculty of Veterinary Medicine, University of Zagreb, 10 000 Zagreb, Croatia

⁸⁰Department of Anthropology and Archaeology, School of Arts and Sciences, National University of Mongolia, Ulaanbaatar 14201, Mongolia

⁸¹Saryarka Archaeological Institute of Buketov Karaganda State University, Karaganda 100074, Kazakhstan

, 34320, Avcılar, Istanbul, Turkey

⁸²Faculty of Humanities (Archaeology), University of Southampton, Avenue Campus, Highfield, Southampton SO17 1BF, UK

⁸³Scientific Research Institute of Archaeology and Steppe Civilizations, Al Farabi Kazakh National University, 050040 Almaty, Kazakhstan

⁸⁴deCODE Genetics, 101 Reykjavik, Iceland

⁸⁵GABI UMR1313, INRA, AgroParisTech, Université Paris-Saclay, Jouy-en-Josas, France

[†]These authors contributed equally to this work.

*Correspondence: ludovic.orlando@univ-tlse3.fr

SUMMARY

Horse domestication revolutionized warfare and accelerated travel, trade and the geographic expansion of languages. Here we present the largest DNA time-series for a non-human organism to date, including genome-scale data from 149 ancient animals and 129 ancient genomes (≥ 1 -fold coverage), 87 of which are new. This extensive dataset allows us to assess the modern legacy of past equestrian civilisations. We find that two extinct horse lineages existed during early domestication, one at the far western (Iberia) and the other at the far eastern range (Siberia) of Eurasia. None of these contributed significantly to modern diversity. We show that the influence of Persian-related horse lineages increased following the Islamic conquests in Europe and Asia. Multiple alleles associated with elite-racing, including at the MSTN “speed gene”, only rose in popularity within the last millennium. Finally, the development of modern breeding impacted genetic diversity more dramatically than the previous millennia of human management.

KEYWORDS

Ancient DNA; Domestication; Animal Breeding; Horses; Mules; Extinct lineages; Selection; Diversity; Management; Equestrian civilizations

TEXT

INTRODUCTION

Horses provided humans with the first opportunity to spread genes, diseases, and culture well above their own speed (Allentoft et al., 2015; Haak et al., 2015; Rasmussen et al., 2014). Horses remained paramount to transportation even after the advent of steam locomotion, and until the widespread use of motor vehicles (Kelekna, 2009). Horses also revolutionized

warfare, pulling chariots at full speed in the Bronze Age, providing the foundation for mounted battle in the early Iron Age, and facilitating the spread of cavalry during Antiquity (Drews, 2004). Today, horses remain essential to the economy of developing countries and to the leisure and racing industries of developed countries (Faostat, 2009).

The earliest archaeological evidence of horse milking, harnessing and corralling is found in the ~5,500 year-old Botai culture of Central Asian steppes (Gaunitz et al., 2018; Outram et al., 2009) (see (Kosintsev and Kuznetsov, 2013) for discussion). Botai-like horses are, however, not the direct ancestors of modern domesticates but of Przewalski's horses (Gaunitz et al., 2018). The genetic origin of modern domesticates thus remains contentious, with suggested candidates in the Pontic-Caspian steppes (Anthony, 2007), Anatolia (Arbuckle, 2012; Benecke, 2006) and Iberia (Uerpman, 1990; Warmuth et al., 2011). Irrespective of the origins of domestication, the horse genome is known to have been reshaped significantly within the last ~2,300 years (Librado et al., 2017; Wallner et al., 2017; Wutke et al., 2018). However, when and in which context(s) such changes occurred remains largely unknown.

RESULTS

Genome dataset

To clarify the origins of domestic horses and reveal their subsequent transformation by past equestrian civilisations, we generated DNA data from 278 equine subfossils with ages mostly spanning the last six millennia ($n=265$, 95%; Figure 1AB; Table S1; STAR Methods).

Endogenous DNA content was compatible with economical sequencing of 87 new horse genomes to an average depth-of-coverage of 1.0-9.3-fold (median=3.3-fold; Table S2). This more-than-doubles the number of ancient horse genomes hitherto characterized. With a total

of 129 ancient genomes, 30 modern genomes and new genome-scale data from 132 ancient individuals (0.01-0.9-fold, median=0.08-fold), our dataset represents the largest genome-scale time-series published for a non-human organism (Tables S2-S4; STAR Methods).

Most specimens were genetically confirmed as horses (175 males, 70 females; Table S1; STAR Methods). Six belonged to other equine species, including three hemiones from Chalcolithic, Bronze and Iron Age sites of Iran and three Roman and Byzantine donkeys (Figure 1A). A total of 27 specimens were genetically assigned to mules (the offspring of a donkey jack and a horse mare), which are difficult to identify in fragmentary fossil records using morphology alone (Schubert et al., 2017). The oldest mules identified are from the La Tène Iron Age site of Saint-Just (France) but they were also found in Roman and medieval Europe as well as Byzantine Turkey.

Changes in horse management through time and their impact on diversity

Previous work comparing the sequence variation present in modern horse genomes and the genomes of 11 ancient horses belonging to the Scythian Pazyryk culture suggested important changes in the management of available genetic resources within the last ~2,300 years (Librado et al., 2017). Our thorough temporal genome sampling allowed us to delineate more precisely when these changes happened. We ensured accurate diversity estimates in ancient horses by only considering genomes sequenced at minimum 1-fold depth-of-coverage and implementing the three following approaches. First, enzymatic treatment against the most prevalent post-mortem DNA damage helped avoid inflating past diversity estimates (STAR Methods). Second, only sites least affected by damage, such as non-CpG dinucleotides and transversion sites, were considered. Last, we checked that diversity measurements were robust both to residual error rates and sequencing depth (Figure S1; STAR Methods).

All modern breeds investigated here showed a ~16.4% median drop in individual heterozygosity levels relative to horses that lived prior to ~200 years ago (Wilcoxon test, p -value = 2.0×10^{-13} ; Figure 2C; Figure S2; STAR Methods). This contrasts with steady heterozygosity levels during the previous four millennia, reflecting that earlier equestrian civilisations managed and maintained higher levels of genetic diversity. A similar trend was found in autosomal π diversity, which severely declined during the most recent time interval with sufficient data to enable calculations (i.e. the last ~400 years). Autosomal π profiles also supported a demographic expansion from La Tène to Roman Europe, possibly pertaining to the growing demand for horses during Roman times (Figure 2A, STAR Methods).

The recent decays of autosomal π diversity and heterozygosity suggest a severe reduction in horse breeding stock within the last few centuries, parallel to the significant changes in agricultural practices underpinning modern studs. This reduction in effective size is expected to have increased mutational loads genome-wide by reducing the efficacy of purifying selection (Cruz et al., 2008; Schubert et al., 2014a). To test this, we calculated conservative estimates for the mutational loads at homozygous sites within protein-coding genes and accounting for possible inbreeding differences (Librado et al., 2017) (calculations at heterozygous sites were proven impracticable, in agreement with (Pedersen et al., 2017); Figure S3AB). As expected, mutational load estimates correlated with reduced selection, as measured from differential diversity patterns at non-synonymous and synonymous sites, and from sites classified as deleterious and non-deleterious on the basis of their evolutionary conservation across multiple vertebrate species (STAR Methods). We found mutational loads increasing during the last ~200 years, parallel to changes in breed reproductive management (~4.6% median load increment; Wilcoxon test, p -value = 8.3×10^{-12} ; Figure 2D). Our data

therefore support the contention that reproductive strategies implemented in the last few centuries reduced the chance to eliminate deleterious variants from domestic horse stock.

The choice of stallions for reproduction and its impact in the last 2,000 years

The Y-chromosomal diversity is extremely limited in modern horses (Lindgren et al., 2004) but was greater in the past (Librado et al., 2017; Lippold et al., 2011), indicating that specific stallion lines have become increasingly prominent. Previous work showed that this process started ~900 BCE-400 CE ((Before) Common Era), however, on the basis of only four polymorphic SNPs (Wutke et al., 2018). We thus leveraged our 105 stallions and the ~1,500 orthologous polymorphic sites recovered at monocopy regions to gain further temporal resolution for this reduction in Y-chromosome diversity (STAR Methods). We considered all past time intervals of 250 years represented by a minimum of 3 males in Asia and in Europe separately, to limit the impact of geographic structure. This revealed that Y-chromosome nucleotide diversity (π) decreased steadily in both continents during the last ~2,000 years but dropped to present-day levels only after 850-1,350 CE (Figure 2B; Figure S2E; STAR Methods). This is consistent with the dominance of a ~1,000-700 years-old oriental haplogroup in most modern studs (Felkel et al., 2018; Wallner et al., 2017). Our data also indicate that the growing influence of specific stallion lines post-Renaissance (Wallner et al., 2017) was responsible for as much as a 3.8-10.0-fold drop in Y-chromosomal diversity.

We then calculated Y-chromosome π estimates within past cultures represented by a minimum of three males to clarify the historical contexts that most impacted Y-chromosomal diversity. This confirmed the temporal trajectory observed above as Byzantine horses (287-861 CE) and horses from the Great Mongolian Empire (1,206-1,368 CE) showed limited yet larger-than-modern diversity. Bronze Age Deer Stone horses from Mongolia, medieval

Aukštaičiai horses from Lithuania (C9th-C10th) and Iron Age Pazyryk Scythian horses showed similar diversity levels (0.000256 - 0.000267, Figure 2A). However, diversity was larger in La Tène, Roman and Gallo-Roman horses, where Y-to-autosomal π ratios were close to 0.25. This contrasts to modern horses, where marked selection of specific patriline drives Y-to-autosomal π ratios substantially below 0.25 (0.0193-0.0396, Figure 2A). The close-to-0.25 Y-to-autosomal π ratios found in La Tène, Roman and Gallo-Roman horses suggest breeding strategies involving an even reproductive success among stallions or equally-biased reproductive success in both sexes (Wilson Sayres et al., 2014).

Influence of Persian lines post C7th/C9th

We next tracked evidence for animal exchange between past cultures by mapping genetic variation through space and time. We included all samples belonging to a particular archaeological culture, as long as they collectively accumulated a minimal genome depth-of-coverage of 2-fold ($n = 186$, Table S5). TreeMix reconstructions (Pickrell and Pritchard, 2012) revealed that modern Shetland and Icelandic ponies were most closely related to a group of north European horses including pre-Viking Pictish horses from C6th/C7th Britain, Viking horses, and one C9th/C10th horse from Estonia (Saardjave) (Figure 3; STAR Methods). This is in line with the historical expansion of Scandinavian seafaring warriors in the British Isles and Iceland between the late C8th/C11th (Brink and Price, 2008). These horses formed a sister clade to mainland European horses spanning the Iron Age to the C7th and a number of cultures, including in the La Tène and (Gallo-)Roman periods. Other modern European native breeds (e.g. Friesian, Duermener, Sorraia and Connemara) were found to belong to yet another clade, first appearing in Europe at Nustar, Croatia in the C9th, but not present at that time in northern Europe (Aukštaičiai, Lithuania). This suggests the introduction of new domestic lineages to the south of mainland Europe between the C7th/C9th,

a time strikingly coincident with the peak of Arab raids on the Mediterranean coasts, including Croatia (Skylitzes and Wortley, 2010). This, and the earliest identification of this clade within two Sassanid Persian horses from Shahr-I-Qumis, Iran (C4th/C5th), supports the growing influence of oriental bloodlines in mainland Europe following at least the C9th.

Moving focus to Asia, steppe Iron Age Pazyryk Scythian and Xiongnu horses appear related to Karasuk horses, locally present in the Minusinsk Basin of South Siberia during the late Bronze Age (Mallory and Adams, 1997). This lineage of horses survived at least until the C8th in Central Asia at Boz Adyr, Kyrgyzstan. However, Mongolian horses from the Uyghur (C7th/C9th, Khotont_UCIE2012x85_1291) and the Great Mongolian Empire (C13th) clustered together with C9th horses from Kazakhstan (Gregorevka4_PAVH2_1192 and Zhanaturmus_Issyk1_1143) within the group descending from the two Shahr-I-Qumis Sassanid Persian horses. Therefore, the population shift observed in Europe during the C7th/C9th was also mirrored in Central Asia and Mongolia.

Gait, speed and selection

We next aimed to identify possible differences in the traits selected prior to and after the C7th/C9th transition. Only one subset of horses provided sufficient data for calculating the Population Branch Statistics (PBS) (Yi et al., 2010) considering at least 10 individuals above 1-fold depth-of-coverage per archaeological site (Tables S6-S7; STAR Methods). It consisted of 11 Bronze Age Deer Stone horses (representing the pre-C7th/C9th Asian group), 11 Gallo-Roman horses (pre-C7th/C9th European horses) and 17 Byzantine horses (post-C7th/C9th). Enrichment analyses of the genes overlapping the top-1,000 50 kb windows revealed that functional categories related to cervical and thoracic vertebrae were over-represented in Byzantine horses (adjusted p -values ≤ 0.05 ; Figure 4A; STAR Methods). Eleven genes within

the HOXB/C clusters, instrumental for the development of the main body plan and the skeletal system (Pearson et al., 2005), featured amongst the windows showing the strongest PBS values (Figure 4A). These findings were robust to the number of outlier windows considered and the significance threshold retained was conservative relative to neutral expectations (STAR Methods). Therefore, our results provide evidence for selection towards changes in the skeletal morpho-anatomy of the post-C7th/C9th horses related to Sassanid Persians.

We further explored temporal shifts in the traits that are commonly selected by modern breeders. We retraced allelic trajectories at key genomic locations associated with or causal for locomotion, body size and coat-coloration phenotypes. We also tracked known variants underlying genetic disorders through time (Figure S5; STAR Methods). Allele frequencies were calculated every 1,000 years (step size = 250 years) and restricted to the lineage leading to modern domesticates (DOM2, Figure 4BC). Mutations causing genetic disorders were extremely rare, including the GYS1 H allele responsible for a severe myopathy in Quarter horses and other heavy and saddle horse breeds. This allele was almost absent across all archaeological sites and, thus, not particularly advantageous for past breeders despite the increased glycogen storage muscular capacity conferred in starch-poor diets (McCue et al., 2008). Spotted and dilution alleles also remained at low frequencies, in contrast to the MC1R chestnut coat-coloration allele, which was relatively common, except at the end of the Middle Ages (Figure 4B, Figure S6). The DMRT3 allele that causes ambling and improves speed capacity in Icelandic horses (Kristjansson et al., 2014) was first seen in a Great Mongolian Empire horse (TavanTolgoi_GEP14_730) and slowly gained in frequency thereafter (Figure S5). Interestingly, the MSTN “speed” gene was amongst the PBS selection candidates in the post-C7th/C9th branch (Figure 4A). We found that a number of alleles involved in racing

performance, including at MSTN and PDK4/ACN9 (Hill et al., 2010), rose in frequency in the last 600-1,100 years (100-1,100 and 600-1,600 years ago; Figure 4B). Allele frequencies at these three loci also varied significantly more through time than other mutations genome-wide (Figure 4C). Altogether, this supports that speed capacity was increasingly selected in the last millennium.

Discovering two divergent and extinct lineages of horses

Domestic and Przewalski's horses are the only two extant horse lineages (Der Sarkissian et al., 2015). Another lineage was genetically identified from three bones dated to ~43,000-5,000 years ago (Librado et al., 2015; Schubert et al., 2014a). It showed morphological affinities to an extinct horse species described as *Equus lenensis* (Boeskorov et al., 2018). We now find that this extinct lineage also extended to Southern Siberia, following the Principal Component Analysis (PCA), phylogenetic and f_3 -outgroup clustering of a ~24,000 years-old specimen from the Tuva Republic within this group (Figures 3 and 5A; Figure S7A). This new specimen (MerzlyYar_Rus45_23789) carries an extremely divergent mtDNA only found in the New Siberian Islands some ~33,200 years ago (Orlando et al., 2013) (Figure 6A; STAR Methods), and absent from the three bones previously sequenced. This suggests that a divergent ghost lineage of horses contributed to the genetic ancestry of MerzlyYar_Rus45_23789. However, both the timing and location of the genetic contact between *E. lenensis* and this ghost lineage remain unknown.

PCA revealed that native Iberian horses (IBE) from the 3rd and early 2nd mill. BCE cluster separately from *E. lenensis*, Przewalski's horses (and their Botai-Borly4 ancestors) and the lineage leading to modern domesticates (DOM2) (Figure 5A; STAR Methods). This indicates that a fourth lineage of horses existed during the early phase of domestication (Gaunitz et al.,

2018; Outram et al., 2009). Members of this lineage possess their own distinctive mtDNA haplogroup (Figure 6A; STAR Methods) and are represented by two Spanish pre-Bell Beaker Chalcolithic settlements (Cantorella and Camino de Las Yeseras) and a Bronze Age village (El Acequión), with archaeological contexts compatible with both wild and domestic status.

Modelling demography and admixture of extinct and extant horse lineages

Phylogenetic reconstructions without gene flow indicated that IBE differentiated prior to the divergence between DOM2 and Przewalski's horses (Figure 3; STAR Methods). However, allowing for one migration edge in TreeMix suggested closer affinities with one single Hungarian DOM2 specimen from the 3rd mill. BCE (Dunaujvaros_Duk2_4077), with extensive genetic contribution (38.6%) from the branch ancestral to all horses (Figure S7B). This, and the extremely divergent IBE Y-chromosome (Figure 6B), suggest that a divergent but yet unidentified ghost population could have contributed to the IBE genetic makeup.

To test this and further assess the underlying population history, we explicitly modelled demography and admixture by fitting the multi-dimensional Site Frequency Spectrum in momi2 (Kamm et al. 2018; STAR Methods). The two best-supported scenarios (Figure 5C) provided divergence time estimates on par with previous work, first ~113-119 kya for the *E. lenensis* split (Librado et al., 2015; Schubert et al., 2014a), then ~34-44 kya for that of Przewalski's horse and DOM2 lineages (Der Sarkissian et al., 2015). In both models, IBE and *E. lenensis* show strong genetic affinities, with no less than 93.2%-98.8% genetic input from the former into the branch ancestral to *E. lenensis*, some ~285-333 kya. The magnitude of this pulse could suggest that the two lineages in fact split at that time, but that a more divergent ghost population contributed ~1.2-6.8% ancestry into IBE, pushing the momi2 estimate for the IBE divergence to deeper times (~539-1,246 kya). The strong genetic affinity between

IBE and *E. lenensis* is consistent with the results of Struct-f4, a new method developed here leveraging all possible combinations of f_4 -statistics to provide a 3D representation of ancestral population relationships that is robust to lineage-specific genetic drift (Figure 5B; STAR Methods), as opposed to PCA projections.

Rejecting Iberian contribution to modern domesticates

The genome sequences of four ~4,800-3,900 years-old IBE specimens characterized here allowed us to clarify ongoing debates about the possible contribution of Iberia to horse domestication (Benecke, 2006; Uerpmann, 1990; Warmuth et al., 2011). Calculating the so-called f_G ratio (Martin et al., 2015) provided a minimal boundary for the IBE contribution to DOM2 members (Cahill et al., 2013) (Figure 7A). The maximum of such estimate was found in the Hungarian Dunaujvaros_Duk2_4077 specimen (~11.7-12.2%), consistent with its TreeMix clustering with IBE when allowing for one migration edge (Figure S7B). This specimen was previously suggested to share ancestry with a yet-unidentified population (Gaunitz et al., 2018). Calculation of f_4 -statistics indicates that this population is not related to *E. lenensis* but to IBE (Figure 7B; STAR Methods). Therefore, IBE or horses closely related to IBE, contributed ancestry to animals found at an Early Bronze Age trade center in Hungary from the late 3rd mill. BCE. This could indicate that there was long-distance exchange of horses during the Bell Beaker phenomenon (Olalde et al., 2018). The f_G minimal boundary for the IBE contribution into an Iron Age Spanish horse (ElsVilars_UE4618_2672) was still important (~9.6-10.1%), suggesting that an IBE genetic influence persisted in Iberia until at least the 7th century BCE in a domestic context. However, f_G estimates were more limited for almost all ancient and modern horses investigated (median = ~4.9-5.4%; Figure 7A). Analytical predictions and population modelling with momi2 further confirmed that IBE

contributed only minimal ancestry (~1.4-3.8%) to modern DOM2 horses, and well prior to their domestication (~34-44 kya).

DISCUSSION

Recent advances in ancient DNA research have opened access to the complete genome sequence of past individuals. These have so far mostly improved our understanding of the evolutionary history of our own lineage, based on hundreds of individual whole genomes and genome-scale data from thousands of individuals (Marciniak and Perry, 2017). Our study represents the first effort to apply the available technology at similar scales to a non-human organism. With 129 ancient genomes and genome-scale data from 149 additional ancient animals, our dataset unveils the past complexity of horse evolution, including the recent impact of humans by means of diversity management, selection and hybridization.

We genetically identified two mules within the La Tène Iron Age site of Saint-Just (France). Mules represented invaluable animals to past societies, being more sure-footed, more resistant to diseases, and harder working than horses. They are, however, difficult to identify morphologically from fragmentary material. Our work gives definitive proof that mules have been bred since at least ~2,200 years ago, despite considerable cost implications of producing sterile stock (Laurence, 1999).

We found that Y-chromosome diversity in horses declined steadily within the last ~2,000 years, with male reproductive success becoming skewed following the (Gallo-) Roman period. This indicates that breeders increasingly chose specific stallions for breeding from the Middle Ages onwards, consistent with the dominance of a ~700-1,000 years-old Arabian/oriental haplogroup in most modern studs (Felkel et al., 2018; Wallner et al., 2017).

Together with the increasing affinity to Sassanid Persian horses detected in the genomes of European and Asian horses after the C7th/C9th, this suggests that the Byzantine-Sassanid wars and the early Islamic conquests significantly impacted breeding and exchange. The legacy of these historical events has persisted until now as the majority of the modern breeds investigated here clustered within a phylogenetic group related to Sassanid Persian horses. During the same time period, the horse phenotype was also significantly reshaped, especially for locomotion, speed capacity and morpho-anatomy. Whether this partly or fully reflects the direct influence of Arabian/oriental lines requires further tests.

Most strikingly, we found that while past horse breeders maintained diverse genetic resources for millennia after they first domesticated the horse, this diversity dropped by ~16% within the last 200 years. This illustrates the massive impact of modern breeding and demonstrates that the history of domestic animals cannot be fully understood without harnessing ancient DNA data. Importantly, recent breeding strategies have also limited the efficacy of negative selection and led to the accumulation of deleterious variants within the genome of horses. This illustrates the genomic cost of modern breeding. Future work should focus on testing how much recent progress in veterinary medicine and the improving animal welfare have contributed to limit the fitness impact of deleterious variants.

In addition to the two extant lineages of horses, we report two other lineages at the far eastern and western range of Eurasia, in Iberia (IBE) and Siberia (*E. lenensis*). Their genomes suggest the presence of other yet unidentified ghost populations. The IBE and *E. lenensis* lineages are now extinct but lived at the time horses were first domesticated. None of them, however, contributed significant ancestry to modern domesticates. Interestingly, Upper Paleolithic cave paintings in Europe have often been proposed to depict Przewalski's horses due to striking morphological resemblance (Leroi-Gourhan, 1958). Our sample set included

one horse from the Goyet cave, Belgium dated to ~35,870 years ago. Although characterized at limited coverage (0.49-fold), D-statistics revealed closer genetic affinity to IBE and DOM2 than to the ancestors of Przewalski's horses ($-15.5 < Z\text{-scores} < -2.4$). European cave painting is, therefore, unlikely to depict Przewalski's horses. It may instead represent the ancestors of the tarpan, assuming that this taxonomically contentious lineage neither represents domestic horses turned feral nor domestic-wild hybrids but truly wild horses that went extinct in the late C19th (Groves, 1994).

Iberia was suggested as a possible domestication center for horses on the basis of both archaeological arguments (Benecke, 2006) and geographic patterns of genetic variation in modern breeds (Uerpmann, 1990; Warmuth et al., 2011). Previous ancient DNA data were limited to short mtDNA sequences of pre-Bronze Age to medieval specimens (Lira et al., 2010), and remained indecisive regarding the contribution of Iberia to horse domestication. Our work shows that IBE horses have not genetically contributed to the vast majority of DOM2 domesticates investigated here, ancient or modern alike, excepting one horse in Bronze Age Hungary, possibly following the Bell-Beaker phenomenon, and an additional one in Iron Age Iberia. Population modelling also confirmed limited contribution within modern domesticates, largely pre-dating domestication. Therefore, IBE cannot represent a main domestication source. Given that other candidates in the Eneolithic Botai culture from Central Asia do not represent DOM2 ancestors (Gaunitz et al. 2018), the origins of the modern domestic horse remain open.

Future work must focus on mapping genomic affinities in the 3rd and 4th mill. BCE, especially at other candidate regions for early domestication in the Pontic-Caspian (Anthony, 2007) and Anatolia (Arbuckle, 2012; Benecke, 2006). Finer mapping of the Persian-related influence at around the time of the Islamic conquest and the diversity hotspots in place prior to modern

stud breeding will also improve our understanding of the source(s) and dynamics underlying the makeup of modern diversity.

ACKNOWLEDGEMENTS

We thank the reviewers for insightful comments and suggestions that helped improve the manuscript. We thank the staff of the Danish National High-Throughput DNA Sequencing Center for technical support; Rachel Ballantyne, Maude Barme, Lucie Cottier, Jean-Marc Fémolant, Stéphane Frère, Gaëtan Jouanin, Patrice Meniel, Nicolas Morand, Anaïs Ortiz, Ollivier Putelat, Vida Rajkovaca, Julie Rivière, Opale Robin, Noémie Tomadini, Jean-Hervé Yvinec and the National Museum of Iceland for providing access to osteological material; Bazartseren Boldbat for his help and guidance; Laurent Frantz, Dan Bradley and Greger Larson for critical reading of the manuscript; Clio Der Sarkissian and Luca Ermini for preliminary analyses, technical support and insightful discussion. BB was supported by the Taylor Family-Asia Foundation Endowed Chair in Ecology and Conservation Biology. ML was supported by a Marie-Curie Individual Fellowship (MSCA-IF-67852); LL was supported by the Estonian Research Council (PRG29); CL was supported by FCT (SFRH/BPD/100511/2014); PK, NR and OM were supported by the Ministry of Educations and Science of Russian Federation (33.1907, 2017/II4) and the Russian Scientific foundation (18-18-00137); TMB was supported by the BFU2017-86471-P (MINECO/FEDER, UE) grant, the U01 MH106874 grant, Howard Hughes International Early Career, Obra Social "La Caixa" and Secretaria d'Universitats i Recerca del Departament d'Economia i Coneixement de la Generalitat de Catalunya; VP was supported by Russian Science Foundation (16-18-10265). This research received support from the SYNTHESYS Project (www.synthesys.info/), which is financed by European Community Research Infrastructure Action under the Seventh Framework "Capacities" Programme. This work was supported by

the Danish National Research Foundation (DNRF94); Initiative d'Excellence Chaires d'attractivité, Université de Toulouse (OURASI); the International Highly Cited Research Group Program (HCRC#15-101), Deanship of Scientific Research, King Saud University; the Villum Fonden miGENEPI research project, the Swiss National Science Foundation (CR1311_140638); the Research Council of Norway (project 230821/F20); the investigation grant HAR2016-77600-P, Ministerio de Economía y Competitividad, Spain, and; the National Science Foundation (ANS-1417036). This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 681605).

AUTHOR CONTRIBUTIONS

LO conceived the project and designed research; AF, CG and NK performed ancient DNA laboratory work, with input from CGa, ML, CMC and LO; ASO performed DNA sequencing; KH, PL and LO designed, coordinated and performed computational analyses, with input from AF for hybrid detection and mitochondrial analyses, and from SF, BW and GB for Y-chromosome analyses; MA, DA, NBa, JBar, JBay, NBe, EBS, LBR, FB, SB, BB, DB, JBu, LD, HD, PBD, MACVC, SDE, CD, ND, MMO, SE, DE, HF, CFR, MG, JHH, RK, MK, PKo, TK, PKu, HL, JAL, JL, CLLV, AL, LL, AL, CL, AMA, RMS, VM, EM, BM, OM, FAM, AM, ANE, HN, AHP, VP, KP, MP, PRS, ARP, NR, ASZ, ES, RS, AS, JS, AS, NS, ASA, SSS, IS, SS, JS, NS, KT, WT, WRT, TTV, ST, DT, SU, EU, AV, SVL, OV, CV, JW, VZ, BC, SL, MM and AKO provided samples and/or information about archaeological/historical contexts; SA, AHA, KAR, BS, JS, EW and LO provided reagents and material; AF, KH, PL and LO prepared figures and tables; AF, CG, KH, PL and LO wrote the supplementary information, with input from JBar, FB, BB, MG, CL, HD, JL, LL,

MK, PRS, ARP, AS, BC, SL and MM; LO wrote the paper, with input from AKO, PL, and all other co-authors.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

- Abad, Ò.E., i Garra, A.M., and Bieto, E.T. (2011). Cantorella (Maldà, Urgell), un nou assentament a l'aire lliure del neolític final-calcolític i del bronze ple a la vall del Corb.
- Allentoft, M.E., Sikora, M., Sjögren, K.-G., Rasmussen, S., Rasmussen, M., Stenderup, J., Damgaard, P.B., Schroeder, H., Ahlström, T., Vinner, L., et al. (2015). Population genomics of Bronze Age Eurasia. *Nature* 522, 167–172.
- Anthony, D. (2007). *The Wheel, the Horse, and Language*.
- Arbuckle, B.S. (2012). Pastoralism, Provisioning, and Power at Bronze Age Acemhöyük, Turkey. *Am. Anthropol.* 114, 462–476.
- Barlow, A., Hartmann, S., Gonzalez, J., Hofreiter, M., and Paijmans, J.L.A. (2018). Consensify: a method for generating pseudohaploid genome sequences from palaeogenomic datasets with reduced error rates.
- Barrett, J.H.B. (2012). Being an islander: production and identity at Quoygrew, Orkney, AD 900-1600 (McDonald Institute for Archaeological Research).
- Barrett, J.H., and Slater, A. (2009). New Excavations at the Brough of Deerness: Power and Religion in Viking Age Scotland. *Journal of the North Atlantic* 81–94.
- Benecke, N. (2006). On the beginning of horse husbandry in the southern Balkan Peninsula—the horse bones from Kirklareli-Kanhgecit (Turkish Thrace). *Equids in Time and Space: Papers in Honour of Véra Eisenmann* 9, 92.
- Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Series B Stat. Methodol.* 57, 289–300.
- Bertašius, M., and Daugnora, L. (2001). Viking age horse graves from Kaunas region (Middle Lithuania). *Int. J. Osteoarchaeol.* 11, 387–399.
- Biglari, F., and Taheri, K. (2000). The discovery of Upper Paleolithic remains at Mar Kuliyan and Mar Dalan Cave, Rawansar. *Essays on the Archaeology, Geology, Geography and Culture of Rawansar Area*. Kermanshah: Taq-E Bostan Publications. P 7–27.
- Blasco, C., Rios, P., and Liesau, C. (2011). Yacimientos calcolíticos con campaniforme de la región de Madrid: nuevos estudios (Universidad Autónoma de Madrid Madrid).
- Boeskorov, G.G., Potapova, O.R., Protopopov, A.V., Plotnikov, V.V., Maschenko, E.N., Shchelchkova, M.V., Petrova, E.A., Kowalczyk, R., van der Plicht, J., and Tikhonov, A.N. (2018). A study of a frozen mummy of a wild horse from the Holocene of Yakutia, East Siberia, Russia. *Mammal Research*.
- Briggs, A.W., Stenzel, U., Johnson, P.L.F., Green, R.E., Kelso, J., Prüfer, K., Meyer, M., Krause, J., Ronan, M.T., Lachmann, M., et al. (2007). Patterns of damage in genomic DNA sequences from a Neandertal. *Proc. Natl. Acad. Sci. U. S. A.* 104, 14616–14621.

- Briggs, A.W., Stenzel, U., Meyer, M., Krause, J., Kircher, M., and Pääbo, S. (2010). Removal of deaminated cytosines and detection of in vivo methylation in ancient DNA. *Nucleic Acids Res.* 38, e87.
- Brink, S., and Price, N. (2008). *The Viking World* (Routledge).
- Cahill, J.A., Green, R.E., Fulton, T.L., Stiller, M., Jay, F., Ovseyanikov, N., Salamzade, R., St John, J., Stirling, I., Slatkin, M., et al. (2013). Genomic evidence for island population conversion resolves conflicting theories of polar bear evolution. *PLoS Genet.* 9, e1003345.
- Charlesworth, B. (2009). Effective population size and patterns of molecular evolution and variation. *Nat. Rev. Genet.* 10, 195–205.
- Corbin, L.J., Blott, S.C., Swinburne, J.E., Vaudin, M., Bishop, S.C., and Woolliams, J.A. (2010). Linkage disequilibrium and historical effective population size in the Thoroughbred horse. *Anim. Genet.* 41 Suppl 2, 8–15.
- Cruz, F., Vilà, C., and Webster, M.T. (2008). The legacy of domestication: accumulation of deleterious mutations in the dog genome. *Mol. Biol. Evol.* 25, 2331–2336.
- Der Sarkissian, C., Ermini, L., Schubert, M., Yang, M.A., Librado, P., Fumagalli, M., Jónsson, H., Bar-Gal, G.K., Albrechtsen, A., Vieira, F.G., et al. (2015). Evolutionary genomics and conservation of the endangered Przewalski's horse. *Curr. Biol.* 25, 2577–2583.
- Drews, R. (2004). *Early Riders: The Beginnings of Mounted Warfare in Asia and Europe* (Routledge).
- Durand, E.Y., Patterson, N., Reich, D., and Slatkin, M. (2011). Testing for ancient admixture between closely related populations. *Mol. Biol. Evol.* 28, 2239–2252.
- Dyson, R.H. (1989). The Iron Age architecture at Hasanlu: an essay. *Expedition* 31, 107.
- Dyson, R.H. (1999). The Achaemenid painted pottery of Hasanlu IIIA. *Anatolian Studies* 49, 101–110.
- Faostat, F.A.O. (2009). *Statistical databases*. Food and Agriculture Organization of the United Nations.
- Felkel, S., Vogl, C., Rigler, D., Jagannathan, V., Leeb, T., Fries, R., Neuditschko, M., Rieder, S., Velie, B., Lindgren, G., et al. (2018). Asian horses deepen the MSY phylogeny. *Anim. Genet.* 49, 90–93.
- Gamba, C., Hanghøj, K., Gaunitz, C., Alfarhan, A.H., Alquraishi, S.A., Al-Rasheid, K.A.S., Bradley, D.G., and Orlando, L. (2016). Comparing the performance of three ancient DNA extraction methods for high-throughput sequencing. *Mol. Ecol. Resour.* 16, 459–469.
- Gaunitz, C., Fages, A., Hanghøj, K., Albrechtsen, A., Khan, N., Schubert, M., Seguin-Orlando, A., Owens, I.J., Felkel, S., Bignon-Lau, O., et al. (2018). Ancient genomes revisit the ancestry of domestic and Przewalski's horses. *Science* 360, 111–114.
- Germonpré, M. (2004). Influence of climate on sexual segregation and cub mortality in Pleniglacial cave bear. *The Future from the Past. Archaeozoology in Wildlife Conservation*

and Heritage Management.

Green, R.E., Krause, J., Briggs, A.W., Maricic, T., Stenzel, U., Kircher, M., Patterson, N., Li, H., Zhai, W., Fritz, M.H.-Y., et al. (2010). A draft sequence of the Neandertal genome. *Science* 328, 710–722.

Groves, C.P. (1994). Morphology, habitat, and taxonomy. Przewalski's Horse--The History and Biology of an Endangered Species. *BOYD L* 39–60.

Haak, W., Lazaridis, I., Patterson, N., Rohland, N., Mallick, S., Llamas, B., Brandt, G., Nordenfelt, S., Harney, E., Stewardson, K., et al. (2015). Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* 522, 207–211.

Hall, S.J.G. (2016). Effective population sizes in cattle, sheep, horses, pigs and goats estimated from census and herdbook data. *Animal* 10, 1778–1785.

Haller, B.C., and Messer, P.W. (2017). SLiM 2: Flexible, Interactive Forward Genetic Simulations. *Mol. Biol. Evol.* 34, 230–240.

Hansman, J., and Stronach, D. (1970). A Sasanian Repository at Shahr-I Qūmis. *J. R. Asiat. Soc. GB Irel.* 142–155.

Hill, E.W., Gu, J., McGivney, B.A., and MacHugh, D.E. (2010). Targets of selection in the Thoroughbred genome contain exercise-relevant gene SNPs associated with elite racecourse performance. *Anim. Genet.* 41, 56–63.

Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P.L.F., and Orlando, L. (2013). mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* 29, 1682–1684.

Jónsson, H., Schubert, M., Seguin-Orlando, A., Ginolhac, A., Petersen, L., Fumagalli, M., Albrechtsen, A., Petersen, B., Korneliussen, T.S., Vilstrup, J.T., et al. (2014). Speciation with gene flow in equids despite extensive chromosomal plasticity. *Proc. Natl. Acad. Sci. U. S. A.* 111, 18655–18660.

Jonvel, R. (2014). Démuin « Le village ». Rapport final d'opération. Autour du château médiéval (IXe-XVIIe siècle) (UnivArchéo, UPJV, Amiens).

Kamm, J.A., Terhorst, J., Durbin, R., and Song, Y.S. (2018). Efficiently inferring the demographic history of many populations with allele count data. *bioRxiv*.

Keane, T.M., Creevey, C.J., Pentony, M.M., Naughton, T.J., and McInerney, J.O. (2006). Assessment of methods for amino acid matrix selection and their use on empirical data shows that ad hoc assumptions for choice of matrix are not justified. *BMC Evol. Biol.* 6, 29.

Kelekna, P. (2009). *The horse in human history* (Cambridge University Press Cambridge).

Korneliussen, T.S., Albrechtsen, A., and Nielsen, R. (2014). ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics* 15, 356.

Kosintsev, P.A. (2015). Bone remains from the settlement Belkaragay I. In *Ancient Turgay and the Great Steppe: Part and Whole.*, A.Z. Beisenov, ed. (Kostanay - Almaty), pp. 142–

Kosintsev, P., and Kuznetsov, P. (2013). Comment on “The Earliest Horse Harnessing and Milking.” *Tyragetia. Serie Nouă* 405–408.

Kousathanas, A., and Keightley, P.D. (2013). A comparison of models to infer the distribution of fitness effects of new mutations. *Genetics* 193, 1197–1208.

Kovalev, A.A., Erdenebaatar, D., and Rukavishnikova, I.V. (2016). A ritual complex with deer stones at Uushigiin Uvur, Mongolia: composition and construction stages. *Archaeology, Ethnology and Anthropology of Eurasia* 44, 82–92.

Kristjansson, T., Bjornsdottir, S., Sigurdsson, A., Andersson, L.S., Lindgren, G., Helyar, S.J., Klonowski, A.M., and Arnason, T. (2014). The effect of the “Gait keeper” mutation in the DMRT3 gene on gaiting ability in Icelandic horses. *J. Anim. Breed. Genet.* 131, 415–425.

Lang, V. (2012). The Bronze and Early Iron Ages in Estonia. In Chapter 2. Settlement Sites and Settlement in the Late Bronze and Early Iron Ages, (University of Tartu Press), pp. 63–65.

Laurence, R. (1999). *The Roads of Roman Italy: Mobility and Social Change*.

Leroi-Gourhan, A. (1958). *L’art pariétal: langage de la préhistoire* (Editions Jérôme Millon).

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*.

Librado, P., Der Sarkissian, C., Ermini, L., Schubert, M., Jónsson, H., Albrechtsen, A., Fumagalli, M., Yang, M.A., Gamba, C., Seguin-Orlando, A., et al. (2015). Tracking the origins of Yakutian horses and the genetic basis for their fast adaptation to subarctic environments. *Proc. Natl. Acad. Sci. U. S. A.* 112, E6889–E6897.

Librado, P., Gamba, C., Gaunitz, C., Der Sarkissian, C., Pruvost, M., Albrechtsen, A., Fages, A., Khan, N., Schubert, M., Jagannathan, V., et al. (2017). Ancient genomic changes associated with domestication of the horse. *Science* 356, 442–445.

Liesau, C. (2005). Arqueozoología del caballo en la antigua Iberia. *Gladius* 25, 187–206.

Liesau, C. (2017). Fauna in Living and Funerary Contexts of the 3rd Millennium BC in Central Iberia. In *Key Resources and Sociocultural Developments in the Iberian Chalcolithic.*, M. Bartelheim, P. Bueno, and M. Kunst, eds. (Tübingen Publishing), pp. 107–128.

Lindgren, G., Backström, N., Swinburne, J., Hellborg, L., Einarsson, A., Sandberg, K., Cothran, G., Vilà, C., Binns, M., and Ellegren, H. (2004). Limited number of patrilineal lines in horse domestication. *Nat. Genet.* 36, 335–336.

Lippold, S., Matzke, N.J., Reissmann, M., and Hofreiter, M. (2011). Whole mitochondrial genome sequencing of domestic horses reveals incorporation of extensive wild horse diversity during domestication. *BMC Evol. Biol.* 11, 328.

Lira, J., Linderholm, A., Olaria, C., Brandström Durling, M., Gilbert, M.T.P., Ellegren, H.,

- Willerslev, E., Lidén, K., Arsuaga, J.L., and Götherström, A. (2010). Ancient DNA reveals traces of Iberian Neolithic and Bronze Age lineages in modern Iberian horses. *Mol. Ecol.* *19*, 64–78.
- Logvin, A.V., and Shevnina, I.V. (2015). Settlement Belkaragaj 1. In *Ancient Rugay and the Great Steppe: Part and Whole.*, A.Z. Beisenov, ed. (Kostanay - Almaty), pp. 122–144.
- Lõugas, L. (1997). Post-glacial development of vertebrate fauna in Estonian water bodies: a palaeozoological study. In *Dissertationes Biologicae Universitatis Tartuensis*, (Tartu University Press),.
- Mallory, J.P., and Adams, D.Q. (1997). *Encyclopedia of Indo-European Culture* (Taylor & Francis).
- Marciniak, S., and Perry, G.H. (2017). Harnessing ancient genomes to study the history of human adaptation. *Nat. Rev. Genet.* *18*, 659–674.
- Martin, S.H., Davey, J.W., and Jiggins, C.D. (2015). Evaluating the Use of ABBA–BABA Statistics to Locate Introgressed Loci. *Mol. Biol. Evol.* *32*, 244–257.
- Mashkour, M., and Mohaseb, F.A. (2019). An impressive collection of equid remains in Shahr-I Qumis (Semnan- Iran). Preliminary archaeozoological results. *Journal of Iranian Archaeology* 22–32.
- McCue, M.E., Valberg, S.J., Miller, M.B., Wade, C., DiMauro, S., Akman, H.O., and Mickelson, J.R. (2008). Glycogen synthase (GYS1) mutation causes a novel skeletal muscle glycogenosis. *Genomics* *91*, 458–466.
- McDonald, J.H., and Kreitman, M. (1991). Adaptive protein evolution at the Adh locus in *Drosophila*. *Nature* *351*, 652–654.
- McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., et al. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* *20*, 1297–1303.
- Meyer, M., and Kircher, M. (2010). Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb. Protoc.* *2010*, db.prot5448.
- Miller, B., Allard, F., Erdenebaatar, D., and Le, C. (2006). A Xiongnu tomb complex: Excavations at Gol Mod 2 Cemetery, Mongolia (2002--2005). *Mongolian J. Anthropol. Arc. Ethnol.* *2* 1–21.
- Negahbān, E.O. (1974). Sagzabad Excavation Report. *Iran* *12*, 216.
- Nieto Espinet, A. (2016). Seguint les traces de la transhumància. Aproximació teòrica a partir dels resultats arqueozoològics de la fortalesa dels Vilars (Arbeca, Garrigues). *Revista D'arqueologia de Ponent*, 2016. Núm. 26, P. 11-34.
- Olalde, I., Brace, S., Allentoft, M.E., Armit, I., Kristiansen, K., Booth, T., Rohland, N., Mallick, S., Szécsényi-Nagy, A., Mittnik, A., et al. (2018). The Beaker phenomenon and the genomic transformation of northwest Europe. *Nature* *555*, 190–196.

- Onar, V., Pazvant, G., Alpak, H., İnce, N.G., Armutak, A., and Kiziltan, Z.S. (2013). Animal skeletal remains of the Theodosius harbor: general overview. *Turkish Journal of Veterinary and Animal Sciences* 37, 81–85.
- Orlando, L., Ginolhac, A., Zhang, G., Froese, D., Albrechtsen, A., Stiller, M., Schubert, M., Cappellini, E., Petersen, B., Moltke, I., et al. (2013). Recalibrating Equus evolution using the genome sequence of an early Middle Pleistocene horse. *Nature* 499, 74–78.
- Outram, A.K., Stear, N.A., Bendrey, R., Olsen, S., Kasparov, A., Zaibert, V., Thorpe, N., and Evershed, R.P. (2009). The earliest horse harnessing and milking. *Science* 323, 1332–1335.
- Paradis, E., Claude, J., and Strimmer, K. (2004). APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics* 20, 289–290.
- Patterson, N., Moorjani, P., Luo, Y., Mallick, S., Rohland, N., Zhan, Y., Genschoreck, T., Webster, T., and Reich, D. (2012). Ancient admixture in human history. *Genetics* 192, 1065–1093.
- Pearson, J.C., Lemons, D., and McGinnis, W. (2005). Modulating Hox gene functions during animal body patterning. *Nat. Rev. Genet.* 6, 893–904.
- Pedersen, C.-E.T., Lohmueller, K.E., Grarup, N., Bjerregaard, P., Hansen, T., Siegismund, H.R., Moltke, I., and Albrechtsen, A. (2017). The Effect of an Extreme and Prolonged Population Bottleneck on Patterns of Deleterious Variation: Insights from the Greenlandic Inuit. *Genetics* 205, 787–801.
- Pickrell, J.K., and Pritchard, J.K. (2012). Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genet.* 8, e1002967.
- Pollard, K.S., Hubisz, M.J., Rosenbloom, K.R., and Siepel, A. (2010). Detection of nonneutral substitution rates on mammalian phylogenies. *Genome Res.* 20, 110–121.
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., de Bakker, P.I.W., Daly, M.J., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81, 559–575.
- Ramsey, C.B. (2009). Bayesian Analysis of Radiocarbon Dates. *Radiocarbon* 51, 337–360.
- Rasmussen, M., Anzick, S.L., Waters, M.R., Skoglund, P., DeGiorgio, M., Stafford, T.W., Jr, Rasmussen, S., Moltke, I., Albrechtsen, A., Doyle, S.M., et al. (2014). The genome of a Late Pleistocene human from a Clovis burial site in western Montana. *Nature* 506, 225–229.
- Rohland, N., Harney, E., Mallick, S., Nordenfelt, S., and Reich, D. (2015). Partial uracil-DNA-glycosylase treatment for screening of ancient DNA. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 370, 20130624.
- Schmid, D., Peter, M., and Deschler-Erb, S. (2011). Crise, culte et immondices: le remplissage d'un puits au 3ème siècle à Augusta Raurica. In *L'Empire Romain En Mutation - Répercussions Sur Les Villes Dans La Deuxième Moitié Du 3e Siècle*, R. Schatzmann, and S. Martin-Kilcher, eds. (éditions monique mergoïl montagnac), pp. 125–132.
- Schubert, M., Jónsson, H., Chang, D., Der Sarkissian, C., Ermini, L., Ginolhac, A.,

- Albrechtsen, A., Dupanloup, I., Foucal, A., Petersen, B., et al. (2014a). Prehistoric genomes reveal the genetic foundation and cost of horse domestication. *Proc. Natl. Acad. Sci. U. S. A.* *111*, E5661–E5669.
- Schubert, M., Ermini, L., Der Sarkissian, C., Jónsson, H., Ginolhac, A., Schaefer, R., Martin, M.D., Fernández, R., Kircher, M., McCue, M., et al. (2014b). Characterization of ancient and modern genomes by SNP detection and phylogenomic and metagenomic analysis using PALEOMIX. *Nat. Protoc.* *9*, 1056–1082.
- Schubert, M., Lindgreen, S., and Orlando, L. (2016). AdapterRemoval v2: rapid adapter trimming, identification, and read merging. *BMC Res. Notes* *9*, 88.
- Schubert, M., Mashkour, M., Gaunitz, C., Fages, A., Seguin-Orlando, A., Sheikhi, S., Alfarhan, A.H., Alquraishi, S.A., Al-Rasheid, K.A.S., Chuang, R., et al. (2017). Zonkey: A simple, accurate and sensitive pipeline to genetically identify equine F1-hybrids in archaeological assemblages. *J. Archaeol. Sci.* *78*, 147–157.
- Sheikhi Seno, S., Mashkour, M., and Sardari, A. (2011). Subsistence Economy of the Lapui settlement of Tepe Mehr Ali Fars on the basis of the archaeozoological analysis. *Journal of Iranian Archaeology* *2*, 39–60.
- Skylitzes, J., and Wortley, J. (2010). *John Skylitzes: A Synopsis of Byzantine History, 811–1057: Translation and Notes* (Cambridge University Press).
- Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* *30*, 1312–1313.
- Uerpmann, H.-P. (1990). Die Domestikation des Pferdes im Chalkolithikum West- und Mitteleuropas, mit 10 Textabbildungen. *Madriider Mitteilungen* 109–153.
- Vieira, F.G., Lassalle, F., Korneliussen, T.S., and Fumagalli, M. (2016). Improving the estimation of genetic distances from Next-Generation Sequencing data. *Biol. J. Linn. Soc. Lond.* *117*, 139–149.
- Vukičević, T.T., Papeša, A.R., Alić, I., Kabalin, A.E., Ostović, M., and Kužir, S. (2017). Contribution to understanding Avar burials with equids in Croatia: detailed archaeozoological analysis. *Revue de Médecine Vétérinaire* *168*, 73–80.
- Wade, C.M., Giulotto, E., Sigurdsson, S., Zoli, M., Gnerre, S., Imsland, F., Lear, T.L., Adelson, D.L., Bailey, E., Bellone, R.R., et al. (2009). Genome sequence, comparative analysis, and population genetics of the domestic horse. *Science* *326*, 865–867.
- Wallner, B., Palmieri, N., Vogl, C., Rigler, D., Bozlak, E., Druml, T., Jagannathan, V., Leeb, T., Fries, R., Tetens, J., et al. (2017). Y Chromosome Uncovers the Recent Oriental Origin of Modern Stallions. *Curr. Biol.* *27*, 2029–2035.e5.
- Wang, J., Vasaikar, S., Shi, Z., Greer, M., and Zhang, B. (2017). WebGestalt 2017: a more comprehensive, powerful, flexible and interactive gene set enrichment analysis toolkit. *Nucleic Acids Res.* *45*, W130–W137.
- Warmuth, V., Eriksson, A., Bower, M.A., Cañon, J., Cothran, G., Distl, O., Glowatzki-Mullis, M.-L., Hunt, H., Luís, C., do Mar Oom, M., et al. (2011). European domestic horses

originated in two holocene refugia. *PLoS One* 6, e18194.

Williamson, R.J., Josephs, E.B., Platts, A.E., Hazzouri, K.M., Haudry, A., Blanchette, M., and Wright, S.I. (2014). Evidence for widespread positive and negative selection in coding and conserved noncoding regions of *Capsella grandiflora*. *PLoS Genet.* 10, e1004622.

Wilson Sayres, M.A., Lohmueller, K.E., and Nielsen, R. (2014). Natural selection reduced diversity on human y chromosomes. *PLoS Genet.* 10, e1004064.

Wutke, S., Sandoval-Castellanos, E., Benecke, N., Döhle, H.-J., Friederich, S., Gonzalez, J., Hofreiter, M., Lõugas, L., Magnell, O., Malaspinas, A.-S., et al. (2018). Decline of genetic diversity in ancient domestic stallions in Europe. *Sci Adv* 4, eaap9691.

Xu, X., and Arnason, U. (1994). The complete mitochondrial DNA sequence of the horse, *Equus caballus*: extensive heteroplasmy of the control region. *Gene* 148, 357–362.

Yi, X., Liang, Y., Huerta-Sanchez, E., Jin, X., Cuo, Z.X.P., Pool, J.E., Xu, X., Jiang, H., Vinckenbosch, N., Korneliussen, T.S., et al. (2010). Sequencing of 50 human exomes reveals adaptation to high altitude. *Science* 329, 75–78.

Youn, M., Kim, J.C., Kim, H.K., Tumen, D., Navaan, D., and Erdene, M. (2007). Dating the Tavan Tolgoi Site, Mongolia: Burials of the Nobility from Genghis Khan's Era. *Radiocarbon* 49, 685–691.

FIGURE TITLES AND LEGENDS

Figure 1. Equine archaeological remains.

(A) Location of archaeological sites. Pie charts are proportional to the total number of specimens providing DNA data compatible with the determination of sex, species and hybrid status. The names and temporal ranges (years ago) of the sites where hybrids and non-caballine species could be genetically identified are indicated.

(B) Temporal distribution of ancient specimens. Eight individuals showing uncertain age determination are not included. See Tables S1-S4 for further information.

Figure 2. Genetic diversity patterns.

(A) Nucleotide diversity (π) estimates and Y-to-autosomal π ratio per equestrian culture. The dashed red line indicates Y-to-autosomal π ratios of 0.25, corresponding to the expected ratio under even male reproductive success.

(B) Autosomal and Y-chromosome π estimates through time. See Figure S2E for more details.

(C) Individual error-corrected heterozygosity estimates. Only transversions were considered to minimize the impact of post-mortem DNA damage. See also Figures S1-S2.

(D) Conservative individual mutational loads from homozygous sites. Violin plots contrast the heterozygosity levels and genetic loads present in ancient (pink) and modern (blue) genomes belonging to the DOM2 lineage. See Figure S3, and Table S5 for further information.

Figure 3. TreeMix phylogenetic relationships.

The tree topology was inferred using a total of ~16.8 million transversion sites and disregarding migration. The name of each sample provides the archaeological site as a prefix,

and the age of the specimen as a suffix (years ago). Name suffixes (E) and (A) denote European and Asian ancient horses, respectively. See Table S5 for dataset information.

Figure 4. Selection targets through time.

(A) Population Branch Statistics (PBS) along the genome of 17 Byzantine horses, compared to 11 Gallo-Roman and 11 Deer Stone horses. The underlying tree topology consists of three groups with sufficient data and representing pre-C7th/C9th horses in Asia and Europe and post-C7th/C9th horses descending from Sassanid Persians. We used non-overlapping 50 kb genomic bins, and genes underlying enrichment for functional categories related to vertebral changes are indicated. These include *Sf3b1* and seven HOXB/C genes. *Hoxc11*, *Hoxb7*, *Hoxb13* and *Hoxc12* are not annotated as related to vertebral modifications, but embedded within the two independent clusters of HOXB/C genes. The *MSTN* speed gene, one selection candidate in Byzantine horses, is also highlighted. See Figure S4 and Tables S6-S7 for further information.

(B) Temporal allele trajectories for six SNPs associated with racing performance and locomotion patterns.

(C) Variance in allele frequency over time for the 57 SNPs investigated, categorized according to their impact on racing performance, body conformation, diseases and coat-color variations. The red dashed line delimits the 95th percentile of the variance distribution obtained from all SNP positions segregating genome-wide. See Figures S5-S6 for the full list of the SNPs investigated.

Figure 5. Genetic affinities.

(A) Principal Component Analysis (PCA) of 159 ancient and modern horse genomes showing at least 1-fold average depth-of-coverage. The overall genetic structure is shown for the first

three principal components, which summarize 11.6%, 10.4% and 8.2% of the total genetic variation, respectively. The two specimens MerzlyYar_Rus45_23789 and Dunaujvaros_Duk2_4077 discussed in the main text are highlighted. See Figure S7 and Table S5 for further information.

(B) Visualization of the genetic affinities amongst individuals, as revealed by the struct-f4 algorithm and 878,475 f4 permutations. The f4 calculation was conditioned on nucleotide transversions present in all groups, with samples were grouped as in TreeMix analyses (Figure 3). In contrast to PCA, f4 permutations measure genetic drift along internal branches. They are thus more likely to reveal ancient population substructure.

(C) Population modelling of the demographic changes and admixture events in extant and extinct horse lineages. The two models presented show best fitting to the observed multi-dimensional SFS in momi2. The width of each branch scales with effective size variation, while coloured dashed lines indicate admixture proportions and their directionality. The robustness of each model was inferred from 100 bootstrap pseudo-replicates. Time is shown in a linear scale up to 120,000 years ago and in a logarithmic scale above.

Figure 6. Phylogenetic reconstructions based on uniparental markers.

Tip labels are respectively composed of individual sample names, their reference number as well as their age (years ago, from 2017). Red, orange, light green, green, dark green and blue refer to modern horses, ancient DOM2, Botai horses, Borly4 horses, Przewalski's horses and *E. lenensis*, respectively. Black refers to wild horses not yet identified to belong to any particular cluster in absence of sufficient genome-scale data. Clades composed of only Przewalski's horses or ancient DOM2 horses were collapsed to increase readability.

(A) Best Maximum Likelihood tree retracing the phylogenetic relationships between 270 mitochondrial genomes.

(B) Best Y-chromosome Maximum Likelihood tree (GTRGAMMA substitution model) excluding outgroup. Node supports are indicated as fractions of 100 bootstrap pseudoreplicates. Bootstrap supports inferior to 90% are not shown. The root was placed on the tree midpoint. See Table S5 for dataset information.

Figure 7. Influence of native Iberian horses within DOM2 domesticates.

(A) Estimates of native IBE ancestry in DOM2 horses, based on the fraction of polymorphisms shared between IBE and DOM2 horses relative to Botai and Borly4 horses, and the level of polymorphisms shared between two IBE horses relative to Botai and Borly4 horses. The ratio of these values approximates a minimal boundary for the fraction of genomic ancestry present in DOM2 genomes pertaining to IBE or a closely related lineage. Consistent estimates are retrieved when replacing Botai with Borly4 horses, a ~5,000 years-old group directly descending from Botai.

(B) Admixture tests. The f_4 -statistics in the form of (outgroup,(IBE,(DOM2,Botai-Borly4))) and (outgroup,(*E. lenensis*,(DOM2,Botai-Borly4))) are provided. Negative values indicate excess of shared derived polymorphisms between IBE (or *E. lenensis*) and DOM2. More negative values indicate a more likely contribution of IBE (than *E. lenensis*) into DOM2. Testing all DOM2 individual genomes provided negative values, except two samples (Saadjave_Saa1_1117 and Friesian_0296A_0), which are not represented and for which other unidentified ancestry components could be present.

SUPPLEMENTAL FIGURE TITLES AND LEGENDS

Figure S1. Distribution of error rates in modern and ancient horses following three filtering approaches, related to Figure 2.

(A) Histogram of overall error rate estimates for samples in dataset (a) (n=159), based on random sampling of a single allele per site. All horses were grouped into modern (turquoise) or ancient horses (red). Three approaches were used to estimate the individual error rates, “All sites” makes use of all sites for which sequencing data was present, “All sites except CpG’s” masks data in CpG contexts according to the reference genome, and finally, “Transversions Only” excludes all sites observed as transitions. The vertical line represents the maximal allowed error rate (0.0005) cut off for analyses including “All sites except CpGs”.

(B) Histogram of overall error rate estimates for samples in the DOM2 clade (dataset (b), n=126), with the same filtering procedure as in Figure S1A, but based on posterior genotype probabilities.

Figure S2. Individual heterozygosity levels estimated through time for the 126 horses in dataset (b), related to Figure 2.

Heterozygosity levels were computed by summing the posterior probability of carrying a heterozygote per site and individual. Sites with missing data were disregarded at the individual level.

(A) Individual heterozygosity levels. Color coding refers to the overall error rate per site using the corresponding filters, see section ‘*Individual heterozygosity*’ The top panel shows individual heterozygosity over time, using all covered genomic sites, masking CpG dinucleotide sites according the reference (middle panel), and excluding transitions (bottom

panel).

(B) Distributions of the individual heterozygosity levels. Violin plots with integrated box plots were traced in R and grouped into modern (0-400 years ago) and ancient (≥ 400 years ago).

(C) Individual error-corrected heterozygosity levels. Same as (A) but individual heterozygosity levels were corrected by subtracting the error rate per site. Panels represent the same filtering approaches as in Figure S2A.

(D) Distributions of the individual error-corrected heterozygosity levels.

(E) Autosomal and Y-chromosome nucleotide diversity π estimates through time. We stratified the samples in the Asian clade into four different time windows (0-400, 400-900, 900-2200, 2200- 500 years ago) following (Wutke et al., 2018). Heterozygosity was calculated for the “All except CpG sites” dataset.

Figure S3. Mutational loads: estimator assessment and correlation to purifying selection, related to Figure 2.

(A) Evaluating the mutational load estimator at homozygous sites.

(B) Evaluating the mutational load estimator at heterozygous sites.

For panels (A) and (B), mutations were classified as nearly neutral ($0 < N_e x s \leq 1$), slightly ($1 < N_e x s \leq 10$), mildly ($10 < N_e x s \leq 100$) or strongly deleterious ($1 < N_e x s \leq 10$), according to their simulated selection coefficient (s) and population size (N_e). The latter was determined in the ancestral population (referred to as P_{Anc}), prior to the corresponding demographic collapses to $N_e = 100$ (P₁₀₀ population), $N_e = 200$ (P₂₀₀) and $N_e = 500$ (P₅₀₀). Five individuals were sampled from each population, and labelled as I1-I5. The homozygous load increased with stronger population declines, especially due to the random fixation of

deleterious mutations that, in PAnc, were slightly deleterious (*ie.* in PAnc, they were less likely to be fixed because the efficacy of negative selection $N_e \times s$ was higher). Despite heterozygosity levels dropped following N_e reductions, the load *per* heterozygous site remained steady, indicating heterozygous load has limited power to identify demographic collapses.

(C) Inverse correlation between mutational loads estimated at homozygous sites in modern horse genomes and the accumulated strength of purifying selection over generations, estimated as described in STAR Methods

(D) Individual differences between non-synonymous (dN) and synonymous (dS) substitutions in DOM2 horses.

(E) Positive correlation between mutational loads estimated at homozygous sites and the difference between non-synonymous (dN) and synonymous (dS) substitutions.

Figure S4. Distributions of PBS scores (Yi et al., 2010), related to Figure 4.

We considered a 3-population tree consisting of 11 Bronze Age Deer Stone horses (representing the pre-C7th/C9th Asian group), 11 Gallo-Roman horses (pre-C7th/C9th European horses) and 17 Byzantine horses (post-C7th/C9th). All individual genomes showed above 1-fold depth-of-coverage (Tables S6-S7; STAR Methods). The genome-wide distributions of PBS scores in all three branches within 50kb sliding windows is shown in blue. Neutral distributions are calculated from 50,000 bootstrap pseudo-replicates at four-fold degenerate sites (red) and intergenic positions defined to be located at least five kilobases away from gene bodies (green). Conservative cut-offs retained to identify selection candidates are shown as horizontal lines.

Figure S5. Heatmap summarizing the genotypes of 159 ancient and modern horses at 57

SNPs associated with key phenotypic traits, related to Figure 4.

The 57 SNPs are grouped into coat-color variants, genetic diseases, racing and locomotion capabilities, and body size conformation. Similarly, horses are grouped based on their population or cultural context. Grey and red tiles denote horses that were heterozygous and homozygous for the allele associated with the phenotypic trait respectively. White cells indicate the horse carried two copies of the non-causative allele. Tiles with vertical crosses mark non-genotyped SNPs. Causative alleles supported by a single read were conservatively assumed to be absent, to mitigate potential sequencing errors.

Figure S6. Temporal allele trajectories for 57 SNPs underlying traits commonly selected by modern breeders, including key genomic loci associated with or causal for locomotion, conformation, coat-coloration phenotypes and diseases, related to Figure 4.

Figure S7. Outgroup-f3 statistics and TreeMix phylogenetic inference, assuming one migration edge, related to Figure 3 and Figure 5.

(A) Outgroup-f3 statistics considering 160 horses included in dataset (d). The f3-outgroup statistics were calculated in the form (X, Y; Outgroup), where X and Y represent pairwise combinations (n=12,561). The f3-outgroup statistic was computed including the outgroup species (*E. africanus somaliensis*) and with 50,757,656 nucleotide transversions, using an in-house C++ script.

(B) TreeMix phylogenetic relationships assuming one migration edge (weight=0.386). The tree topology was inferred using a total of ~16.8 million transversion sites. The name of each sample provides the archaeological site as a prefix, and the age of the specimen as a suffix (years ago). Name suffixes (E) and (A) denote European and Asian ancient horses, respectively. See Table S5 for further information on the datasets used in each analysis.

STAR METHODS

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by Lead Contact, Ludovic Orlando (ludovic.orlando@univ-tlse3.fr).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Following low-depth sequencing, a total of 278 ancient equids reached a nuclear DNA coverage higher than 0.01X and were investigated for sex and species identification using the methodology implemented in Zonkey (Table S1, (Schubert et al., 2017)). Additionally, 87 pure horses recovered from museum and/or private collections, palaeontological and archaeological sites spread across Eurasia, showing moderate to high endogenous DNA content 0.06-0.78 were selected for whole-genome sequencing at a depth-of-coverage higher than 1-fold. The following section describes the archaeological contexts associated with all equids sequenced in this study (Tables S1-S3). The full name of each specimen is composed of the excavation site, followed by the sample name and age (in years ago from 2017), as estimated from direct radiocarbon dating (Table S1) or inferred from the archaeological context.

Belgium (Goyet A1)

The sample Goyet_Vert311_35870, a proximal fragment of a metatarsus, was unearthed from the first bone horizon (A1) of the third cave of Goyet, which was excavated for the first time in 1868 by Edouard Dupont. The paleolithic Goyet cave is part of a larger cave system located in the Belgian Mosan basin. Cave bear, reindeer and wild horse remains are the most common represented animals at the site (Germonpré, 2004). Excavations also revealed samples Goyet_Vert293_UpperPalaeolithic, Goyet_Vert300_31750 and Goyet_Vert304_UpperPalaeolithic.

China

The sample Fengtai_Fen4_2820 originates from the multilayered dwelling of the Kayue site of Fengtai (Province Qinghai), which is located at the rim of a large valley and consists of two phases. An early phase consisting of mainly wooden houses, dated to ~1190-920 BCE and a later phase composed of mud brick constructions, dated to ~980-750 BCE. The presence of permanent houses and the substantial amount of remains of domesticated plant grains like wheat and barley found at the site indicate a relative advanced mixed agropastoral economy.

Croatia (Bapska, Nuštar, Otok)

The late Neolithic settlement of Bapska-Gradac is located in Eastern Croatia, 4.5 km south of the Danube river. The site consists of different layers that have been associated with the Sopot and Vinča cultures and potentially also the Starčevo culture based on tools and pottery finds. Radiocarbon dating of individual BapskaGradac_BAPSKA_1305 revealed that the remain was intrusive and dated to ~1305 years ago (C7th-C8th).

The late Avar period cemetery of Nuštar is located in continental Croatia and is dated to the C8th and C9th. Samples Nustar_4_1187 and Nustar_5_1187 were unearthed during a rescue excavation in 2011 from one of the only two graves containing human and horse remains. Both graves were found near the southwest edge of the cemetery and were oriented east-west. Within the burial, the human remains were located on the right side, leaving the horse remains on the left side. Numerous grave goods such as iron knives, bronze belts decorated with floral motifs and horse equipment were found. Some pathologies were detected on Nustar_5_1187, and are exclusive to the spine, including thoracolumbar transitional vertebrae, spondylitis ankylopoetica (first two lumbar vertebrae) and spondylosis chronica deformans on four thoracic vertebrae and one lumbar vertebra. Further morphological analysis showed that the horse skeleton found in grave 4 (Nustar_4_1187) was probably five and a half to six years old at the time of death and had a slightly larger height to the withers of 143 cm, whereas Nustar_5_1187 reached an age of *circa* seven years and had average withers height of 139 cm (Vukičević et al., 2017).

The sample Otok_OTOK16_1307 was found at the archaeological site of Otok-Gradina, near the city Vinkovci during an excavation carried out in 1970. Based on the grave goods, the site was dated to the late C8th-early C9th. Among the 22 graves found, there were only two burials (Grave 4 and 16) containing the remains of humans and horses. The sample used in this study was found within grave context 16 alongside a 40 to 50 years old male skeleton. The morphology of the bones identified the horse as a mare, which was confirmed genetically. Further morphological analyses estimated the withers height of the mare to 139 cm (Vukičević et al., 2017).

Estonia (Otepää hill-fort, Ridala, Saadjärve)

The sample Estonia_Ote2_1184 originates from the archaeological site of Otepää Hill-fort, which is located on an eponymous upland area in the south of Estonia. This site covers a time period ranging from the Iron Age to the Middle Ages. However, the majority of the archaeological bone material and artifacts found at the site are associated with the late Iron Age. The site produced numerous tools (needles, knife handles, spinning-whorls, combs), weapons (arrowheads), ornaments (tusks, pendants, brooch), toys (die, toggles) and other unidentified objects. Those artifacts have been mostly made out of bones and antler of domestic and wild animals, but teeth (mainly canines) were also used to manufacture pendants. The bone material found at Otepää Hill-fort consists of cattle, pig, elk, bird, sheep, goat and horse remains. Among the horse bone material, there were also two bones showing signs of processing.

The sample Ridala_Rid2_2717 was recovered from a fortified settlement on a moraine ridge close to the coastal zone of Saaremaa Island (west Estonia), which was at the time of the settlement a coastal island. Excavations were carried out in 1961 by Aita Custin and in 1963 by Artur Vassar. The archaeological site covers an area of around 4500 m² and was dated to the C8th-C7th. Until now, approximately one tenth (435 m²) of the area has been excavated. A total of 2020 bone fragments have been recovered, 75% of which belonging to domestic animals (sheep/goat, pigs, cattle, and horses) and 25% to wild animals (seals). Sheep and goat bone fragments are the most frequent while horse bones represent the least frequent of all faunal remains. The horse remains recovered from this site belong to eight different individuals. Two were foals, two were slaughtered before the age of three and the other four were between two and four and a half years old. The presence of remains from exclusively juvenile individuals suggests their use for food consumption (Lang, 2012)).

Sample Saadjarve_Saal_1117 was excavated in 1984 at the settlement site of Saadjärve in eastern central Estonia, which is located 17 km north of the city of Tartu. Next to the remains of elks, cattle, goats, sheep, pigs, foxes, beavers, water voles, pikes and

breams, fourteen horse remains were recovered together with a range of the remains of freshwater fish, such as perch and burbot (Lõugas, 1997) .

France (Beauvais: Maladrerie Saint-Lazare and rue de L'Isle-Adam, Boinville-en-Woëvre, Boves “Chemin de Glisy”, Capesterre, Chartres “Boulevard de la Courtille”, Evreux “Clos-au-Duc 3 rue de la Libération – 2007”, Longueil-Annel, Mâcon “Rue Rambuteau”, Metz “Place de la République”, Saint-Claude, Saint-Laurent Blangy “Actiparc 2002”, Vermand 2005 and Saint-Just-en-Chaussée)

La Maladrerie Saint-Lazare, located in Beauvais, Northern France, was a leper colony founded during the late C11th or early C12th. It remained in activity until the French revolution, when it was closed and then sold to the French State. Horse Beauvais_GVA122_417 was sampled from a petrosal bone excavated in 2013 from a latrine transformed into a waste pit (US 8057). The site from rue de L'Isle-Adam at Beauvais is a former convent, excavated in 1992, dating back to the C15th. A deep hole containing equine bones, including sample Beauvais_GVA375_567, was found next to the church of the convent. The bone remains include pieces of rachis and various dislocated anatomical parts.

Boinville-en-Woevre is an ancient Gallo-Roman villa, located in Meuse, France. Fifteen pits containing the remains of some large equids have been found in the *pars rustica* of the villa. Some of those pits contained several individuals, potentially buried simultaneously. In total, 22 individuals dating back to the C2nd and C3rd have been discovered, including Boinville_GVA125_1817, an 8-to-10-year-old male, genetically identified as a donkey (Schubert et al., 2017).

Boves « chemin de Glisy » corresponds to a large archaeological area of Northern France. Several settlements dated from Iron Age and Roman periods have been studied. The excavation of a large pit (6.80m deep), corresponding to a Gallo-Roman quarry (C3th) revealed several carcasses of animals (sheep, equids), including individual Boves_GVA191_1717.

The site of Roseau, located in Capesterre Belle-Eau (Guadeloupe), is associated with both some European remains dating back to the C16th-C17th, and some pre-colonial Amerindian remains dating back to the C11th-C15th. Individual Capesterre_LIS2_417 is associated with the European settlement.

Twenty-four horses sequenced in this study were excavated from the archaeological site of “Boulevard de la Courtille C277”, situated in the outskirts of Chartres, France. Hundreds of well-preserved scattered equids skulls were found lying on the grounds, probably employed to drain excessive water in the antic city of *Autricum* (present-day Chartres). These 24 individuals are Chartres_GVA1_1917, Chartres_GVA2_1917, Chartres_GVA4_1917, Chartres_GVA9_1917, Chartres_GVA26_1917, Chartres_GVA28_1917, Chartres_GVA36_1917, Chartres_GVA39_1917, Chartres_GVA43_1917, Chartres_GVA47_1917, Chartres_GVA48_1917, Chartres_GVA53_1917, Chartres_GVA56_1917, Chartres_GVA60_1917, Chartres_GVA64_1917, Chartres_GVA64_1917, Chartres_GVA68_1917, Chartres_GVA75_1917, Chartres_GVA78_1917, Chartres_GVA81_1917, Chartres_GVA95_1917, Chartres_GVA111_1917, Chartres_GVA112_1917 and Chartres_GVA115_1917.

The archaeological site of Démuin is located in Somme, France and revealed an occupation period extending from the C9th to the Renaissance. The equine bones represent 2.8% of more than 5,000 remains collected (Jonvel, 2014). Individuals Demuin_GVA401_917 and Demuin_GVA402_917 were sampled from isolated bones discovered in two grain silos transformed into trash.

The site “clos-au-Duc 3 rue de la liberation” in Evreux, France, is a funeral site dated to the C1st to C3rd. Several excavated pits contained parts of equid skeletons, including samples Evreux_GVA130_1817, Evreux_GVA132_1817, Evreux_GVA133_1817, Evreux_GVA135_1817 and Evreux_GVA140_1817. There does not seem to be any ritual connection between humans and equids since dead horses have been shown to represent waste. However, the site is connected to a rendering activity.

The archaeological site of Longueil-Annel, located in the middle valley of Oise, between the towns of Noyon and Compiègne, has been associated with different occupations throughout the Neolithic, Bronze Age and Iron Age. However, individual LongueilAnnel_GVA129_267 comes from a skeleton found in a modern pit dating back to the C18th.

The “Rue Rambuteau” site is situated in the outskirts of the antique city of Mâcon, Eastern France, and dates back to the C3rd. The excavated area is a large dirt quarry containing a great amount of animal remains, mainly equids, but also cattle, dogs and pigs. In total, 1,497 parts of dead equids were thrown in this waste area, representing at least 16 individuals, Macon_GVA201_1767.

Individual Metz_GVA321_492 was unearthed from the “Place de la République” site in Metz, France. The excavation, spreading over 1,375 m², has allowed the identification of numerous archaeological levels and remains, from the C1st to the C19th.

The excavations at Saint-Claude “Cité de la connaissance” (Basse-Terre, Guadeloupe), have revealed some pre-columbian occupations and some ancient sugar refineries. Some equine remains could be identified in levels dating back to the second half of the C18th, including SaintClaude_GVA381_242, sampled from a petrous bone belonging to a 7-to-10-year-old individual.

Actiparc at Saint-Laurent-Blangy, Northern France, is a large area of 300 hectares excavated as part of the construction of a craft activity area. Excavation campaigns revealed several types of settlements, including rural habitats, indigenous farms and a necropolis. The time periods represented range from the ancient La Tène period in the Iron age, to the Roman period (3rd century BCE – 4th century CE). The bones of horses Actiparc_GVA124_2143, Actiparc_GVA307_2127, Actiparc_GVA308_2312, Actiparc_GVA309_2302 and Actiparc_GVA311_2253 all come from waste pits.

The gallic sanctuary of “Les Rossignols” in Saint-Just-en-Chaussée (Oise, France), excavated in 1994-1995, was occupied from the final La Tène period in the Iron Age (D1-D2) up until the Roman High Empire. The most remarkable remains were ditches filled with horse bones with pieces of chariots and harnesses, and some human bones. Samples SaintJust_GVA212_2162, SaintJust_GVA219_2162 and SaintJust_GVA242_2250 were recovered from such ditches.

Saint-Quentin is an archaeological site in Aisne, France, from which a large number of bovine and equine samples have been recovered. However, the precise nature of the assemblages has not been defined yet, it is therefore unsure whether the site can be associated with some cultural activities or was rather used as a ditch. Individuals SaintQuentin_GVA237_1917 and SaintQuentin_GVA238_1917 were sampled from petrosal bones dated to the C1st-C2nd.

The excavations at Vermand, in the department of Aisne, Northern France, were carried out in 2015 and delivered remains of an ancient Roman way, as well as the remains of equines, including individual Vermand_GVA199_1742, which was excavated from structure 15 of the site.

Georgia (Dariali)

Tamara fort is situated in the Dariali Gorge of Northern Georgia, next to the Russian border, on a raised landform of the west bank of the Tergi river. It was occupied from the Sasanian period to the Medieval period. Excavations at the site indicate several occupations mainly between c. 400-1000 CE and a late reoccupation between the late C13th and early C15th. Individual Dariali_Georgia2_317 most likely dates back to this medieval period or a post-medieval one.

Germany (private collections, Schloßvippach)

The samples Mainz_Mzr1_1373 and FrankfurtHedderheim_Fr1_1863 were both sampled from a private collection of Prof. emer. Helmut Hemmer (Johannes Gutenberg-University Mainz). It was compiled between the 1960s to 1980s, mainly consisting of stray finds of single loose bones found during excavations of Roman sites of the Rhine-Main area, Germany. Sample Mainz_Mzr1_1373 is a calvarium found at the construction site of the Mainz University clinic and was radiocarbon dated to 1,373 years ago. Sample FrankfurtHedderheim_Fr1_1863 is a calvarium found in Frankfurt-Hedderheim (Nida) in a soil filled well shaft and was radiocarbon dated to 1,863 years ago.

Schloßvippach is an Early Bronze Age site located in Germany, dating back to 1600-2200 BCE and composed of a settlement of long dwelling houses and a burial site with a large number of graves. Excavations have revealed some ceramics, bronze tools and jewels, as well as animal bone remains, including the horse sample Schloßvippach_Svi6_3917. Addendum to (Gaunitz et al., 2018), regarding the Roman horse from Augsburg-Haunstetten (Haunstetten_1979): Archaeozoological analysis revealed a 5-8-year-old stallion with a withers height of 124±3 cm according to Kiesewalter. The specimen shows splint, a premature ankylosis between the inside splint bone and the cannon bone of the right foreleg. This was probably caused by too early and too much exercise. The teeth are showing cementum hypoplasias in the occlusal plane and transversal enamel hypoplasia near to the dentinoenamel-junction. The latter is indicating unspecific stress at the end of enamel formation.

Iceland (Berufjörður and Granastaðir)

The sample Berufjordur_VHR102_1067 was excavated in 1898 by Daniel Bruun and Brynjúlfur Jónsson at site of Berufjörður in Barðastrandasýsla (Westfjörds), Iceland and is dates to the Viking Age (ca. 850-1050 CE). Due to the incomplete documentation of the site the exact association of the horse and human burials at the site is not clear. Only a few horse teeth with fragments of a maxilla were kept from this particular horse burial, the horse was 5-to-7-year-old at death. The tooth sampled for this study was a maxillary molar. The tooth has been radiocarbon dated to cal. (2 σ) 890–1015 CE.

Individual Granastadir_VHR031_1057 was sampled from a maxillary molar from Granastaðir, an early Viking farmstead in Northern Iceland. The molar has been radiocarbon dated to cal. AD (2 σ) 895–1025. The site was excavated by Bjarni F. Einarsson between 1987 and 1991 (Einarsson, 1994). A collection of animal bones was recovered from the site both from a midden and from within the excavated buildings mostly representing domestic animals.

Iran (Belgheis, Kulian Cave, Sagzabad, Shahr-i-Qumis, Tepe Hasanlu, Tepe Mehr Ali)

The citadel of Belgheis is located three kilometers away from the modern city of Esfarayen, in North East Iran. It covers an area of 180 hectares and was occupied from the beginning of the Islamic period until the C18th. Individual Belgheis_TrBWBX116_485 sequenced in this study was sampled from a left third lower molar recovered from Tr BWBX116, Unit 12, Depth 515-235, dated to the Seljukid-Ilkhanid periods (C11th-C14th). Radiocarbon dating, however, further indicated that the specimen belonged to the C16th.

Kulian Cave is located near the city of Rawansar, about 52 km northwest of Kermanshah, in west Iran. The site contains Pleistocene and Holocene archaeological deposits (Biglari and Taheri, 2000). The cave is about 20 m long and consists of two chambers. A petrosal bone of a horse sequenced for this study, KulianCave_MV178_1694, was found in the inner chamber. Animal bone accumulation in this chamber most likely originated from carnivore activity and natural death. The equid belongs to a female individual and was dated to the time of reign of Sasanian king Shapur II (309–379 CE).

Samples Sagzabad_SAGS27_3117 and Sagzabad_SAGxPit22_3117 were excavated from the archaeological site of Sagzabad, located in the central district of Buin Zahra in the Qazvin plain, 140 km west of Tehran, Iran. Multiple archaeological campaigns have evidenced a continuous occupation from the Late Bronze Age to the Iron Age II (Negahbān, 1974). The large animal assemblage is composed of more than 10,000 identified bones and shows the importance of domestic herbivores ovi-caprines and cattle, followed by an important contribution of domestic equids.

Shahr-i-Qumis is a site in North East Iran, consisting of several isolated mounds spread across an area of 28 kilometers. It dates back to the Parthian and Sasanian periods, although some recent radiocarbon datings of faunal remains tend to show a longer period of occupation, from the 8th century BCE to the 8th century CE. The site has been identified as Hekatompylos, the capital of the Parthian Empire and major hub of the Silk Road and Great Khorasan Road. Excavations at Shahr-i-Qumis revealed a very large quantity of equine skeletons, including sample ShahrIQumis_AM115_1557 (Hansman and Stronach, 1970). The radiocarbon date obtained for this sample place it either during the kingdom of Yazdegerd II (438–457 CE) or his brother Peroz I (457–484 CE). At the beginning of the C5th, nomadic group and in particular the Hephthalites or White Huns attacked Persia several times, invading parts of eastern Persia for several years. These events may have had also an impact of the equine population. A large set of equine bones from Shahr-i-Qumis has been studied during these last years at the British Institute of Persian studies in Tehran and currently a morphometric geometric project is ongoing on this material (Mashkour and Mohaseb, 2019).

Tepe Hasanlu is a fortified site located in Solduz Valley of Western Azerbaijan province, Northwestern Iran. The site was occupied from the Late Neolithic to the Iron Age and consisted of two distinct parts: a High Mound and a Low Mound (Dyson, 1989). A total of four individuals sequenced in this study originate from this site. Two horse samples were recovered from the citadel of Iron Age II (1,050-800 BCE) associated with the Mannaeans kingdom and destroyed by the Urartians during a battle around 800 BCE. While individual TepeHasanlu_3461_2930 was unearthed from a rough soil deposit, TepeHasanlu_3394_2808 was found together with thousands of artifacts and faunal remains, within the deposit and collapse of buildings, likely used as horse stables (Dyson, 1989). Individuals TepeHasanlu_1140_2682, TepeHasanlu_3459_2667 and TepeHasanlu_V31E_2667 date back to the Urartian occupation period that followed the destruction of the citadel. After a hiatus, period IIIa related to the Achaemenid Dynasty (550-330 BCE), for which no substantial architectural remains have been found. Period II is also a debated issue but

generally assigned to the Seleucid or Parthian period, post-Achaemenid (Dyson, 1999). These historical periods were very short at Hasanlu, chronologically between 400 to 270 BCE. Four samples, including TepeHasanlu_2327_2352, TepeHasanlu_2529_2352, TepeHsanlu_2689_2352 and TepeHasanlu_3398_2352, belong to this Historic Era.

Tepe Mehr Ali is located in the province of Fars, Southwest of Iran. The site belongs to the Lapui culture, dated to the Chalcolithic (6th-4th mill. BCE), and shows an over-representation of domestic animals such as cattle, sheep and goats but also a significant number of wild herbivores, such as hemiones and gazelles (Sheikhi Seno et al., 2011). Individual TepeMehrAli_Trj12x31_CopperAge was genetically identified as a pure hemione specimen (Schubert et al., 2017).

Kazakhstan (Belkaragay, Halvai)

Belkaragay is a Copper Age site located in the Kostanay region, Kazakhstan. The area of excavation spreads over 1,000 m² and is composed of ten house-like structures, in which many different animal bones were discovered. Those were attributed to various species, including the wolf, the Saiga antelope, the fox, the hemione and the horse, such as samples Belkaragay_NB13_CopperAge and Belkaragay_NB15_CopperAge included in this study (Kosintsev, 2015; Logvin and Shevnina, 2015).

Sample Halvai_KSH4_4017 was excavated from Kurgan Halvai 5 (pit number 4), which located on the left bank of the Tobol branch of the Karatomar Reservoir in Northern Kazakhstan (Kostanay Region), located 500 meters to the north-east of the Sintashta kurgan Halvai 3. The kurgan was 30 meters in diameter and 80 cm in height. Pit number 4, which is associated with the Sintashta culture of the Bronze Age, was located directly in the center of the kurgan. The horse skull was found close to the edge of the southern wall of the pit. In addition to the horse skull, the pit also contained human remains, belonging to a female, and other grave goods, such as a zoomorphic stone altar, stone tips, pebble fragments and fragments of a vessel. Another individual, Halvai_KSH5_2542, was excavated from Kurgan Halvai 3 (pit number 8A). The horse skeleton was found together with the remains of an approximately 50 years old woman and a sheep. Based on the position of the skeleton and stratigraphic information, the burial can be assumed to have been constructed during the Early Iron Age.

Kyrgyzstan (Boz-Adyr)

The Boz-Adyr burial ground is located on the slope of the Ak-Bakshy mountain range, located in the Issyk-Kul region, Kyrgyzstan. The burials are located in mounds, which is characteristic for the funeral traditions of the nomadic and semi-nomadic populations of Tien Shan and Semirechye, especially during the C12th-C15th. During the excavation in 2014, three burial mounds showing next to human remains also the skeletons of horses were discovered (burial mounds number 10, 16 and 19). The burial rites associated with those graves are characteristic for the Turkic period of the C6th-C9th.

Burial mound 16 was a swampy rock-earthen embankment of a circular shape with a diameter of 5 meters. At a depth of 150-160 cm, the remains of a decapitated adult male, accompanied by a horse, were discovered. The skeleton of horse sample BozAdyr_KYRH8_1267 was supported by two boulders, with its legs bent and the head pointing west. The neck was bend facing north. Alongside with the horse remains, bit wear, iron stirrups and the remains of a wooden saddle were also discovered.

Burial mound number 19 is located 15 meters to the west of burial mound number 16. At a depth of 140 cm, the undisturbed remains of a human and horse,

BozA dyr_KYRH10_1267, were discovered. The human skeleton was found lying on its back with the head pointing to the east, and the legs bent pointing to the right. The horse was positioned on it, with the abdomen turned to the right and the head pointing to the west.

Lithuania (Marvelė cemetery)

Marvelė is a large medieval cemetery located in Kaunas, Lithuania, and dates back to the C8th to the C11th. It consists of 211 human burials that also contain about 250 horse remains, either whole skeletons, head and forelegs only or scattered remains (Bertašius and Daugnora, 2001). The vast majority of buried horses are young adults under ten years and can be associated with some ritual offerings, which seemed to be common in the Baltic Aukštaičiai people during that era. Individuals Marvelė_01_1138, Marvelė_02_1138, Marvelė_05_1138, Marvelė_16_1138, Marvelė_18_1189, Marvelė_21_1087, Marvelė_22_1138, Marvelė_27_1138 and Marvelė_32_1144 were all excavated from Marvelė cemetery.

Moldova (Miciurin)

The horse sample Miciurin_Mic2_3267 was excavated from the Bronze Age site of Miciurin, Moldova (1500-1000 BCE). The discovery of some slag fragments and some mill remains indicate the presence of Bronze craftsmanship, and to some extent mill cereal agriculture at Miciurin.

Mongolia (Gol Mod II, Khatuu 2, Olon-Kurin-Gol (Olon Guuriin Gol), Uushgiin Uvur, Talvan Tolgoi, Khotont)

The Gol Mod II site is a cemetery located in Central North Mongolia, north of the Khangai mountains, that dates back to the 3rd century BCE to the 1st century CE (Miller et al., 2006). All sequenced horses from this site come from Grave #1, which apparently belonged to an aristocratic figure of the Xiongnu period. With a length of 86 meters, the whole burial structure is one of the largest burial structures of Central Asia. The grave has been robbed, but still yields with rather astonishing bronze, glass and gold artifacts, which suggests a burial ritual for kings of the Xiongnu period. At a depth of 10 meters, remains of 25 different horses were excavated, including GolModII_Mon23_2007, GolModII_Mon24_1993, GolModII_Mon25_2011, GolModII_Mon26_1995 and GolModII_Mon27_2011, which were all sampled from petrous bones.

Khatuu 2 is an Iron Age site located in the Mongolian Altai, associated with the nomad culture of Pazyryk and dating back to the 4th and 3rd centuries BCE. Sample Khatuu_Kha2_t1_2312 was excavated from a tomb that contained the skeletons of one human and one horse.

The undisturbed and permafrozen Pazyryk burial complex of Olon-Kurin-Gol (Olon Guuriin Gol) was excavated in the summer of 2006 by a German-Mongolian-Russian expedition team. The Kurgan is located close to the upper Olon-Kurin-Gol River on the southern slope of the Saylyugem Mountains, Mongolian Altai. The grave contained the partially-mummified, fully-dressed remains of a Scythian warrior and two horse skeletons. The samples OlonKurinGol_OKG1_2367 and OlonKurinGol_OKG2_2367 originate from such horses. Along with the horse remains, the harness and saddle of the horses were found.

The individuals UushgiinUvur_Mon37_3085, UushgiinUvur_Mon39_3085, UushgiinUvur_Mon40_3085, UushgiinUvur_Mon41_3085, UushgiinUvur_Mon42_3130, UushgiinUvur_Mon43_3120, UushgiinUvur_Mon44_3085, UushgiinUvur_Mon45_3080,

UushgiinUvur_Mon79_3085, UushgiinUvur_Mon87_3117 and UushgiinUvur_Mon89_3085 were all sampled from petrosal bones excavated from the Uushgiin Uvur site, a large Bronze Age complex of deer stones and ritual mounds, located in front of the Ulaan Uushig Mountain, Mongolia. There are about 30 deer stones, arranged in a row from north to south, surrounded by many ritual structures with solid or ring-shaped rock coverings. Mongolian-Russian archaeologists within the Uushgiin Uvur – 2013 Project, run by the Department of Archaeology of Ulaanbaatar University of Mongolia and the Institute of Archaeology, Russian Academy of Sciences, excavated a ritual complex with deer stones (Kovalev et al., 2016). All deer stone statues are made of light gray, reddish brown or blue-gray granite. Based on the horse bones, the Uushgiin Uvur Deer Stone Complex was dated to 1,312-810 BCE.

Tavan Tolgoi is a steppe cemetery located in South East Mongolia, associated with the Mongolian Empire and dated to 1,206-1,368 CE (Youn et al., 2007). It is characterized by a post and unique stone carvings that indicate the presence of no less than 14 graves. Eight graves were excavated in 2004 and 2005 by a team from the National University of Mongolia and revealed artifacts dating to the Mongolian Empire period. Grave 4 contained horse stirrups, a birch bark quiver, male adult human remains as well as sheep and horse cranial bones, from which the genomes of TavanTolgoi_GEP13_730, TavanTolgoi_GEP14_730 and TavanTolgoi_GEP21_730 were sequenced. In 2012, a Chinese-Mongolian team of archaeologists excavated a mound on the top of the mountain Bayantsogt, located in Khotont Soum in the Arkhangai Aimag province. The site is elevated at 1.75 meters above sea level. The mound itself is approximately 2.6 meters tall with the basal diameter of about 26 meters and shows a 24 meters-long and six meters-wide ridge on its southeast face. Graverobbers have left a hole starting from the top. The excavation yielded wooden artifacts, sheep and/or goat bones, along with some horse bones, including a petrosal bone of individual Khotont_UCIE2012x85_1291. The external features and internal organization of the mound indicate that it belonged to ancient Uyghurs (611-840 CE).

Poland (Bruszcewo)

Individual Bruszcewo_Bru4_3917 was excavated from the Early Bronze Age site of Bruszcewo, located in Poland around 60 kilometers south of Poznan and dating back to 2,200-1,600 BCE.

Portugal (Santarém)

Alcáçova of Santarém is placed in the right bank of the lower Tagus estuary, it is an important and large site (50,000 m²) located in a plateau overlooking the river, where an extensive diachronic occupation (since Late Bronze Age until nowadays) has been identified. The sample 254, an inferior premolar 4, was recovered in the excavations carried out in 2001, at layer [14], corresponding to the Islamic occupation, according to the ceramic material, well dated to the C11th-C12th.

Russia (Altata, Arzhan II, Balagansk, Bateni – Karasuk, Derkul, Kokorevo, Krasnaya Gorka, Lebyanzhinka IV, Merzly Yar, Oktyabrsky, Potapovka I, Sayangorsk, Sintashta)

Altata is a settlement located in the Saratov region, Russia, encompassing a total area of 100 to 200 m², from which different artifacts associated with the Neolithic could be recovered. Animal remains found at this site have been identified aurochs, foxes and horses, including sample Altata_NB31_Neolithic.

The Scythian burial mound or the so-called kurgan of Arzhan represents the youngest of its kind and can be associated with the Aldy-Bel culture, which dates back to the boundary of the 7th and 6th century BCE. The elite funeral complex is located only nine kilometers away from Arzhan I in the Uyuk hollow, from which ArzhanI_Arz3_2767 was excavated. The undisturbed kurgan is 80 meters in diameter and 2 meter-high and consists of 27 graves, from which individuals ArzhanII_Arz15_2642 and ArzhanII_Arz17_2642 were unearthed. Additionally, a special burial including fourteen harnessed horses was found during the excavation expedition from 2001 to 2003, which includes individuals ArzhanII_Rus9_2500 and ArzhanII_Rus11_2500.

Individual Balagansk_Rus19_2017 was excavated from Balagansk, a village located close to the Angara river in the Irkutsk region, flooded by the Bratsk reservoir. It represents a Central Asian trading settlement associated with the Ust-Talka culture.

The sample Bateni_Rus16_3318 was retrieved from a right distal metatarsal bone and belongs to the Late Bronze Age Karasuk culture (1,500-800 BCE), which followed the Andronovo culture in the South of Siberia. It covered an area from the Aral Sea to the Yenisei river on the East and to the Altai mountains and Tien Shan in the South. Karasuk communities are known to be farmers who practiced a mixture of agricultural and stockbreeding of cattle, sheep and horse. They are especially known for their metallurgy, in particular for their daggers and knives (Mallory and Adams, 1997). The bone sample was excavated close to the Bateni settlement, Republic of Khakassiya. The excavation site is not accessible anymore due to the flooding of the Krasnoyarsk Reservoir.

Derkul is a settlement dating back to the Neolithic, situated in the Orenburg region of Russia, and from which only one house-like structure has been excavated. Both Neolithic artefacts and animal bones, including remains of individuals Derkul_NB2_Neolithic and Derkul_NB4_Neolithic, could be unearthed from this small site.

Kokorevo I is a late Upper Palaeolithic site associated with the Kokorevo culture, located by the Upper Yenisey river, Russia. The horse sample Kokorevo_Rus3_14450 was found in level 4, and is thus thought to be slightly older than level 3, dated to 14,450 years BP. This Pleistocene horse most likely represents a wild individual.

Horse sample KrasnayaGorka_Rus48_1446 was sampled from a left humerus unearthed from the Krasnaya Gorka burial mound, Republic of Tyva, Russia, and dating back to the C6th. No further information could be recovered regarding this site.

Individual LebyazhinkaIV_NB35_Neolithic was unearthed from the settlement of Lebyazhinka IV, Samara region, Russia. Based on excavated artifacts, this site has been associated with the Neolithic and the Copper Age. Animal bone remains found on site have been associated with beavers, horses, Siberian roe deers, aurochs and elks.

The Paleolithic horse MerzlyYar_Rus45_23789 was sequenced from bone fragments recovered from an outcrop in the Todza depression, close to the village of Seiba, Republic of Tuva, Russia. It is not associated with any archaeological context.

Oktyabrsky is a village located on the lower Volga river, in Ustin district, Kalmyk Republic of Kalmyk, Russia. The burial ground is associated with the Sarmat culture, but as many steppe cemeteries, it contains a number of additional younger graves. Individual Oktyabrsky_Rus37_830 was sampled from a left metatarsal bone of a complete horse carcass, unearthed from grave 1 of the burial mound number 16. A second individual, Oktyabrsky_Rus38_659, was unearthed from grave 1 of burial mound 3. These burial grounds are associated with Sarmatian time in general, but some bone remains, including the two individuals included in this study, date back to more recent times, namely the early post-Khazar and the late Golden Horde periods.

Sintashta is a late Bronze Age archeological site located in Chelyabinsk Oblast, Russia, and is associated with the eponymous Sintashta culture. This site includes five

cemeteries and no less than 40 graves. The horse samples Sintashta_NB44_3577 and Sintashta_NB45_3577 were recovered from grave number 19, together with a chariot and two other horses.

Individual PotapovkaI_1_3900 was sampled from the left mandible of mare excavated from Kurgan 3 of the settlement of Potapovka I, by the Sok River, in the Samara region, Russia. This site is associated with the Bronze Age Potapovka culture and has revealed a large number of pieces of metalwork. The horse skull was placed over the body of a decapitated woman, potentially as an offering.

The city of Sayanagorsk in the Republic of Khakassiya is located in the south part of the Minusinsk basin at the left bank of the Sayan Mountains. Sample Sayangorsk_Rus41_2677 originates from the right tibia of a horse which was unearthed during the excavations from the site Ai-Dai 1 (mound 3, grave 1). This site is associated with the Tagar culture (7th to 2nd centuries BCE) in South Siberia, which was preceded by the Karasuk culture.

Slovakia (Sebastovce)

Individual Sebastovce_131_1317 was sampled from a tooth, excavated from the medieval Avarian_Slavonic cemetery of Šebastovce, Slovakia, dating back to the C7th-C8th.

Spain (Camino de las Yeseras, Cantorella, Capote, El Acequi3n, Els Vilars)

The large settlement of Camino de las Yeseras, located in San Fernando de Henares, Madrid, Spain, was occupied between the beginning of the 3rd and the 2nd millennia BCE. Domestic structures such as hut pits have been associated with both pre-Beaker phase and Bell-Beaker culture. Because of its remarkable size and strategic location - near rivers and flint mines, it is thought to have played a major role in central Iberia (Blasco et al., 2011). The central area of Camino de las Yeseras is a big structure with a sunken floor of *circa* 600 m², with more than 50 adjunct structures and a stratigraphy of about 2 meters. One of the excavated units yielded 1772 faunal remains, including both domestic and wild animals. The horse CaminoDeLasYeseras_CdY2_4678 was excavated from the central area of the site and radiocarbon dated to 2,861-2,496 BCE, which is associated with the Pre-Bell-Beaker phase of the site (Liesau, 2017).

The Cantorella settlement is situated in the Corb valley, Catalonia, Spain, discovered in March 2010 and dated to 3,670-1,800 BCE. It was inhabited twice in prehistory, once during the Final Neolithic - chalcolithic and once during the Bronze Age, and a distance of 200 meters separates the two settlements. Although Cantorella faunal archaeological records are scarce compared to Iron Age Catalan sites, horse remains are very abundant relative to other animals, especially in the Late Neolithic - Chalcolithic occupancy (Abad et al., 2011). A number of 15 silos of that period revealed the presence of *Equus sp.*, which might represent an autochthonous species of horse. Horse remains were present in approximately 50% of the final Neolithic structures but absent from all but two Bronze Age silos. Cantorella_UE2275x2_4791 was sampled from a petrous bone unearthed from silo SJ-191 that contained two horse skulls from Final Neolithic - Chalcolithic period.

The sample Capote_Cap102_2464 is an upper molar tooth. It was found in a bone depot, inside of a house at the hill-fort of Castrej3n de Capote, a Celtic fortified village of the Iberian Late Iron Age (4th-1st century BCE). This house, "HE-A", is an archetypical household of two rooms, the first bigger and used for several functions (cooking, eating, etc.) and the second and smaller, for storing and sleeping. In this pattern of household, the first room used to be a hearth and a quern, and in this room, there is a central big hearth that was stratigraphically dated to the middle of 2nd century BCE. Below the layer sequence of this depot, there is an older sequence dated from the 4th century to the first half of the 2nd century

BCE, and above, a later sequence from the second half of the 2nd century BCE to the first half of the 1st century BCE. The three sequences show similar household remains but only the intermediate one has faunal remains, mainly from horses, cows and pigs. Although this hill-fort is well-known in Celtic archaeology for an important very well-preserved shrine, where collective banquets were held by the time of the sequence IIb, the charcoal and bones remains of HE-A are believed to belong to the household field.

The two individuals belonging to the Bronze Age village of El Acequión were sampled from petrous bones (ElAcequion_Spain38_4055 and ElAcequion_Spain39_3993). El Acequión is located on the margins of an eponymous drained endorheic lake. Excavated features in the inner precinct include huts, pavements and dump yards, where most of the faunal remains derive. An extraordinary number of horse bones have been found on site, many of them exhibiting chop and hack marks, which gave rise to the hypothesis that, despite its chronology, some of these horses could represent domestic animals (Liesau, 2005). No human remains have been recorded.

The fortress of Els Vilars, located in the Segre valley, Catalonia, Spain, is a complex defensive site, occupied between 750 BCE and 325 BCE. The fortress has been built and developed throughout four distinct periods: Vilars 0 (775-700/675 BCE), Vilars I (700-675-550 BCE), Vilars II (550-425 BCE), Vilars III (425-375/350 BCE) and Vilars IV (375/350-325 BCE). The pre-Iberian phases (Vilars 0 and I) yielded little faunal archaeological record and very few equids (representing only ~0.8% of animal remains) as they were most likely associated with a pastoral economy based on subsistence activities. Both the quantity of domestic animal and the proportion of horse remains excavated increased throughout the following periods, with horses representing more than 10% of excavated animal bones of Vilars III (Nieto Espinet, 2016). Intriguingly, people inhabiting Els Vilars started burying horse foetuses by Vilars I, an unprecedented ritual practice that seemed to have become more common throughout Vilars II. Individual ElsVilars_UE4618_2672 was sampled from an adult bone in a domestic unit associated with Vilars I.

The sample Vicerrectorado_VIR175_1717, a superior premolar, comes from the excavation developed in a plot of the Roman city of *Lucus Augusti*, capital of the *Conventus Lucensis* and administrative center of northern *Gallaecia* in the northwest of the Iberian Peninsula. In this place, a long occupational sequence has been recognized, but the stratigraphic unit (UE-2101) in which this sample was recovered is assigned to low-imperial moments (C4th) considering the archaeological material present. In addition to equine remains, the fauna corresponds mainly to domestic species (cattle, pigs, sheep, goats, dogs) as well as some wild animals, mainly deer.

Sweden

(Uppsala)

Individual Uppsala_Upps02_1317 was sampled from a tooth excavated from a medieval site dating back to the C7th-C9th, and situated in Uppsala, Sweden.

Switzerland (Augusta Raurica, Stein am Charregass, Solothurn Vigier)

The Roman archaeological site of Augusta Raurica is located on the Rhine bank 20 kilometers east of Basel, Switzerland, and was an important Roman trade centre throughout the C1st, C2nd and C3rd. Individual AugustaRaurica_JG160_1817 was recovered from an underground fountain in Insula 8, an ancient residential and artisanal area of the site (Schmid et al., 2011). The fountain was established around 80 CE together with an 11 meters-deep well and an access tunnel typical of a Roman province architecture. The well was then likely used as a disposal structure and filled on several times with artifacts, tools and animal remains throughout the C2nd and C3rd. Individuals AugustaRauricaSchmidmatt_NBxK9279_1717 and

AugustaRauricaSchmidmatt_NBxP9261_1782 were unearthed from building Schmidmatt in Augusta Raurica, which represents one of the best preserved Roman buildings of Switzerland.

The site of Stein am Rhein Charregass was a Roman military camp on a hill close to the river Rhine. It was built as defence at the border of the Roman Empire at the end of the C3rd to the end of C4th. Large amounts of animal bones were retrieved from a ditch around the camp, including sample Charregass_NBxRa849_1667.

The site Solothurn Vigier was a Roman vicus and trading post during the C1st–C4th by the river Aare. It is located along the axis of the Roman regional capital Aventicum and Augusta Raurica at the river Rhine. A bridge indicates a changing place for transport animals, which is supported by the presence of cattle and many equine remains, such as SolothurnVigier_NB63_1867 SolothurnVigier_NB175_1817 and SolothurnVigier_NB699_1867.

Turkey (Yenikapi)

Yenikapi is a site located in present-day Istanbul that used to be the major Byzantine harbour of Theodosius, founded by emperor Theodosius I during the 4th century BCE. Yenikapi thus represents the major harbour of the Byzantine period, and a prominent Late Antiquity trade hub in the Mediterranean basin. Excavations revealed 36 wrecked ships, many artifacts and tools, and over 20,000 animal bones, mostly horses, mules and donkeys, but also cattle, sheep and pigs. Horse skeletons, representing ~32.6% of all animal skeletal remains, mainly belong to young male individuals no older than ten years (Onar et al., 2013). These include 12 stallions sequenced in this study, all sampled from petrosal bones: Yenikapi_Tur140_1289, Yenikapi_Tur141_1430, Yenikapi_Tur142_1396, Yenikapi_Tur145_1156, Yenikapi_Tur146_1730, Yenikapi_Tur150_1443, Yenikapi_Tur167_1443, Yenikapi_Tur168_1443, Yenikapi_Tur169_1443, Yenikapi_Tur170_1443, Yenikapi_Tur171_1689, Yenikapi_Tur173_1443, Yenikapi_Tur175_1443, Yenikapi_Tur176_1443, Yenikapi_Tur181_1443, Yenikapi_Tur189_1443, Yenikapi_Tur191_1443, Yenikapi_Tur193_1443, Yenikapi_Tur194_1360, Yenikapi_Tur206_1443, Yenikapi_Tur229_1443, Yenikapi_Tur243_1443, Yenikapi_Tur244_1443, Yenikapi_Tur246_1443, Yenikapi_Tur271_1443, Yenikapi_Tur273_1443, Yenikapi_Tur276_1443 and Yenikapi_Tur277_1443.

United Kingdom (Brough of Deerness, Quoygrew, Whitehall Roman Villa and Witter Place)

Brough of Deerness is a Pictish (pre-Viking Age) and Viking Age settlement set atop a roughly 30 meters-high sea stack located in the East Mainland of Orkney. All horse samples analysed in this study (BroughOfDeerness_VHR010_1417, BroughOfDeerness_VHR011_1367, BroughOfDeerness_VHR037 and BroughOfDeerness_VHR062_1417) are of Pictish date (C6th–C7th) and originate from excavation area C (Barrett and Slater, 2009).

The sample Quoygrew_VHR017_1117 was excavated from the site of Quoygrew, a rural settlement situated on the island of Westray in Orkney, Scotland, that was occupied from the C10th until the 1930s. The sample was from Context C010 (of C11th–C12th date) within excavation area C of a coastal shell and fish-bone midden (Barrett, 2012).

Individual WhitehallRomanVilla_UK08_1667 was sampled from a petrous bone of a horse excavated from a Roman settlement near Nether Heyford, Northamptonshire, United Kingdom, dating back to the Roman times (C3rd and C4th).

Witter Place is a post-medieval/pre-modern site, dating back to the C17th to the C19th and located just outside the northern walls of the city of Chester, United Kingdom. Excavated in 2001, the site shows a very large number of animal remains, mainly of cattle and horse (representing 93% of identifiable remains) but also dog, and to a smaller extent goat, sheep, pig and other wild mammal species, such as rabbit or hare. Witter place most likely represents a butchery site involved in animal-based industry and a possible tanning complex as suggested by the discovery of tanning pits. Most horses were old individuals over 20 years, showing some hack and chop marks for ~3% of them. While all parts of the horse skeleton could be recovered, the distribution of bones was uneven, with a lot of long bones but very few phalanxes. Individuals WitterPlace_UK15_217, WitterPlace_UK16_217, WitterPlace_UK17_267, WitterPlace_UK18_267, WitterPlace_UK19_267 and WitterPlace_UK20_217 were recovered from petrosal bones found on site.

Uzbekistan (Yerqorqan/Erkurgan)

Sample Yerqorqan_YER28_2853 was recovered from Yerqorqan (also referred to as Erkurgan), an ancient city surrounded by walls, located in Southern Uzbekistan, North of Qarshi. It was first settled during the second half of the 2nd millennium BCE, then occupied until the Sassanian period and destroyed by the Turks during the C6th. Excavations revealed the presence of a large palace located within a citadel, as well as temples and a mausoleum, thus showing the political and religious influence of the ancient city. Moreover, excavated blacksmith tools, pottery and currencies tend to suggest that this site used to be a major trading hub.

Museum

Individual Museum_Earb6_89 is an English Thoroughbred racehorse stallion, known as Dark Ronald, preserved in a museum in Halle, Germany. He was born in 1905 and shot to death in 1928, when old and suffering from colic. Early in his life, he fathered many thoroughbred and sport horses, and hence has had a critical influence on the whole horse warmblood breeding industry until now. DNA was sequenced from a petrosal bone of Dark Ronald's skull. Individual Museum_Earb5_105 was sampled from a petrosal bone of an English Thoroughbred stallion, also from a museum in Halle, Germany, that was born in 1891 and died in 1912.

Comparative dataset

We assembled a comparative panel consisting of 42 ancient and 30 modern genomes published in the literature (Table S3). Additionally, the genomes of namely a Somali wild ass and/or a domestic donkey were included as outgroups (Jónsson et al., 2014; Orlando et al., 2013). The panel of ancient horse genomes consists of three wild extinct horses from a now-extinct lineage dating back to ~5000-42000 years ago (Librado et al., 2015; Schubert et al., 2014a), four horses from Botai and five from Borly4, dated to ~5,000-5,500 years ago, one mare associated with the Sintashta culture (~4,000 years ago), two stallions from Arzhan I dating back to ~2,700 years ago, 11 stallions associated with the Scythian site of Berel', ~2,300 years ago, one horse from the C19th Yakutia and one historical Przewalski's individual, also from the C19th. The panel of modern horse includes six genomes of Przewalski's horses and those of 24 domestic horses from 18 domestic breeds, namely one Arabian horse, one Connemara Pony, one Duellmener, one Franche-Montagnes, one Friesian, one Hanoverian, one Saxon-Thuringian Heavy Warmblood, one Jeju pony, one Marwari, one Morgan, one American Quarter, one Sorraia, one Standardbred breed, as well as two

Icelandic horses, two Shetland ponies, two Mongolian individuals, two Thoroughbreds and three Yakutian horses.

METHOD DETAILS

DNA extraction and genome sequencing

Drilling and DNA extractions of osseous material were carried out in the ancient DNA facilities of the Centre for GeoGenetics, University of Copenhagen (Denmark), and laboratoire AMIS CNRS UMR 5288, Université de Toulouse III Paul Sabatier (France). DNA was extracted from 115-730 mg of bone or tooth powder, following method Y from (Gamba et al., 2016), with slight modifications. The powder was first digested for 1h at 37°C, in 4 ml of lysis buffer, composed of EDTA 0.45M, N-lauryl Sarcosyl 0.5% and Proteinase K 0.25 mg/ml. Following this pre-digestion step, the recovered pellets underwent a second digestion overnight at 42°C in an identical fresh lysis buffer. DNA was then concentrated and purified from the supernatant fraction of this second digestion. The vast majority of DNA extracts were also incubated with USERTM enzyme mix (NEB[®], 0.235 units/ μ L) at 37°C for 3 hours in order to remove uracil residues and thus reduce the impact of nucleotide mis-incorporations due to post-mortem cytosine deamination, typical of ancient DNA (Briggs et al., 2007).

DNA extracts were subsequently constructed into Blunt-End DNA libraries, following (Meyer and Kircher, 2010) as modified in (Gamba et al., 2016) (method A) or for a limited number of samples (Cantorella_UE2275x2_4791, ElAcequion_Spain38_4058, ElAcequion_Spain39_3993, ElsVilars_UE4618_2672), a slightly different procedure (method B) that differed from method A in the addition of 7-nucleotides-long unique indices within each P5 and P7 adapters prior to ligation to DNA, the sequence of which was obtained from (Rohland et al., 2015). A quantitative real-time Polymerase Chain Reaction (qPCR) was then carried out on 20X dilution duplicates of each library to determine the required PCR cycle number for amplification. Subsequently, libraries were amplified for 4-16 cycles (Table S4), following (Gamba et al., 2016). Each PCR reaction included 1 unit of AccuPrimeTM Pfx DNA polymerase, 3 to 6 μ l of unpurified DNA library and 1-2 μ l of 5 μ M custom PCR primers, one of which containing a unique external 6-bp index used for post-sequencing sequence demultiplexing. PCR products were then purified using either Minelute columns (Qiagen(C)) or Agencourt AMPure beads, eluted in 25 μ l of EB (10 mM Tris-Cl, pH 8.5) supplemented with 0.05% Tween, and quantified on TapeStation 2100/4200 or Bioanalyzer instruments (Agilent technologies). Finally, sequencing was performed at the Danish National High-Throughput DNA Sequencing Centre, either on Illumina HiSeq2500 for the vast majority of samples (method A), or the Illumina HiSeq4000 for samples (Cantorella_UE2275x2_4791, ElAcequion_Spain38_4058, ElAcequion_Spain39_3993, ElsVilars_UE4618_2672) (method B). Sequence trimming, mapping, filtering and base calibration at damaged sites were carried out following the methodology from (Gaunitz et al., 2018).

Multiple independent amplifications carried out for most DNA libraries to limit PCR duplicates, aiming at limiting sequencing costs. In total, one to seven libraries were generated for each sample selected for whole-genome sequencing, among which the vast majority were USERTM-treated (322/326, representing ~98.8%). The list of these indexed libraries with corresponding sequencing effort and clonality can be found in Table S4.

Radiocarbon dating

Unless indicated otherwise, AMS-radiocarbon dating of the samples was performed at the Keck Carbon Cycle Accelerator Mass Spectrometry Laboratory, UC Irvine. Bone or tooth pieces between 1.01 g and 2.24 g were sampled in the bone laboratory in the ancient DNA facilities of the Centre for GeoGenetics and sent for subsequent dating of ultrafiltered collagen. Sample preparation backgrounds were estimated, and subtracted, based on measurements of ^{14}C -free mammoth and whale bone. Calibration was carried out using OxCalOnline (Ramsey, 2009) and the IntCal13 calibration curve. Calibrated dates are provided in Table S1. Collagen yields obtained for the samples Ridala_Rid2_2717, Marvele_01_1117, Saadjarve_Saa1_1117, Dariali_Georgia2_317, Halvai_KSH4_4017, Halvai_KSH5_2542, were insufficient for radiocarbon dating. Their respective age was, thus, determined based on their archaeological context.

QUANTIFICATION AND STATISTICAL ANALYSIS

Read alignment, rescaling and trimming

We generated one to seven DNA libraries per sample, the majority of which were built on DNA extracts treated with USERTM enzyme mix (322/326, ~98.8%). This enzyme mix was used to limit the impact of mis-incorporations at deaminated cytosines in downstream analyses (Briggs et al., 2010). For each library, sequencing reads were parsed through PALEOMIX version 1.1.1 (Schubert et al., 2014b) with default parameters, except seeding that was disabled. Through this pipeline, reads were trimmed for low quality termini and known adapter sequences, then filtered out if shorter than 25 nucleotides, and finally aligned against the horse mitochondrial (Genbank Accession number NC_001640) (Xu and Arnason, 1994), and the horse nuclear reference sequence (EquCab2) (Wade et al., 2009) appended for 2,797 Y-chromosome contigs (Wallner et al., 2017). AdapterRemoval2 (Schubert et al., 2016) was used to trim reads and BWA version 0.5.9-r26-dev (Li and Durbin, 2009) (Li and Durbin, 2009) to map reads, excluding alignments whose mapping qualities were inferior to 25. Finally, PCR duplicates were removed and reads were locally realigned around indels using the IndelRealigner procedure from GATK (McKenna et al., 2010).

The software mapDamage2 (Jónsson et al., 2013) was used to check for the presence of nucleotide mis-incorporation profiles characteristic of ancient DNA data at the library level, randomly selecting 100,000 reads. We observed the expected increase of C to T (G to A) mis-incorporation rates at read starts (read ends) for both USERTM-treated and non-USERTM-treated data. Furthermore, genomic positions preceding read starts were higher in purines in non-USERTM read alignments, consistently with post-mortem DNA fragmentation being depurination-driven. In USERTM-treated read alignments, these positions were enriched in cytosine residues, in line with the excision of deaminated cytosines by the sequential activities of Uracil DNA glycosylase and Endonuclease VIII enzymes present in the USERTM mix. In order to limit the impact of remnant mis-incorporations in downstream analyses, we applied the computational procedure combining end trimming and base quality rescaling based on post-mortem DNA damage profiles, as described in (Gaunitz et al., 2018).

Uniparental markers

Mitochondrial DNA

Mitochondrial haplotypes were called following the procedure from (Gaunitz et al., 2018), restricting to positions showing at least 3-fold coverage after mapDamage rescaling and trimming, and a minimal base quality score (post-rescaling) of 25. The sequence data were divided into six independent partitions, including first codon positions (partition 1; 3,802 sites), second codon positions (partition 2; 3,799 sites), third codon positions (partition 3; 3,799 sites), the control region (partition 4; 961 sites), ribosomal RNAs (partition 5; 2,555 sites) and transfer RNAs (partition 6; 1,518 sites). We then constructed a dataset including all mitochondrial sequences analyzed by Gaunitz and colleagues (2018) plus all novel sequences reported in this study (dataset 1, 393 sequences). This dataset was used for phylogenetic reconstruction in RAxML (Stamatakis, 2014), using, for each partition, the GTRGAMMAI substitution model. This model was determined following modelgenerator v0.85 under the Bayesian Information Criterion (BIC) (Keane et al., 2006) as the best implemented in the software for partitions 1, 4, 5, and 6 (the best models implemented in the software for partitions 2 and 3 were GTRINV and GTRGAMMA and were nested within the GTRGAMMAI model). Node support was estimated using 100 bootstrap pseudo-replicates. The best-ML phylogenetic tree reconstructed is provided in Figure 6A.

Y chromosome

Read alignment, rescaling and trimming were carried out following the same methodology as described by (Gaunitz et al., 2018). The Y-chromosome haplotype was reconstructed from the high-quality reads aligning against the concatenated Y-chromosome contigs described by (Wallner et al., 2017) and three non-repetitive parts of the Y-chromosome (Genbank Accession Nb. AC215855.2 and JX565703), following the filtering criteria presented by (Gaunitz et al., 2018), relaxed here to include samples showing at least 25% of the total number of sites covered in the most-covered individual (Icelandic_0144A_0). This resulted in the selection of a total number of 139 individuals. In order to ensure orthology, phylogenetic reconstructions on the Y-chromosome data were restricted to the set of contigs present as single copy (*i.e.* excluding collapsed Copy Number Variants), as identified by (Wallner et al., 2017). Additionally, singletons as well as sites not covered in at least half of the total number of males were disregarded in order to reduce the impact of sequencing errors and missing information on the reconstruction, respectively. The phylogenetic analyses were performed using a GTRGAMMA substitution model in RAxML (version 8.2.4) (Stamatakis, 2014). The GTRGAMMA model was the second best-supported model identified in ModelGenerator v0.85 (Keane et al., 2006) under both Akaike and Bayesian Information Criteria, and the best of those implemented in RAxML. Figure 6B shows an easy-to-read collapsed version of the tree rooted on the domestic donkey, which was based on 4,996 variable sites. The overall topology was in agreement with those reported by (Gaunitz et al., 2018) based on a more limited number of samples. Node support was estimated using 100 bootstrap pseudo-replicates.

Autosomal and sex chromosomes

For all datasets considered in subsequent analyses of autosomal and sex chromosomes, we applied the following filters using ANGSD (Korneliussen et al., 2014):

- Minimum base quality: 20 (-minQ 20)
- Minimum mapping quality: 25 (-minMapQ 25)
- Remove all no primary, duplicated, and unmated reads (-remove_bads 1)
- Remove all reads showing multiple hits (-uniqueOnly 1)
- Downgrade quality scores around indels (-baq 1)
- Adjust mapping qualities in regions with excessive mismatches (-C 50)

We also used ANGSD (Korneliussen et al., 2014) to generate two different datasets based on autosomal sequencing data. The first dataset (a) consists of pseudo-haploid genomes (hereafter referred to as the pseudo-haploid dataset), generated by random sampling of reads, following the methodology used in (Gaunitz et al., 2018). Briefly, from the pileup of reads passing all quality filters, we sampled a single read at every genomic site for every horse. This dataset contains 160 samples that, prior to read random sampling, reached a sequencing depth-of-coverage equal to or greater than 1-fold, including *E. africanus somaliensis* as an outgroup (Table S5).

The second dataset (b) was prepared following the filtering of nine samples (Berel_BER04_D_2300, Friesian_0296A_0, Garbovat_Gar3_3574, Berel_BER07_G_2300, Berel_BER12_M_2300, Oktyabrsky_Rus37_830, Saadjarve_Saa1_1117, Tachtiperda_TP4_3604, Yerqorqan_YER28_2853) showing high overall error rate estimates (noCpG error estimate; $\epsilon \geq 0.0005$ errors per site) (Figure S1). In contrast to dataset (a), which was based of pseudo haploid calls, dataset (b) is based on posterior probabilities of genotypes ('posterior probability genotype dataset') of 126 modern and ancient samples with a depth-of-coverage equal to or greater than 1-fold. The 126 samples fall within the second domestic clade (DOM2), as defined by both the bootstrapped neighbour-joining tree and TreeMix results. Allele frequencies of at least 60 out of the 126 individuals were used as prior (-doMajorMinor 1 -doMaf 1 -beagleProb 1 -doPost 1 -GL 2). This dataset was used to estimate individual heterozygosities, autosomal nucleotide diversity profiles across cultures and through time (Table S5).

Additionally, we used ANGSD (Korneliussen et al., 2014) to generate pseudo-haploid calls for the two sex chromosomes, conditioning on males with a depth-of-coverage of at least 1-fold. We used these datasets to calculate nucleotide diversity profiles across cultures and through time. We excluded the nine samples with high overall error rates (noCpG error estimate; $\epsilon \geq 0.0005$ per site), to ensure cross-comparison with the analyses based on dataset (b).

Type-specific and overall error rate estimates

We estimated type-specific and overall error rates following the procedure developed in (Orlando et al., 2013), which leverages a test genome, an outgroup genome (*E. africanus somaliensis*), and a so-called 'perfect' genome (Icelandic_0144A_0). For every genomic site covered by all three samples, a random allele was sampled from the pile of reads passing the quality scores per individual. This matrix of counts was then used to estimate the individual error rates. Briefly, the method uses a maximum likelihood approach to estimate the excess of derived mutations in the test genome, compared to the perfect genome. We estimated the

error rates for all samples included in dataset (a) showing a depth-of-coverage equal to or greater than 1-fold. Three different filtering approaches were used to estimate the error rates, (i) using all covered genomic sites, (ii) masking CpG dinucleotide sites in the reference genome (EquCab2.0) (Wade et al., 2009), (iii) forcing observed transitions from procedure (i) to zero prior to estimation of the overall error rate, in order to recover an error rate based on transversions only. The distribution of error rates in modern and ancient horses following each of the three filtering approaches is shown in Figure S1A. For every individual in dataset (b), we also estimated type-specific and overall error rates using the same high-quality genome and outgroup. For every site, we sampled an allele according to its corresponding genotype posterior probabilities as weights. As for dataset (a), we applied the three filters described above. Distributions of error rates in modern and ancient horses following these filters of dataset (b) are shown in Figure S1B.

Genetic Distance and f3-outgroup statistics

We followed the methodology in (Gaunitz et al., 2018) to assess the amount of shared genetic drift between a pair of individuals using the f3-outgroup statistics (Patterson et al., 2012) in the form $f3(X, Y; \text{outgroup})$, where X, Y represent all possible pairwise permutations of horses. We used dataset (d), which consists of 160 samples (including ass *E. africanus somaliensis* as outgroup) with a sequencing depth-of-coverage equal or higher than 1-fold, and an Upper Palaeolithic individual, Goyet_Vert311_35780 (Table S5). To limit the possible impact of post-mortem DNA damage, we computed pairwise genetic distances on the basis of nucleotide transversions only ($n=50,757,656$) using ngsDist (Vieira et al., 2016) and estimated node support from 100 pseudo-replicate bootstraps. In total, we computed 12,561 f3-outgroup statistics based on those 50,757,656 nucleotide transversion sites (Figure S7A).

Principal Component Analysis (PCA)

Using the same dataset as for the calculation of genetic distances (dataset (a)), we conducted a Principal Component Analysis (PCA). We excluded the outgroup sample (Somali_0226A_0) and positions that segregated as singletons, which represented a total of 45,944,755 nucleotide transversions. The remaining 4,812,870 were used to conduct a PCA with plink v1.90b4.9 (Figure 5A). The fraction of the variance explained by the first three components was 11.6%, 10.4%, and 8.2%, respectively.

Genomic consequences of breeding: an increased genetic load

Population genetics predicts that the efficacy of negative selection depends on $s \times Ne$ product, where s is the selection coefficient and Ne the effective population size. This product is often binned into four categories, representing nearly neutral mutations ($s \times Ne \leq 1$), slightly deleterious ($1 < s \times Ne \leq 10$), mildly deleterious ($10 < s \times Ne \leq 100$), and strongly deleterious ($s \times Ne > 100$) (Kousathanas and Keightley, 2013; Williamson et al., 2014). With Ne recently dropping to *ca.* 100 reproductive horses (Corbin et al., 2010; Hall, 2016), some breeds likely carry a substantial number of alleles with $s < 0.01$, which behave as nearly neutral owing to a limited efficacy of negative selection ($s \times Ne = 0.01 \times 100 = 1$).

Such reduced efficacy of natural selection filtering out deleterious alleles results in an accumulation of harmful mutations (Charlesworth, 2009). Increased genomic load was indeed

observed for domestic horses, relative to a limited ancient DNA dataset including 14 Scythian horse genomes dating back to ~2.3 kya (Gaunitz et al., 2018; Librado et al., 2017). Leveraging the extensive time-series generated in this study (dataset (b)), we investigated the precise historical context in which the strong demographic declines resulting from breeding started to compromise the efficacy of selection, as potentially reflected by a burst in genetic load observed in each individual genome.

Prior to empirical analyses, however, we first thoroughly evaluated available genomic load estimators. In our previous work, the preferred estimator was calculated from protein-coding sites, following the procedure detailed in Librado *et al.* (2017) and Gaunitz *et al.* (2018) (Gaunitz et al., 2018; Librado et al., 2017)

$$LOAD(HOM) = \frac{\sum_{i=1} P(\text{homozygous_alternative})_i \times PhyloP_i}{\sum_{i=1} P(\text{homozygous})_i}$$

where $P(\text{homozygous_alternative})_i$ is the probability that site i is homozygous for the non-constrained nucleotide variant, as estimated by ANGSD (Korneliussen et al., 2014). $PhyloP_i$ is the evolutionary score for position i , which was used as proxy for the phenotypic impact of mutations at this position. PhyloP scores are available at <http://hgdownload.cse.ucsc.edu/goldenPath/hg19/phyloP46way/placentalMammals/> (Pollard et al., 2010). Since the denominator normalizes by the total amount of homozygous positions, calculations account for individual differences in inbreeding. This estimator should thus be interpreted as deleterious load *per* homozygous site. We here similarly define the genetic load in heterozygous positions

$$Load(HET) = \frac{\sum_{i=1} 0.5 \times P(\text{het})_i \times PhyloP_i}{\sum_{i=1} P(\text{het})_i}$$

We evaluated both statistics conducting forward simulations with SLiM 3.0 (Haller and Messer, 2017). In particular, we let an ancient population of 50,000 horses to evolve during 100,000 generations, until reaching mutation-selection-drift equilibrium. A hundred generations ago, the ancestral horse population split into three additional subpopulations, with $N_e = 100, 200$ and 500 , respectively. This range was chosen to mirror recent population sizes estimated for modern horse breeds. Mutation and recombination rates were assumed to be 7.24×10^{-9} mutations per generation and site, and 1×10^{-8} crossovers per generation and site, respectively. Deleterious mutations were supposed to represent $\frac{2}{3}$ of the protein coding sites, and were drawn from an exponential distribution with parameter -0.0001 (average selection coefficient against mutations = 0.0001). Five thousand 5 kb-long fragments were simulated (25 Mb in total), representing 5,000 protein-coding genes. Five individuals were sampled from each population.

We found that the estimator based on homozygous positions scales well with N_e reductions, for all mutation categories, as expected (Figure S3AB). Although the heterozygosity dropped by *ca.* 12% (for $N_e = 500$), 27% ($N_e = 200$) and 44% ($N_e = 100$), the load *per* heterozygous position remained steady. Since simulations provided certain genotype calls with null error rates, a steady heterozygous load cannot reflect methodological challenges associated with calling heterozygous genotypes. Instead, it suggests that the more

random loss or fixation of alleles owing to elevated drift, especially for variants that segregated as slightly deleterious in the ancient horse population, ultimately increased the levels of homozygous load. Therefore, given the limited power provided by heterozygous sites, we estimated genetic load from homozygous sites.

We further validated the homozygous load estimator by investigating its relation to the composite $s \times t$ parameter, where s is the selection coefficient against deleterious alleles and t is the number of generations of selection. Inspired by the McDonald-Kreitman test (McDonald and Kreitman, 1991), we contrasted two classes of sites. The first class was assumed to be neutral, and was comprised of protein-coding sites with a phyloP score lower than 1.5. This threshold was shown to best discriminate zero-fold and four-fold codon sites, the latter of which induces synonymous changes in protein-coding regions, and are often considered to evolve under almost neutrality (Librado et al., 2017). Besides drift and inbreeding, the second class of sites was supposed to be also affected by negative selection, and pertained to protein-coding sites with a PhyloP score greater or equal to than 1.5.

The rationale is that elevated drift, reflected as the loss of heterozygous sites, should result in an equally-balanced increase of both homozygous genotypes. With purifying selection, however, the deleterious homozygous genotype is often filtered out, resulting in the preferential inheritance of non-deleterious homozygous genotypes. More specifically, the genotype frequencies were modelled as:

$$\begin{aligned}
 AA_s &= \frac{AA_n}{AA_n + Aa_n \times (1 - sh)^t + aa_n \times (1 - s)^t} \\
 Aa_s &= \frac{Aa_n \times (1 - sh)^t}{AA_n + Aa_n \times (1 - sh)^t + aa_n \times (1 - s)^t} \\
 aa_s &= \frac{aa_n \times (1 - s)^t}{AA_n + Aa_n \times (1 - sh)^t + aa_n \times (1 - s)^t}
 \end{aligned}$$

where the s or n subindices denote selected and neutral sites respectively. The dominance coefficient is represented by h . This implies that the ratio of homozygous genotypes is:

$$\frac{AA_s}{aa_s} = \frac{AA_n}{aa_n(1 - s)^t}$$

Rearranging

$$st \approx -\ln\left(\frac{AA_n \times aa_s}{AA_s \times aa_n}\right)$$

The $s \times t$ parameter was estimated from high-quality modern horse genomes. We also conditioned on sites that segregated as nucleotide transversions in our extensive panel of ancient horses; *ie.* at the time-scale studied encompassing a few hundred generations, the accumulation of deleterious alleles mainly reflects an increased impact of drift (measured as a loss of heterozygous sites), relative to a presumably-recent relaxation of purifying selection.

Applying this validation method to the selected positions, we found a negative correlation between $s x t$ and genetic load at homozygous sites (Spearman correlation; $r = -0.8736$; p -value = 5.2834×10^{-8} ; Figure S3C). This confirms that a relaxation of negative selection (*ie.* reduction in $s x t$) drove the load increment observed in modern horse genomes. No relation to the average depth-of-coverage was found (p -value = 0.7796).

Similarly, we calculated the substitution rate at functional and nearly-neutrally evolving sites. To simplify calculations, we relied on zero-fold (hereafter, 0d) and four-fold degenerate (4d) codon sites, where nucleotide changes always imply non-synonymous and synonymous changes, respectively. We conditioned on nucleotide transversions, to mitigate the impact of post-mortem damage in ancient DNA samples. For each horse in the DOM2, we calculated

$$dN - dS$$

where dN is the genetic distance, at 0d sites, between pseudo-haploidized version of a DOM2 and the donkey genome. The principle of this subtraction is analogous to the well-known dN/dS statistics. While $dN - dS$ scores greater than zero reflect the action of positive selection, values lower than zero reveal evolution under negative selection. The $dN - dS$ values are more robust than the dN/dS statistics to the presence of undetected sequencing errors, which is typically the case for genomes sequenced at low coverage. Assuming the error rate is approximately constant in 0d and 4d sites, modeling selective pressure as a subtraction allows canceling out the impact of errors on dN and dS estimates, in contrast to modeling through dN/dS statistics, where a constant error rate added to the numerator and denominator could considerably distort the ratio.

Less efficient purifying selection is expected to elevate this $dN - dS$, owing to non-synonymous deleterious mutations being more often fixed. Therefore, if individual deleterious loads increased, we should find such ratios to be larger in modern DOM2 horses than in horses that lived prior to the C19th. This was indeed observed (Figure S3D). The correlation between the $dN - dS$ statistics and the load estimator was actually significant (Pearson correlation; $r = 0.67$; $p = 0$; Figure S3E), despite exploiting considerably different information to estimate either the selective pressure or the accumulation of deleterious mutations, respectively. All together, these findings support that the genomic load only increased in the last centuries, following the horse population decline and population structuration implemented by modern breeding practices.

Individual heterozygosity

We calculated the individual heterozygosity level for all horses within the second domestic clade (DOM2) using dataset (b). DOM2 members were defined based on the TreeMix (Figure 3; Figure S7B). We summed up the posterior probability of being heterozygous per site per individual, excluding sites with missing data for each corresponding individual. To reduce the possible bias driven by misincorporations as a result of deamination in the historic and ancient samples, we conducted a series of analyses, (i) including all covered positions, (ii) masking all data observations overlapping a CpG in the reference genome, (iii) excluding all transitions (Figures S2A-S2B). We then performed the exact same analyses after subtracting mean error rates per site in order to mitigate the total contribution of sequencing errors (Figures S2C-S2D).

We next regenerated the dataset of posterior probabilities of genotypes restricting the analysis to individuals with a depth-of-coverage greater than or equal to 4-fold downsampled to 4-fold to control for the possible effect of depth-of-coverage variation across samples. Individual heterozygosity levels with and without subtracting mean error rates per site were consistent (drop of heterozygosity: 18.5% (p-value: 2.63×10^{-14}) prior to subtracting errors, and 13.2% (p-value: 8.365×10^{-15}) after subtracting errors).

Nucleotide diversity (π) profiles across cultures

We computed the nucleotide diversity (π) of a set of predefined cultures in dataset (b), provided the number of samples is equal to or greater than three (Table S5). We estimated π for the autosomes, the Y chromosome, and the mitochondrial DNA. We used three different approaches to compute π , depending on the ploidy and depth-of-coverage. For autosomes, we extracted the individuals related to each culture from the ‘posterior probabilities of genotypes’ dataset as:

$$\pi = \frac{2p(1-p)n}{n-1}$$

where p is the ancestral allele frequency segregating within the samples of this culture, and n the corresponding sample size. For the Y chromosome, we computed the Site Frequency Spectrum (SFS) from pseudo haploid genomes for all males per culture ($N_{\text{males}} \geq 3$) and computed π from the SFS using an in-house script, following (Librado et al., 2017). Finally, π per culture of the mitochondrial DNA was computed from multiple sequence alignment of haplotypes. To reduce DNA damage related biases for ancient samples, we excluded genomic sites in CpG dinucleotides context for both the autosomes and Y chromosome.

Nucleotide diversity (π) profiles through time

In addition to calculating π for a set of *a priori* defined cultures, we computed π through time, following the same approaches as described in section ‘*Nucleotide diversity (π) profiles across cultures*’, except that datasets were stratified according to the age of the individuals instead of culture. Samples younger than 400 years were represented within a single group, and from 400 years ago to the oldest sample in the second domestic clade, we grouped samples using a step size of 250 years. Groups with fewer than 3 stallions were excluded from the calculations. To reduce the possible bias introduced by different geographic substructure across the temporal windows, we estimated π in Asia and Europe separately (Table S5). As an additional caution, we computed π using the same set of stallions for the autosomes and Y chromosome in each time window. To minimize the impact of damage related biases, present in ancient samples, we excluded genomic sites in CpG dinucleotides for the autosomes and Y chromosome. Lastly, following (Wutke et al., 2018), we also stratified the samples in the Asian clade into four time windows (0-400, 400-900, 900-2200, 2200-5000 years ago) and computed π for the autosomes and Y-chromosome (Figure S2E). We applied the same filters as for the time slicing described above.

Selection targets

Populations Branch Statistic (PBS)

We applied the population branch statistics (PBS) (Yi et al., 2010) to a subset of cultures in the DOM2 domestic clade. PBS can identify genomic regions that underwent recent natural selection by comparing allele frequency changes between two sister populations and an outgroup population. We investigated selection targets in Byzantine horses, representative of Persian-related lines, in comparison to Gallo-Roman (representing European horses prior to pre-Islamic conquests) and Deer Stone (Asian horses prior to pre-Islamic conquests) horses.

First, using `mstatspop v.0.1beta (20180220)` (available at <https://github.com/CRAGENOMICA/mstatspop/archive/master.zip>), we calculated the fixation index (F_{ST}) between pairs of cultures in 50 kb genomic windows with a step size of 10kb, and transformed these into drift units $-\log(1-F_{ST})$. Using the computed drift units, we then calculated the population branch statistics following the methodology in Yi *et al.* (2010) (Yi et al., 2010). We excluded genomic windows with more than 25% missing sites as well as regions showing an excess of the transition/transversion (ti/tv) ratio. The upper boundary, delineating this ti/tv excess, was estimated from the distribution of ti/tv ratios across genomic windows. In particular, windows with a ti/tv greater than the mode of the distribution, plus the distance from the mode to the lowest ratio, were excluded. This was done to minimize the number of false positives due to post-mortem deaminations, which are reflected as nucleotide transitions. From the remaining windows, we selected the top-1,000 genomic windows according to their higher PBS score, as candidate regions for positive selection on each given branch/population. We confirmed that this arbitrarily-high threshold was conservative by estimating the distribution of (nearly-)neutral PBS scores from two categories of sites: (1) intergenic sites (defined as located at least 5kb from the closest gene annotation) and (2) fourfold-degenerate sites within annotated protein coding regions. Two null distributions of PBS scores under (nearly-)neutral evolution were generated by performing 5,000 bootstrap pseudo-replicates, which consisted of random sampling with replacement 50k sites pertaining to both categories (Figure S4). We found, for all three branches, that the highest PBS score for neutral evolving sites was smaller than the lowest identified PBS score among the top-1,000 genomic windows. This indicates that the significance threshold that we selected is conservative, and that the top-1,000 genomic windows provide genuine candidate regions for positive selection. The same was true when considering the top-2,000 genomic windows.

Genes whose protein-coding exons overlapped the top-1,000 or top-2000 PBS windows (Table S6) were submitted to functional enrichment analyses (Table S7), provided that they have a 1:1 ortholog relationship (single-copy genes) to human or mouse genes. Orthology was defined according to annotations in Ensembl Genes 92. Human and mouse orthologs were analyzed for functional enrichment, using the WebGestaltR script, on the following functional databases hosted by WebGestalt 2017 (Wang et al., 2017): *geneontology_Biological_Process_noRedundant*, *geneontology_Cellular_Component_noRedundant*, *geneontology_Molecular_Function_noRedundant*, *pathway_KEGG*, *pathway_Panther*, *pathway_Reactome*, *pathway_Wikipathway*, *disease_Disgenet*, *disease_GLAD4U*, *phenotype_Human_Phenotype_Ontology*. Only functional categories significantly enriched, according to their adjusted p -value < 0.05 (Benjamini and Hochberg, 1995), are reported in Table S7.

Mendelian traits

We investigated 57 SNPs associated with key phenotypic traits, including coat-color variation, genetic disorders, body size configuration, as well as racing and locomotion skills. We explored their origin and evolution across past equestrian civilizations.

We followed the same approach as in (Librado et al., 2017) to chart the frequency of the causative allele in all 159 horses comprising dataset (b). Causative alleles supported by a single read were conservatively considered as absent (Figure S5). A few alleles were not detected in ancient horses, but only identified in modern domesticates, including in *PDK4* and *PONI*, improving racing capabilities. The *DMRT3* allele, causing ambling gaits, was first identified in a heterozygous horse dating back 730 years ago (TavanTolgoi_GEP13_73), from the Great Mongolian Empire.

We next estimated allele trajectories over time, for the 57 SNPs, using a sliding window approach (step = 250 years, span = 1,000 years). For each horse within each time bin, we randomly sampled a single read, provided it passes all quality filters. This enabled us to calculate allele frequencies without genotype uncertainties. This process was repeated 100 times to approximate the sampling variance. The mean over the 100 replicates was plotted, with the shaded area representing twice the standard deviation, which approximately delimits the 95% confidence interval around the mean (Figure 4B; Figure S6).

TreeMix population tree

We created an extended dataset (c) where specimens were grouped mainly by culture, as shown in Table S5. This dataset consisted of some individuals present in dataset (a) plus a number of other specimens for which genome-scale data could be generated, providing a minimum of 2-fold genome coverage per group considered. In total, dataset (c) included 186 horses and an outgroup. Intra-group allele counts (*eg.* within cultures) were calculated supplying the `--within` command to `plink v1.90b4.9` (Purcell et al., 2007). The resulting output was subsequently transformed into TreeMix input format using the `plink2treemix.py` script delivered within the TreeMix package. Sites not covered in all groups were disregarded, yielding a total of 16,829,417 nucleotide transversions.

TreeMix v1.13 (Pickrell and Pritchard, 2012) was run applying the same parameters as in (Gaunitz et al., 2018), with a number of migration edges ranging from 0 to 10. The outgroup was forced to be placed in-between the wild ass and horses. The variance explained by each migration model ranges from 0.9989676 (0 migration edge) to 0.9989983 (10 migration edges), thus indicating a limited improvement for an increasing number of migration edges. For illustrative purposes, the tree with one migration edge was visualized using the R APE package (Paradis et al., 2004).

Struct-f4

The D (Green et al., 2010) and related f (Patterson et al., 2012) statistics have been proven as useful tools to unravel dynamic population structures in the last millennia, including to track introgression from extinct lineages. Methods exploiting f_4 statistics have been already implemented, in ADMIXTOOLS and the more recent admixturegraph packages (Patterson et al., 2012). These two package implementations often require proposing a model to be fit to all f_4 permutations. Proposed models should be based on prior knowledge, in the

form of potential hypotheses that can be subsequently contrasted. Owing to reduced sample sizes, prior knowledge can be however limited or even biased.

Leveraging permutations of the f_4 statistics (Patterson et al., 2012), we here present an additional approach to infer fine-scale population affinities, in the form of 3D visual embedding, which is complementary to existing methods. These will be included in the Struct-f4 package, available upon request.

The first step in the Struct-f4 pipeline is an efficient C implementation to rapidly calculate millions of f_4 permutations from a TreeMix or PLINK file in tped format. The output provides BABA and ABBA counts, alongside with f_4 values, standard errors and Z-scores.

The second program performs 3D visual embedding. More specifically, the f_4 statistics is formally defined as:

$$f_4(A, B; C, O) = (p_A - p_B)(p_C - p_D)$$

where p_A , p_B , p_C and p_O represent the allele frequencies in individuals/populations A, B, C and the outgroup, respectively. The term $(p_A - p_B)$ is then the change in allele frequency between individuals A and B. As drift path between two populations, it can be visualized as a Euclidean distance:

$$d(p_A, p_B) = \sqrt{(x_A - x_B)^2 + (y_A - y_B)^2 + (z_A - z_B)^2}$$

where x_A , y_A and z_A are the x, y and z geometrical coordinates for individual A. Struct-f4 searches for the x, y and z coordinates that minimize the difference between the observed f_4 values and those predicted from the 3D geometrical embedding. Higher dimensionality can be also considered.

Struct-f4 currently implements a cost function based on weighted least squares:

$$\operatorname{argmin} \frac{[d(p_A, p_B)d(p_C, p_O) - f_4(A, B; C, O)]^2}{s^2}$$

The s in denominator stands for the standard deviation of the corresponding f_4 estimate, so that f_4 values estimated with large variances contribute less to the overall cost function. This correction enables to include samples sequenced at very depth-of-coverage. They are implemented in the R programming language, and C++ for intensive calculations.

Struct-f4 minimizes the cost function in two steps. First, `f4_2_3D` runs the SANN algorithm, as implemented in the `optim` R function, for 100,000 MCMC iterations. This provides a sub-optimal approximation to the global maximum. This is followed by ML refinement through the L-BGFS-B algorithm, with stringent convergence criteria (`maxit=50000`, `factr=1e1`). Figure 5B provides a visualization of Struct-f4.

Modelling IBE contribution to DOM2

We used `momi2` (Kamm et al., 2018) to model the population history underlying the main horse lineages characterized, as well as to further assess the possible contribution of IBE to

DOM2. Momi2 leverages the multidimensional Site Frequency Spectrum (SFS) to identify the best population model among a pre-defined set, characterized by a series of demographic changes, split and admixture events. We first sub-selected modern horses to represent the DOM2 and Przewalski's horse lineages

(DOM2: FMontagnes_0065A, Hanoverian_0235A, Marwari_0239A, Shet_0249A_SPH020, Shetland_0250A, Icelandic_0144A, Mongolian_0153A and Mongolian_0215A and Przewalski's horses: Przewalski_0150A, Przewalski_0160A, Przewalski_0159A, Przewalski_0158A, Przewalski_0157A and Przewalski_0151A).

The other lineages considered were represented by one high-coverage sample for *E. lensensis* (Taymyr_CGG10022_42758) and two ancient specimens sequenced to lower coverage for IBE (Cantorella_UE2275x2_4791 and CaminoDeLasYeseras_CdY2_4678). We prioritized high-coverage modern genomes whenever possible to perform strict genotype calling and avoid inflating split times and population size estimates, owing to post-mortem damage and sequencing errors.

The coverage and age for the sample considered are provided in Tables S2 and S3. The donkey genome was used as outgroup to polarize alleles as ancestral or derived. Sites uncovered or heterozygous in the donkey genome were thus skipped.

To estimate the 4D-SFS, we applied all filters described in section 'Autosomal and sex chromosomes' to calculate the genotype probabilities with ANGSD (Korneliussen et al., 2014). Based on these probabilities, we called the most likely genotype. We restricted to sites covered at least in all non-IBE samples, provided they did not exceed the depth-of-coverage threshold reported by the PALEOMIX depths file and consisting to the 99.5% quantile of the per site depth-of-coverage distribution. This intended to eliminate regions which may represent undetected CNVs and erroneously inflate local heterozygosity estimates. We also removed heterozygous sites where the alternative allele is supported by a reduced number of reads, such as sites with a heterozygous allele balance lower or equal than 15%. For instance: in cases where the sequencing depth was eight-fold and only one read supported the alternative allele, this site was skipped as potentially resulting from a sequencing error.

The sequencing depth of the two IBE samples was not sufficiently high to perform accurate genotype calling (<3.68x). Based on the new strategy proposed by Consensify (Barlow et al., 2018), we devised a random-sampling approach that limits the impact of sequencing errors in the calculation of 4D-SFS. Sequencing errors are indeed often reflected as singletons and introduce long terminal branches in the genealogy. This may lead to overestimating both population sizes and split times. Our strategy consisted of randomly sampling three, four and five reads per site and sample. If three out of these three randomly sampled sets agreed, the allele determined by these reads was called. Otherwise, the site was set as missing. This strategy seeks for removing errors solely represented by one or two reads, which are usually incorporated into the data frame when sampling a single read per site. This error-aware sampling strategy is feasible for these two samples because their average sequencing depth is not extremely low, but >2.45x, warranting a sufficient number of sites covered three, four or five times. A pseudo-diploid IBE individual was reconstructed merging the two IBE samples that were pseudo-haploidized.

After filtering, we retained a total of 1,228,613,653 sites. Of these, 2,183,704 correspond to nucleotide transversions and were used to build the 4D-SFS. For momi2, the transversion mutation rate was fixed to be 2.3728×10^{-9} transversions per site and generation, following

previous work(Orlando et al., 2013). We then used momi2 to contrast the fit of five evolutionary scenarios to the 4D-SFS. These models are presented below:

- **Model A** consisted of the topology in the form (IBE, (*E. lenensis*, (Przewalski, DOM2))), disregarding admixture. Each branch in the population tree was allowed to have a different population size, except for the DOM2 demographic trajectories prior to 33,930 years ago, which was fixed according to the PSMC profiles reconstructed by (Librado et al., 2015).
- **Model B1** was equivalent to **Model A**, except that an introgression event from DOM2 into IBE, constrained to occur after the split between the Przewalski's horse and DOM2 lineage.
- **Model B2** was equivalent to **Model B1**, except that the directionality of the introgression event was reversed, from IBE into DOM2.
- **Model C1** consisted of the topology in the form (Ghost, (*E. lenensis*, (Przewalski, (IBE, DOM2)))), with gene flow from the early diverging Ghost population into IBE, as suggested by TreeMix (Figure S7B).
- **Model C2** consisted of the topology in the form (IBE,(*E. lenensis*,(Przewalski, DOM2))), with three gene flow events: (i) from IBE to DOM2; (ii) from IBE to the branch ancestral to (*E. lenensis*, (Przewalski, DOM2)) and (iii) from DOM2 to Przewalski's horses.
- **Model C3** was the same as C2, except that an additional introgression event was introduced from the branch ancestral to DOM2-Przewalski split into the IBE lineage.

Model A consistently recovers population trajectories previously published

To avoid issues related to local minima during likelihood optimization, models were initialized with reasonable starting guesses (from previous work), as well as with three sets of random values. The most likely replicate was selected for each model, albeit only limited likelihood variation was observed across replicates. **Model A** reported a likelihood of -9345832.7844 and a Kullback-Leibler divergence (KLdivergence) between the predicted and observed 4D-SFS of 0.05214. Under this model, the split time between the DOM2 and Przewalski's horse lineages was estimated to have occurred 45,775 ya, whereas the divergence of both *E. lenensis* and IBE lineages were 181,248 and 222,152 ya respectively. The estimated population sizes were $N_e = 30,985$ (DOM2), 7,360 (Przewalski's horses), 64,709 (*E. lenensis*) and 14,899 (IBE). These estimates are in line with previous work (Orlando et al., 2013) and mirror the well-documented demographic collapse in the Przewalski's horse population (Der Sarkissian et al., 2015). These N_e also support population decay in IBE horses, possibly owing to partial isolation within Iberia, an area that could have served as horse refugium during the last glacial maximum (Warmuth et al., 2011). That the split times of *E. lenensis* and IBE are separated by only 40 kya with high N_e implies reduced drift at the time of divergence and is consistent with their relative positioning on the PCA vs. struct-f4 plots (*i.e.* IBE is closer to *E. lenensis* in the latter as this method is less sensitive to drift post-divergence; the PCA places IBE farther away from all other lineages due to recent population collapses).

Limited but significant introgression from IBE into DOM2

Model B1 marginally improved the likelihood to -9345720.5512, as well as the KLdivergence (0.05209) over model A. Yet, the improvement detected was lower than that observed for **Model B2** (likelihood: -9345653.5967; KLdivergence: 0.05206), suggesting that introgression preferentially occurred from IBE to DOM2, rather than in the reverse direction. **Model B2** estimated gene flow from IBE to DOM2 to be ~0.4%, and to have occurred 8,483 ya. We caution that the estimated date for introgression is likely imprecise, as the amount of the underlying gene-flow is extremely limited and thus barely impacts likelihood values.

To further validate the directionality of gene flow, we polarized shared variants as ancestral (A) and derived (D), in order to efficiently exploit diagnostic allele configurations. We assumed the topology (((((DOM2, Dunaujvaros_Duk2_4077), Botai), *E. lenensis*), IBE), DONKEY). Note Dunaujvaros_Duk2_4077 was not included in the momi2 analyses due to its low sequencing depth, but incorporated due to its greater affinities to IBE (Figure 7). Assuming an infinite-sites model (one mutation per site, at most), there is only one scenario whereby IBE and DOM2 (or Dunaujvaros_Duk2_4077) could share ancestral alleles, yielding ADDDAA (DADDAA) patterns. This scenario involves derived mutations happening in the lineage immediately ancestral to *E. lenensis*, and the ancestral allele re-appearing into Duk2 through introgression from IBE.

Leveraging the panel of nucleotide transversions built for TreeMix and with Sintashta_NB46_4023 as representative of DOM2, we contrasted the amount of DADDAA vs. ADDDAA patterns. We indeed found an excess of DADDAA (2,722 vs. 1,817), revealing increased affinities between IBE and Dunaujvaros_Duk2_4077, relative to DOM2, consistent with Figure 7. The absolute difference is small, but revealed significant through 100 bootstrap pseudo-replicates (p -value < 0.01). This supports gene flow directionality from IBE to Dunaujvaros_Duk2_4077.

Detecting the presence of a ghost lineage

That IBE introgressed into Dunaujvaros_Duk2_4077 does not support TreeMix inference assuming one migration pulse. Under this TreeMix tree, IBE and Dunaujvaros_Duk2_4077 are placed as sister taxa, with input from a divergent unsampled (ghost) population into IBE (Figure S7B). Using momi2, we evaluated such topology (**Model C1**), but found no likelihood improvement (-9351206.7300). More importantly, under this model, all the IBE ancestry was inferred to derive from the ghost population, suggesting that IBE should actually occupy the position of the ghost population, and thus that IBE is basal to all caballine lineages investigated here.

Yet, IBE shows an extremely divergent Y chromosome (Figure 6B). Based on the observed vs. expected pairwise genetic distances as proxies for the residuals of the model fit, we propose an alternative explanation. This involves two additional admixture events (**Model C2**). The first consists of gene flow from DOM2 to Przewalski's horses, while the second from IBE to the ancestor of all remaining caballine horses. Implementing **Model C2** substantially improved KLdivergence and likelihood values, to 0.01899 and -9273440.1542, respectively. Optimizing **Model C2** parameters estimated that DOM2 branched off from Przewalski's 43.8 kya, and from *E. lenensis* 118.6 kya. Both such estimates are more in line with previous results based on F-statistics (Orlando et al., 2013). This model also indicates that DOM2 contributed 22% to the ancestor of Przewalski's horses ca. 9.47 kya, suggesting the Holocene optimum, rather than the Eneolithic Botai culture (~5.5 kya), as a period of

population contact. This pre-Botai introgression could explain the Y-chromosome topology, where Botai horses were reported to carry two different segregating haplogroups: one occupied a basal position in the phylogeny while the other was closely related to DOM2 (Figure 6B). Multiple admixture pulses, however, are known to have occurred along the divergence of DOM2 and the Botai-Borly4 lineage, including 2.3% post-Borly4 contribution to DOM2, and a more recent 6.8% DOM2 introgression into Przewalski's horses (Gaunitz et al., 2018). **Model C2** parameters accommodate all these as a single admixture pulse, likely averaging the contributions of all these multiple events.

More surprisingly, IBE was inferred to have massively contributed to the ancestor of all caballine horses, 285.3 kya with 98.8% of its genetic ancestry. The split time between IBE and the rest of caballine horses is then pushed back 1.25 mya, which could be compatible with the very divergent Y haplogroup carried by IBE stallions. This 98.8% should be interpreted as if most of the IBE-DOM2 genomic regions coalesced 285.3 kya, while a 1.2% originated from an even more divergent population. The massive improvement driven by this admixture event strongly supports the existence of a divergent ghost population, which could have participated to the genetic makeup of the caballine horses investigated here.

We caution however that the exact split time of 1.25 mya might vary depending on the existence of other unsampled populations. For example, the f_4 permutations involving Goyet_Vert_311 revealed that it likely represents a different lineage, with close affinities to both IBE and DOM2 at the same time. Although its low sequencing depth precludes its inclusion in momi2 analyses, its 35 kya radiocarbon date suggests it could descend from a population ancestral to the Przewalski-DOM2 split. More specifically, comparison of the observed vs. predicted genetic distances (residuals) of **Model C2** indicated that both Przewalski and DOM2 horses should be modeled as closer to IBE. We incorporated thus another admixture pulse in **Model C3**, from the branch ancestral DOM2 and Przewalski's horses to IBE. We found again significant improvement, with KLdivergence and likelihood values of 0.01738 and -9269933.1010, respectively. **Model C3** predicts a 36.7% contribution into IBE, 65.2 kya. The split between IBE and all remaining horse lineages is then reduced to 539.1 kya, more in line with the scale of the Y-chromosome tree. The N_e of Przewalski's and IBE horses decreased, to *ca.* 3,798 and 8,848 reproductive individuals, respectively. Importantly, **Model C3** is consistent with other models investigated, inferring an IBE to DOM2 introgression of only 1.4%.

Analytical predictions corroborate limited IBE contribution into DOM2

We next analytically estimated whether the admixture proportions estimated from the f_G parameter (Figure 7A) represent reasonable estimates of the possible genetic contribution of IBE into the DOM2 lineage. The following derivations assume three well-established constraints:

- According to the long internal branch leading to the IBE population in TreeMix inference (not shown for illustrative purposes), as well as momi2 parameters for **Model C3**, it is likely that IBE experienced a population decay. Its N_e population size was 8,848 horses, which is slightly larger than the value retrieved for Przewalski's horses, but considerably smaller than those leading to both the *E. lenensis* and DOM2 lineages.

- Under **Model C3**, the split time of *E. lenensis* was estimated to be ~112.7 kya, in line with previous work (Orlando et al., 2013). With a generation time of 8 years, this means that IBE diverged at approximately 15,000 generations ago (kga).
- As shown in Figure 7B, the introgression from IBE to DOM2 should have occurred after the split of Przewalski and DOM2, some 35.4 kya (~4.425 kga).

Given the distance between the early IBE split and its late contribution to DOM2, we here demonstrate that an admixture fraction greater than 15% is highly incompatible with the f_4 values reported in Figure 7B, ranging from -0.00018 for Dunaujvaros_Duk2_4077 to -2.68933⁻⁵ for Jeju_0275A. Durand and colleagues analytically showed that the expected number of ABBA-BABA events, in a (O,(P₃,P₂,P₁)) configuration, can be expressed in terms of explicit coalescent parameters (Durand et al., 2011). Respecting their notation (the sign of f_4 is reverted because we calculated BABA-ABBA):

$$-f_4 = \frac{ABBA - BABA}{\#SNPs} = \frac{3fd(N + tp3 - tGf - t)}{\#SNPs}$$

Isolating the admixture fraction (f)

$$f = \frac{\#SNPs * -f_4}{3d(N + tp3 - tGf - t)}$$

where d is the probability of pairwise coalescence within P₃. The denominator includes (i) N as the size of the population ancestral to P₁, P₂ and P₃; (ii) $tp3$ the split time of P₃, (iii) tGf the time of admixture between P₃ and P₂, (iv) and t is the expected time of coalescence within P₃, given that both lineages indeed coalesced within P₃. Assuming a constant N_3 population size within P₃:

$$d = 1 - \left(1 - \frac{1}{N_3}\right)^{(tp3 - tGf)}$$

and

$$t = \frac{\sum_{t=tGf}^{tp3} \frac{t}{N_3} \left(1 - \frac{1}{N_3}\right)^{(t - tGf)}}{d}$$

Our three constraints imply that $tGf < 5.6$ kga, $tp3 > 10$ k and $N_3 \ll N$. Indeed, while momi2 infers an N_3 of 8,848 IBE horses, the PSMC reconstructions 200-100 kya indicate that the horse population exceeded $N = 100,000$ horses (Librado et al., 2015). Based on this, we evaluated different realistic ranges for N (100-200k reproductive horses) and N_3 (5-50k), $tp3$ (10-62.5 kga) and tGf (0.5-6.2 kga). Within such parametric space, a $f_4 = -0.0001$ (Figure 7B) is compatible with very small admixture proportions. The maximum possible admixture fraction is ~12%, and is obtained under the less realistic values of $N = 100,000$, $N_3 = 45,000$, $tp3 = 10,000$ ga (80,000 ya) and $tGf = 6,000$ ga (48,000 ya). For other parameter combinations, the predicted admixture fraction is ~1-2%, in line with the limited fraction modelled in **Model C3**. Coupled with such constraints, therefore, the small f_4 values in Figure

7B corroborate that IBE contributed to DOM2, but cannot represent the major genetic pool leading to the makeup of modern domesticates.

Species and sex identification

Preliminary sequence data obtained while screening the DNA libraries, typically considering 1-5 million sequencing reads, were processed with the Zonkey package, as part of PALEOMIX (<https://github.com/MikkelSchubert/paleomix>; (Schubert et al., 2017)) in order to identify first-generation equine hybrids and determine the molecular sex of each individual. Among the 278 individuals included in this study, we detected three hemiones, three asses, 27 mules, 245 pure horses. We also identified 196 males (representing ~70.5% of all individuals) and 82 females (~29.5%). Additionally, among pure horses, we identified 175 stallions and 70 mares.

DATA AVAILABILITY

Sequencing data aligned against the horse reference genome EquCab2.0 (Wade et al., 2009) can be found on the European Nucleotide Archive, project XXXX (will be uploaded upon manuscript acceptance).

SUPPLEMENTAL TABLE TITLES AND LEGENDS

Table S1. Ancient equids sequenced in this study, related to Figure 1.

Archaeological, cultural and geographical information for the 278 ancient equids sequenced in this study, as well as their inferred species (including non-caballine and ass components), sex, average depth-of-coverages for their nuclear and mitochondrial genomes, calibrated radiocarbon measurement summary statistics and dating when available.

Table S2. Panel of ancient horses included in at least one nuclear DNA analysis, related to Figure 1.

Archaeological, cultural and geographical information, as well as their respective endogenous content, average sequencing depths for nuclear and mitochondrial genomes, and inferred sex.

Table S3. Comparative panel for nuclear DNA analyses, related to Figure 1.

ID, temporal and cultural background associated with each of the 42 previously published ancient and 31 modern equids included in this study, as well as their inferred sex, accession number and original publication. Nuclear and mitochondrial sequencing depths, prior and after trimming and rescaling, are also provided.

Table S4. Laboratory procedures carried out on each horse sequenced in this study, and that has been included in at least one nuclear DNA analysis, related to Figure 1.

Summary information on DNA extractions, library construction, amplification, and sequencing statistics per library. nuDNA and mtDNA refer to nuclear and mitochondrial DNA, respectively. All values have been calculated based on sequencing output prior to rescaling.

Table S5. Datasets used for autosomal and Y chromosome analyses, related to Figures 2, 3, 5 and 6. For each individual, its allocated culture is indicated, as well as which dataset(s) it is part of.

Table S6. Protein coding genes overlapping the top-1000 PBS windows and the top-2000 PBS windows for each branch in the Byzantine/Gallo-Roman/Deer Stone configuration, related to Figure 4.

Table S7. WebGestalt analysis for protein coding genes overlapping the top-1000 PBS regions and the top-2000 PBS regions, estimated for each branch in the Byzantine/Gallo-Roman/Deer Stone configuration, related to Figure 4.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological Samples		
osteological remain	this study	Actiparc_GVA124_2143
osteological remain	this study	Actiparc_GVA307_2127
osteological remain	this study	Actiparc_GVA308_2312
osteological remain	this study	Actiparc_GVA309_2302
osteological remain	this study	Actiparc_GVA311_2253
osteological remain	this study	ArzhanII_Rus9_2500
osteological remain	this study	AugustaRaurica_JG160_1817
osteological remain	this study	Bateni_Rus16_3350
osteological remain	this study	Beauvais_GVA122_417
osteological remain	this study	Beauvais_GVA375_467
osteological remain	this study	Belgheis_TrBWBX116_485
osteological remain	this study	Boves_GVA191_1717
osteological remain	this study	BozA dyr_KYRH10_1267
osteological remain	this study	BozA dyr_KYRH8_1267
osteological remain	this study	CaminoDeLasYeseras_CdY2_4678
osteological remain	this study	Cantorella_UE2275x2_4791
osteological remain	this study	Chartres_GVA26_1917
osteological remain	this study	Chartres_GVA36_1917
osteological remain	this study	Chartres_GVA4_1917
osteological remain	this study	Chartres_GVA43_1917
osteological remain	this study	Chartres_GVA56_1917
osteological remain	this study	Chartres_GVA81_1917
osteological remain	this study	EIAcequion_Spain39_3993
osteological remain	this study	ElsVilars_UE4618_2672
osteological remain	this study	Evreux_GVA133_1817
osteological remain	this study	Evreux_GVA140_1817

osteological remain	this study	FrankfurtHeddernheim_Fr1_1863
osteological remain	this study	GolModII_Mon23_2007
osteological remain	this study	GolModII_Mon24_1993
osteological remain	this study	GolModII_Mon25_2011
osteological remain	this study	GolModII_Mon26_1999
osteological remain	this study	GolModII_Mon27_2011
osteological remain	this study	Halvai_KSH4_4017
osteological remain	this study	Khatuu_Kha2_t1_2312
osteological remain	this study	Khotont_UCIE2012x85_1291
osteological remain	this study	Macon_GVA201_1767
osteological remain	this study	Mainz_Mzr1_1373
osteological remain	this study	Marvele_01_1138
osteological remain	this study	Marvele_18_1189
osteological remain	this study	Marvele_21_1087
osteological remain	this study	Marvele_32_1144
osteological remain	this study	MerzlyYar_Rus45_23789
osteological remain	this study	Museum_Earb6_89
osteological remain	this study	Noyon_GVA123_717
osteological remain	this study	Nustar_5_1187
osteological remain	this study	Oktyabrsky_Rus37_830
osteological remain	this study	OlonKurinGol_OKG2_2367
osteological remain	this study	Otepaas_Ote2_1184
osteological remain	this study	Ridala_Rid2_2717
osteological remain	this study	Saadjarve_Saa1_1117
osteological remain	this study	Sagzabad_SAGS27_3117
osteological remain	this study	SaintJust_GVA242_2250
osteological remain	this study	Sayangorsk_Rus41_2677
osteological remain	this study	SharlQumis_AM115_1557
osteological remain	this study	TavanTolgoi_GEP13_730

osteological remain	this study	TavanTolgoi_GEP14_730
osteological remain	this study	TavanTolgoi_GEP21_730
osteological remain	this study	UushgiinUvur_Mon37_3085
osteological remain	this study	UushgiinUvur_Mon39_3085
osteological remain	this study	UushgiinUvur_Mon41_3085
osteological remain	this study	UushgiinUvur_Mon42_3130
osteological remain	this study	UushgiinUvur_Mon43_3120
osteological remain	this study	UushgiinUvur_Mon44_3085
osteological remain	this study	UushgiinUvur_Mon45_3080
osteological remain	this study	UushgiinUvur_Mon79_3085
osteological remain	this study	UushgiinUvur_Mon87_3117
osteological remain	this study	WitterPlace_UK15_217
osteological remain	this study	WitterPlace_UK16_217
osteological remain	this study	WitterPlace_UK17_267
osteological remain	this study	WitterPlace_UK18_267
osteological remain	this study	Yenikapi_Tur140_1289
osteological remain	this study	Yenikapi_Tur141_1430
osteological remain	this study	Yenikapi_Tur142_1396
osteological remain	this study	Yenikapi_Tur145_1156
osteological remain	this study	Yenikapi_Tur146_1730
osteological remain	this study	Yenikapi_Tur150_1443
osteological remain	this study	Yenikapi_Tur170_1443
osteological remain	this study	Yenikapi_Tur171_1689
osteological remain	this study	Yenikapi_Tur173_1443
osteological remain	this study	Yenikapi_Tur175_1443
osteological remain	this study	Yenikapi_Tur176_1443
osteological remain	this study	Yenikapi_Tur181_1443
osteological remain	this study	Yenikapi_Tur193_1443
osteological remain	this study	Yenikapi_Tur194_1360

osteological remain	this study	Yenikapi_Tur229_1443
osteological remain	this study	Yenikapi_Tur243_1443
osteological remain	this study	Yerqorqan_YER28_2853
osteological remain	this study	ArzhanII_Arz15_2642
osteological remain	this study	ArzhanII_Arz17_2642
osteological remain	this study	ArzhanII_Rus11_2500
osteological remain	this study	Berufjordur_VHR102_1057
osteological remain	this study	BroughOfDeerness_VHR010_1417
osteological remain	this study	BroughOfDeerness_VHR011_1367
osteological remain	this study	BroughOfDeerness_VHR037_1417
osteological remain	this study	BroughOfDeerness_VHR062_1417
osteological remain	this study	Chartres_GVA111_1917
osteological remain	this study	Chartres_GVA112_1917
osteological remain	this study	Chartres_GVA115_1917
osteological remain	this study	Chartres_GVA28_1917
osteological remain	this study	Chartres_GVA47_1917
osteological remain	this study	Chartres_GVA48_1917
osteological remain	this study	Chartres_GVA53_1917
osteological remain	this study	Chartres_GVA60_1917
osteological remain	this study	Chartres_GVA75_1917
osteological remain	this study	Chartres_GVA9_1917
osteological remain	this study	EIAcequion_Spain38_4058
osteological remain	this study	Evreux_GVA135_1817
osteological remain	this study	Goyet_Vert311_35870
osteological remain	this study	Granastadir_VHR031_1067
osteological remain	this study	Marvele_02_1138
osteological remain	this study	Marvele_05_1138
osteological remain	this study	Marvele_16_1138
osteological remain	this study	Marvele_22_1138

osteological remain	this study	Marvele_27_1138
osteological remain	this study	Nustar_4_1187
osteological remain	this study	OlonKurinGol_OKG1_2367
osteological remain	this study	Quoygrew_VHR017_1117
osteological remain	this study	SaintQuentin_GVA237_1917
osteological remain	this study	Syrgal_Syr1t1c4_2317
osteological remain	this study	UushgiinUvur_Mon40_3085
osteological remain	this study	UushgiinUvur_Mon89_3085
osteological remain	this study	Vermand_GVA199_1742
osteological remain	this study	WitterPlace_UK19_267
Chemicals, Peptides, and Recombinant Proteins		
N-Lauroylsarcosine solution 30% 500ml	Dutscher	348533
Proteinase K 100MG	Thermo Fisher Scientific	10103533
H2O, Molecular Biology Grade, Fisher BioReagents	Thermo Fisher Scientific	10490025
Tween 20 100ML	Thermo Fisher Scientific	10113103
Ethanol, Absolute, Mol Biology Grade	Thermo Fisher Scientific	10644795
5M Sodium Chloride 100ML	Thermo Fisher Scientific	10609823
USER Enzyme	New England Biolabs	M5505L
NEBNext End Repair Module	New England Biolabs	E6050L
Bst DNA Polymerase	New England Biolabs	M0275L
NEBNext Quick Ligation Module	New England Biolabs	E6056L
BSA Molecular Biology Grade	New England Biolabs	B9000S
ROX REFERENCE DYE 500µL	Thermo Fisher Scientific	10635353
SYBR Green I Nucleic Acid 10000X concentrate	Thermo Fisher Scientific	10710004
ACCUPRIME PFX DNA POLYMERASE 100µL	Thermo Fisher Scientific	10472482

Agencourt AMPure XP - 60ml	Beckman Coulter	A63881
Buffer PE	Qiagen	19065
Buffer PB	Qiagen	19066
Buffer EB	Qiagen	19086
DMSO molecular biol 250ML	Thermo Fisher Scientific	10397841
dNTP Set 100mM 100µL	Thermo Fisher Scientific	10336653
EDTA 0.5M pH 8.0 Fisher Bioreagents 500ML	Thermo Fisher Scientific	10182903
Tris HCl, 1M, pH 8.0, 100ML	Thermo Fisher Scientific	10336763
Critical Commercial Assays		
MinElute PCR Purification kit	Qiagen	28006
Tapestation screenTape D1000 HS	Agilent	5067-5584
Bioanalyzer High Sensitivity DNA labchips kit	Agilent	5067-4626
Deposited Data		
Raw and analysed data	this study	will be uploaded upon manuscript acceptance
Software and Algorithms		
ANGSD	(Korneliussen et al. 2014)	https://github.com/ANGSD/angsd
TreeMix	(Pickrell and Pritchard 2012)	https://bitbucket.org/nygcresearch/tremix/wiki/Home
mapDamage	(Jónsson et al. 2013)	https://ginolhac.github.io/mapDamage
Paleomix	(Schubert et al. 2014)	https://github.com/MikkelSchubert/paleomix
mstatspop	unpublished	https://github.com/CRAGENOMICA/mstatspop
NGSdist	(Vieira et al. 2016)	https://github.com/fgvieira/ngsDist
fastme	(Lefort et al. 2015)	http://www.atgc-montpellier.fr/fastme/binaries.php
raxML	(Stamatakis 2014)	https://sco.hits.org/exelixis/web/software/raxml/index.html
PLINK	(Purcell et al. 2007)	https://www.cog-genomics.org/plink2
AdapterRemoval	(Schubert et al. 2016)	https://github.com/MikkelSchubert/adapterremoval
Samtools	(Li et al. 2009)	https://github.com/samtools/samtools
Momi2	(Kamm et al. 2018)	https://github.com/popgenmethods/momi2
Admixtools	(Patterson et al. 2012)	https://github.com/DReichLab/AdmixTools

Figure 1

[Click here to access/download;Figure;Fig1_01022019.pdf](#)

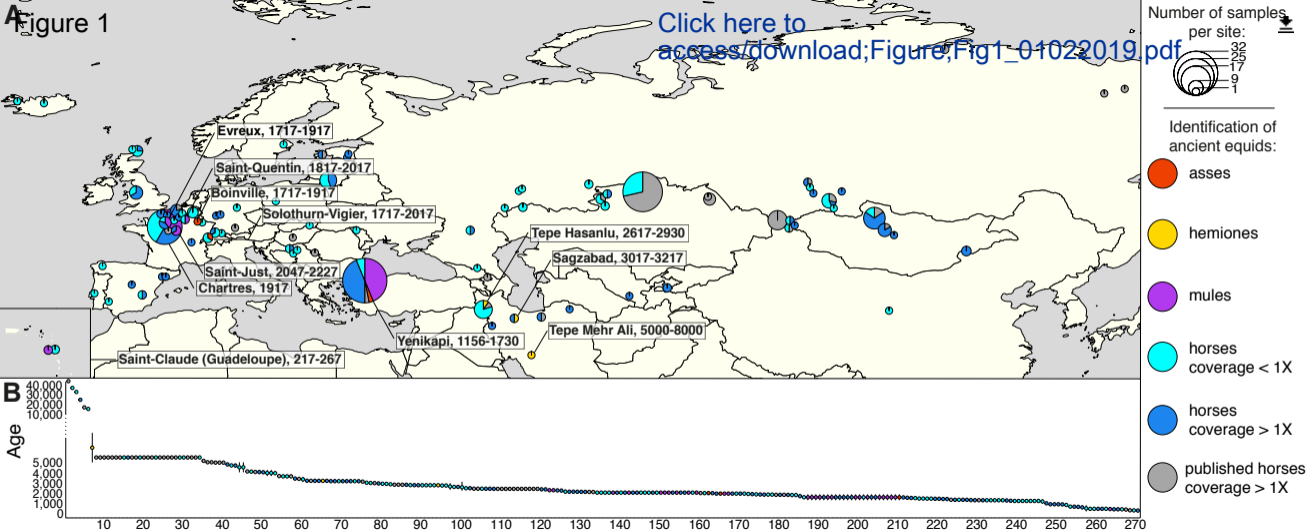
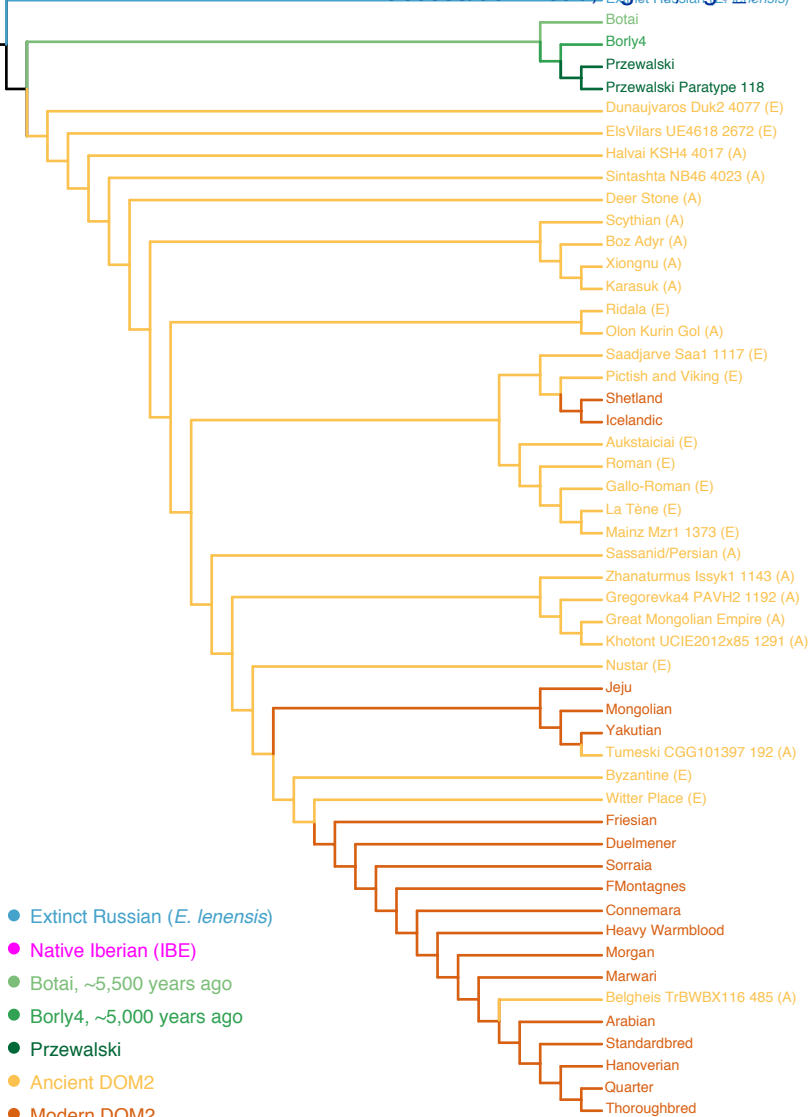




Figure 3

[Click here to](#)
[access/download/Figure3_0](https://www.ancientgenetics.com/access/download/Figure3_0)


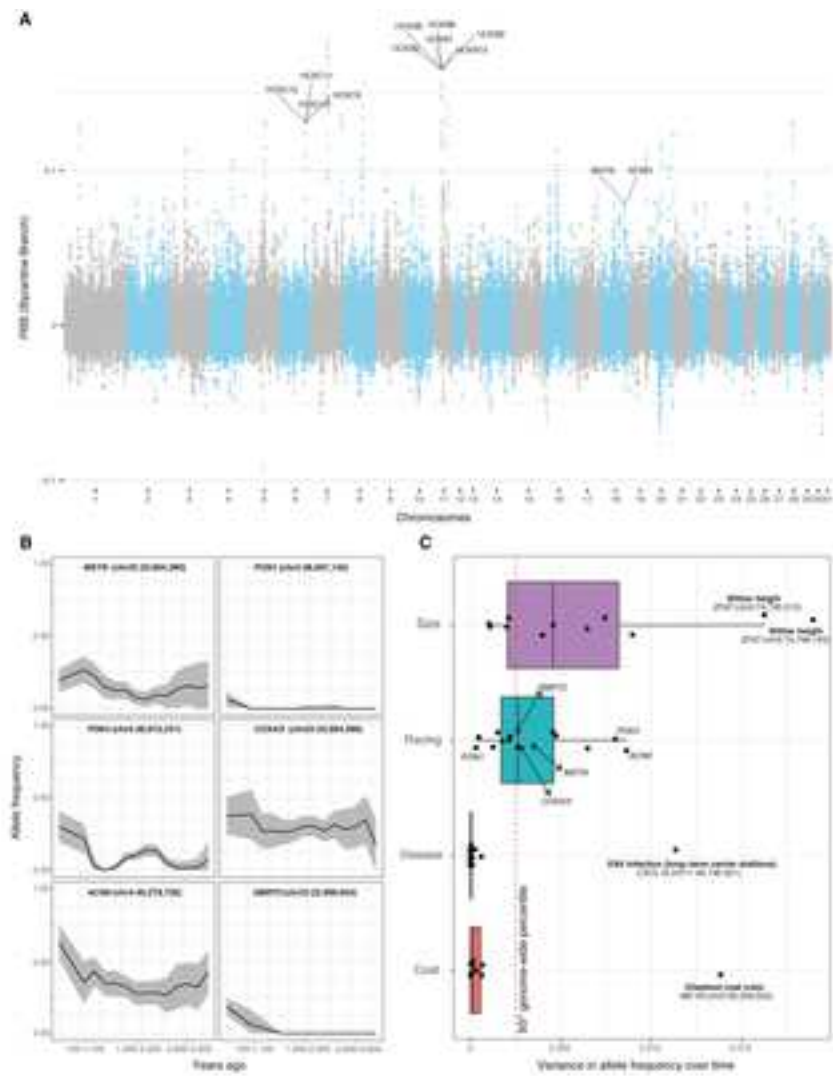
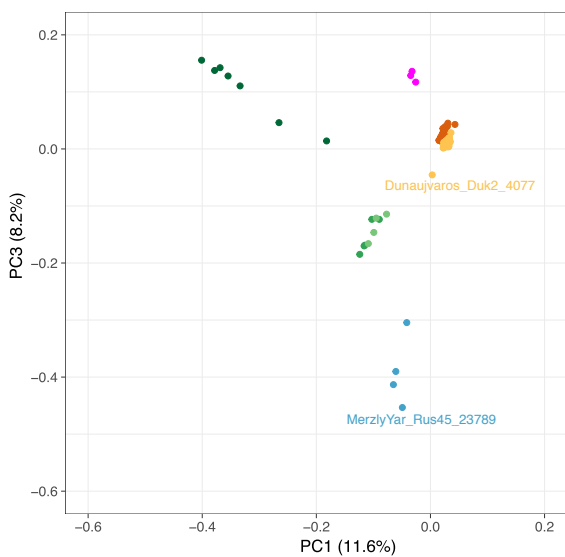
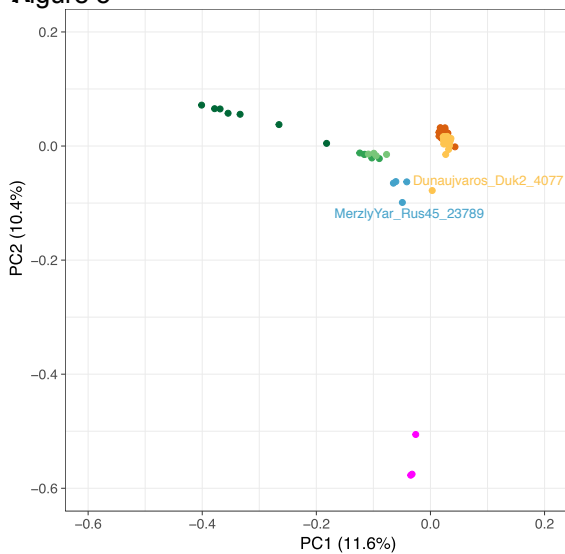
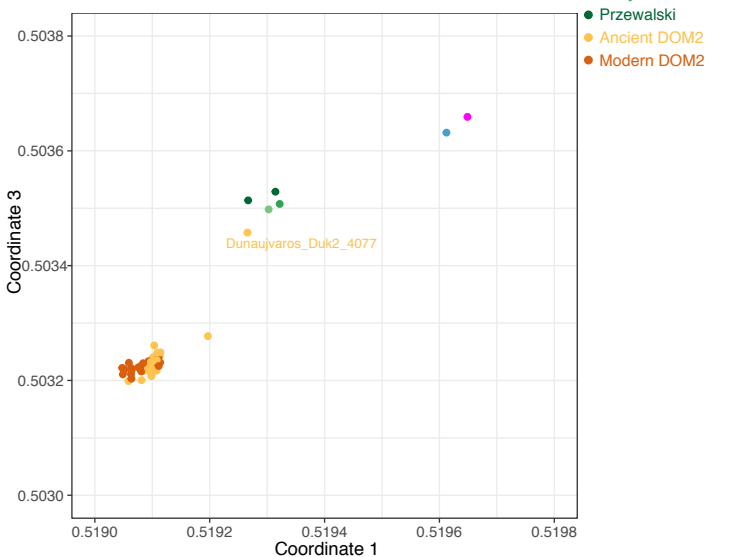
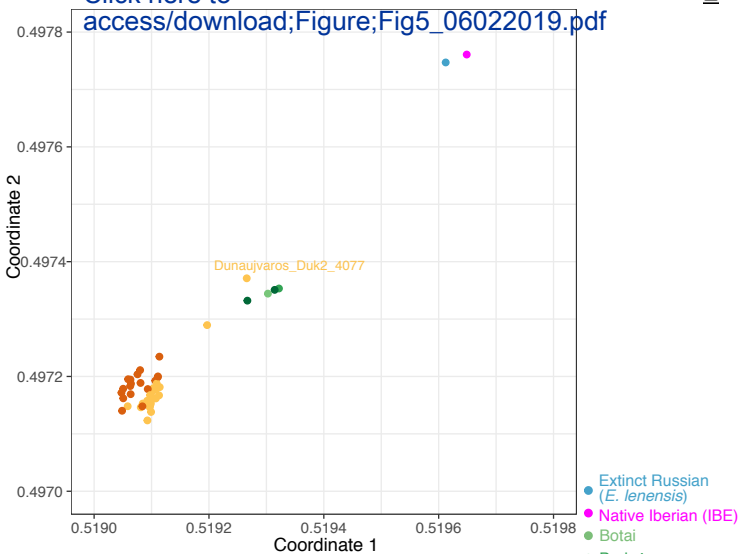


Figure 5



B

[Click here to access/download;Figure;Fig5_06022019.pdf](#)



C

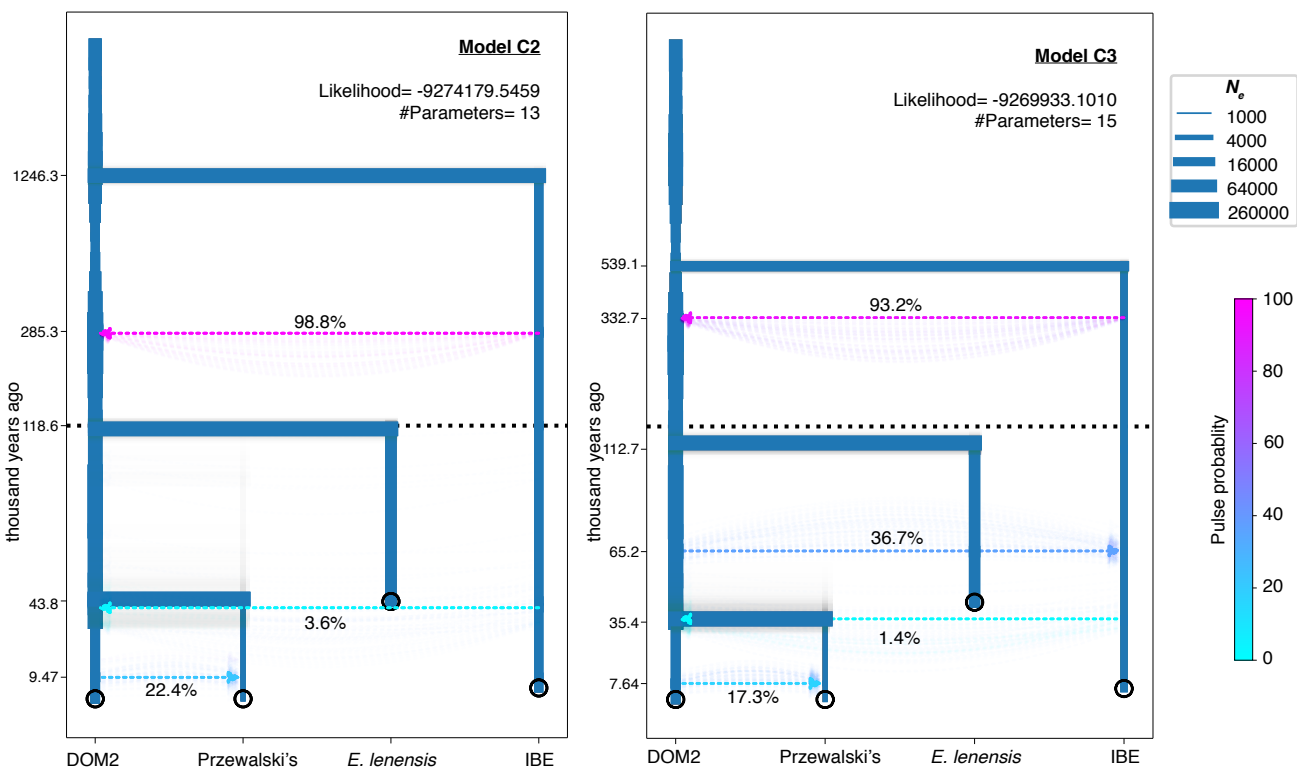
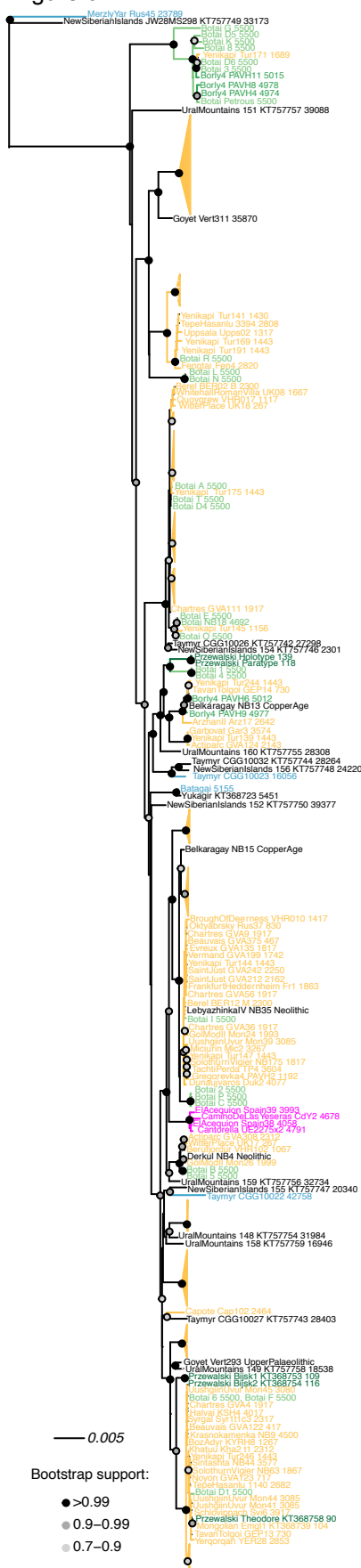
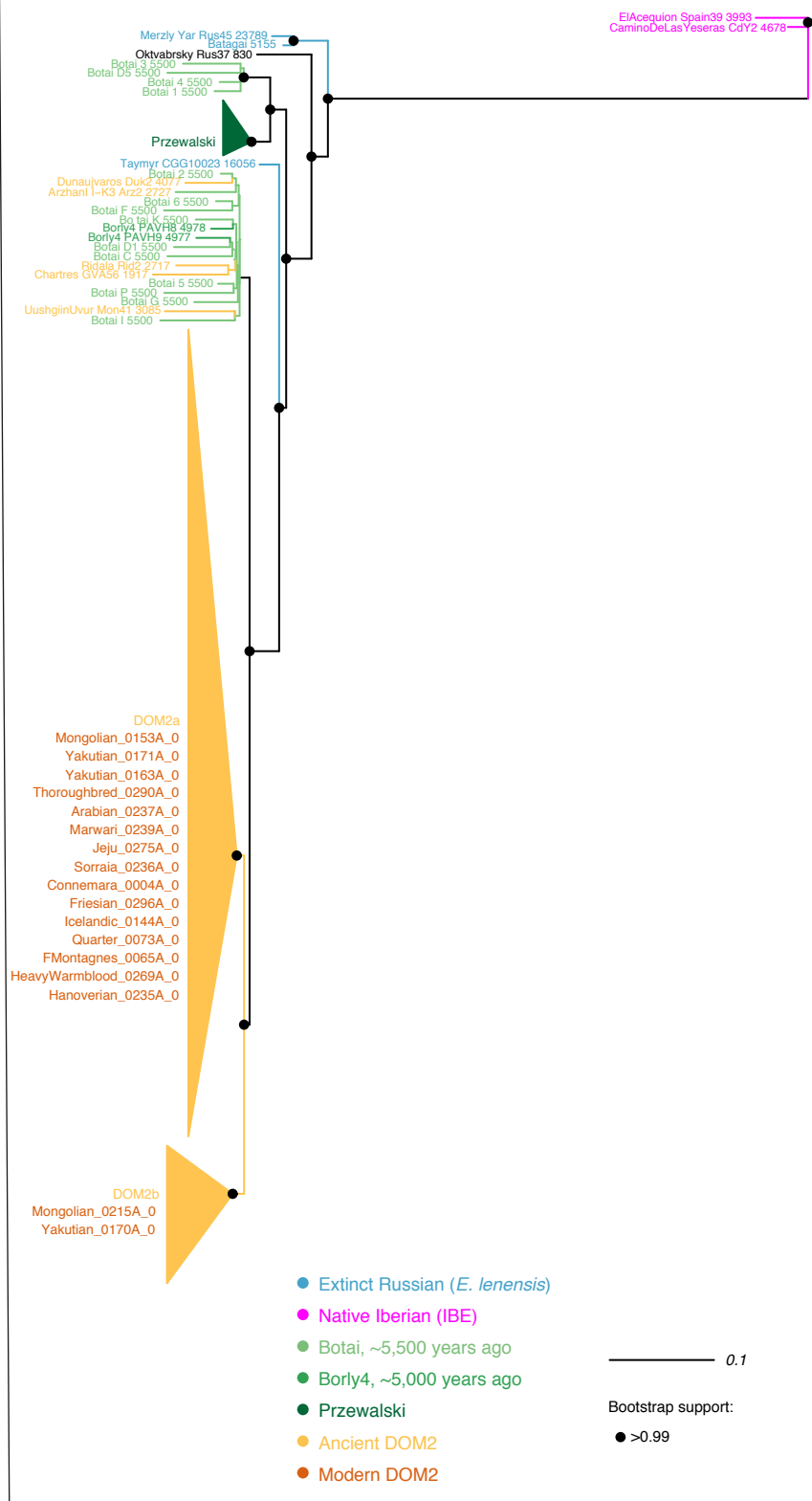


Figure 6



B

Click here to access/download;Figure;Fig6_05022019.pdf

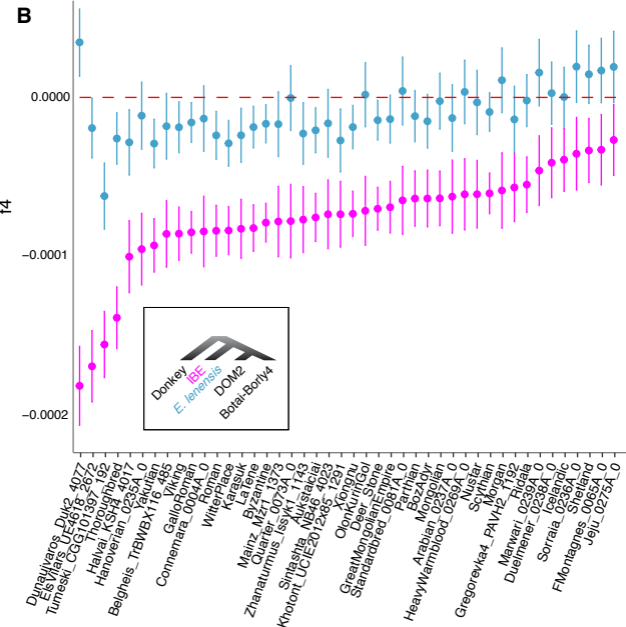
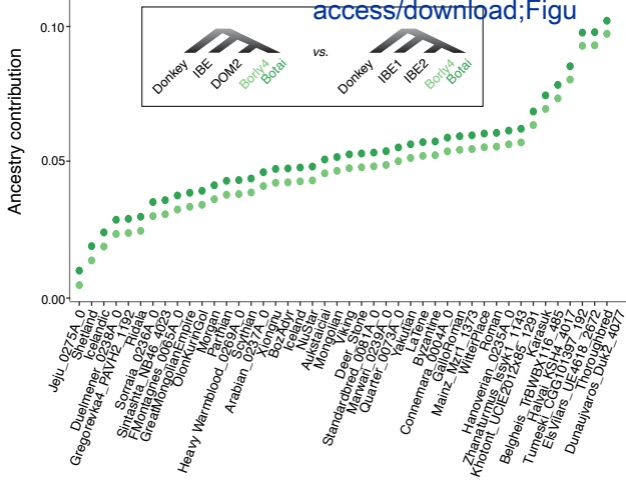


- Extinct Russian (*E. lenensis*)
- Native Iberian (IBE)
- Botai, ~5,500 years ago
- Bory4, ~5,000 years ago
- Przewalski
- Ancient DOM2
- Modern DOM2

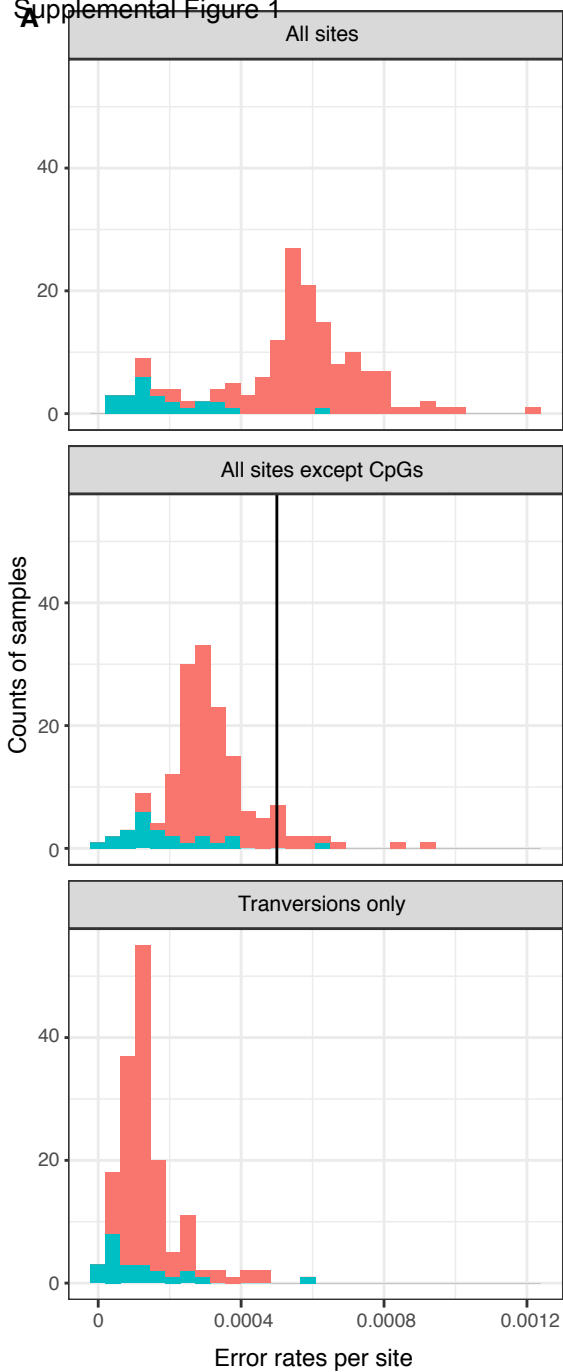
0.1
Bootstrap support:
● >0.99

Figure 7

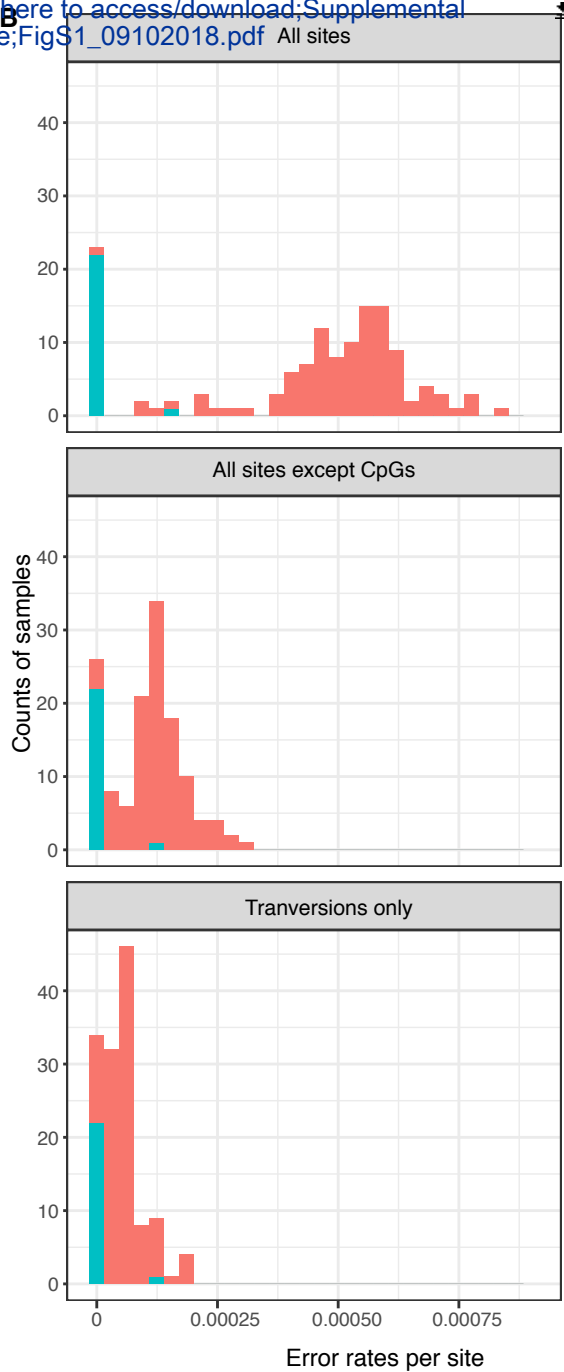
Click here to [access/download;Figure](#)

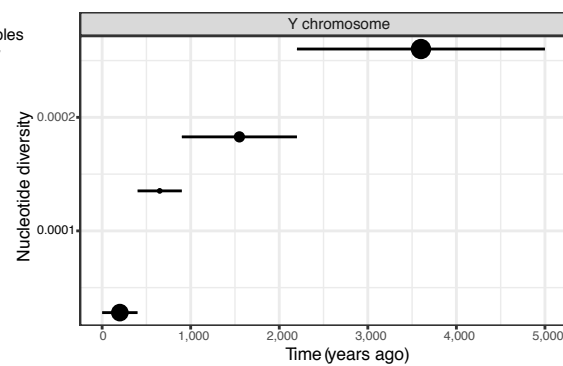
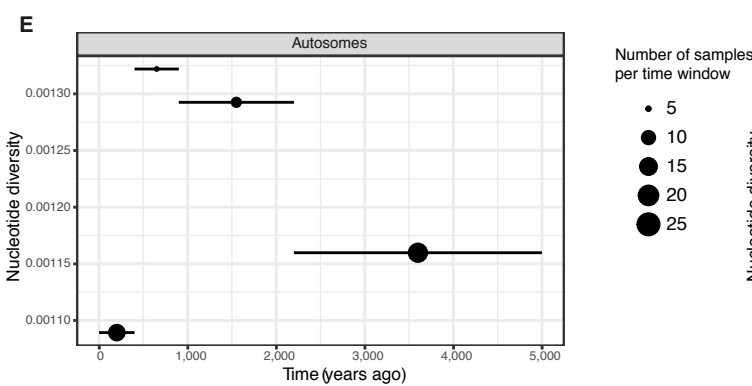
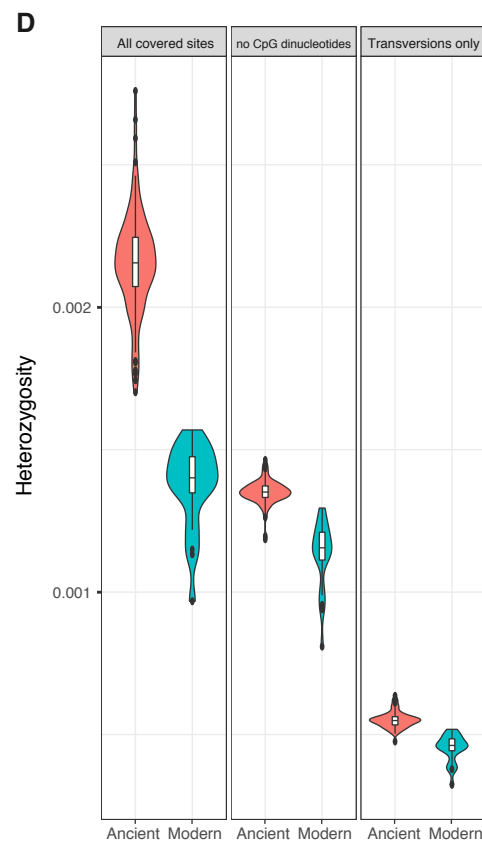
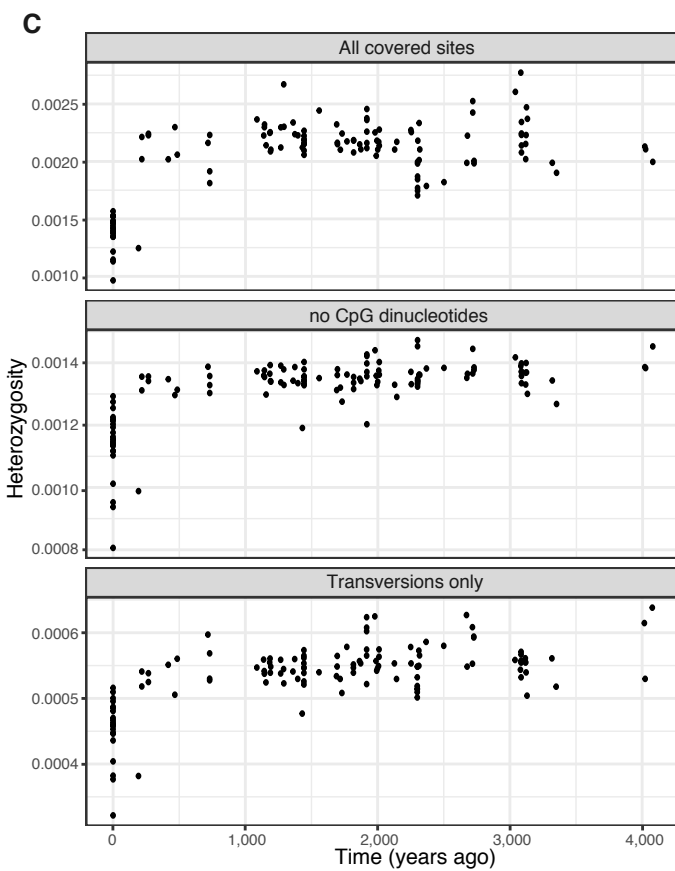
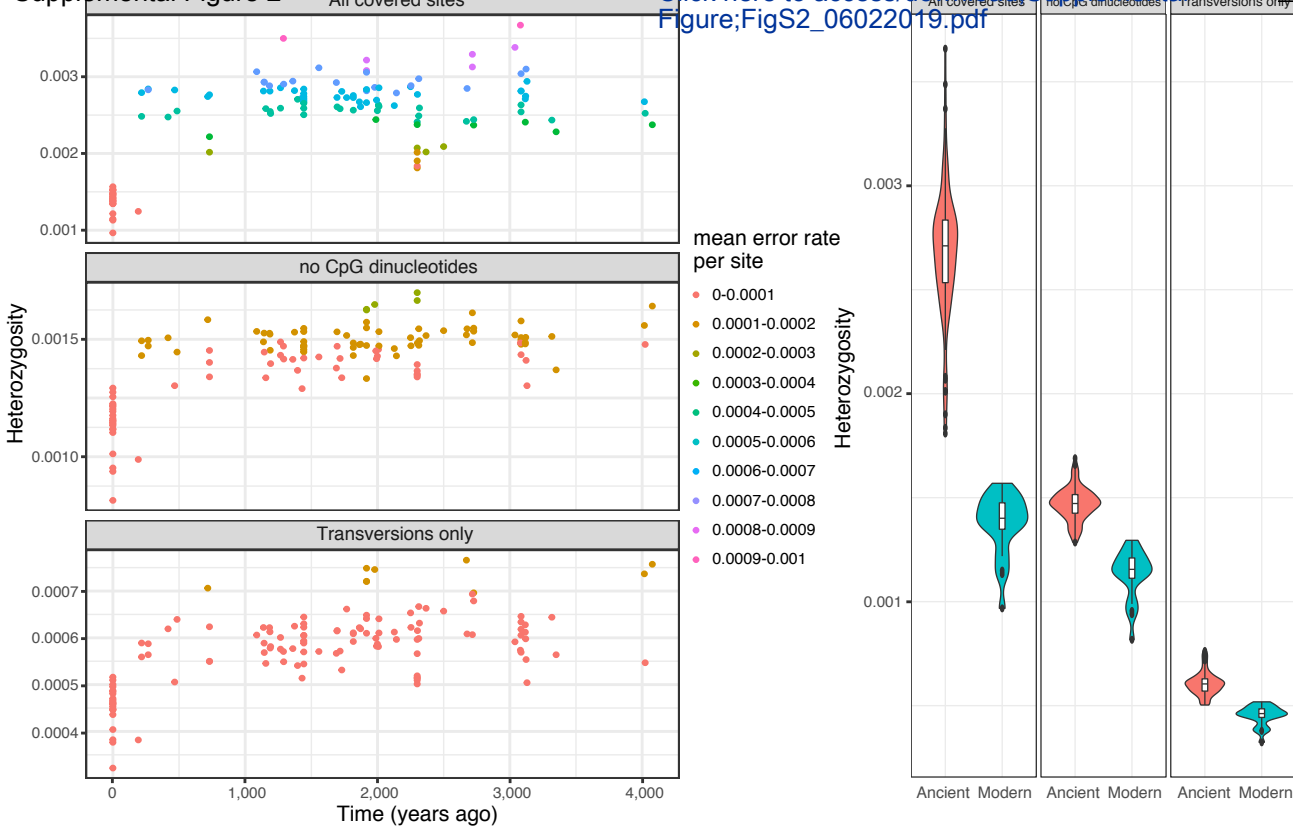


A Supplemental Figure 1



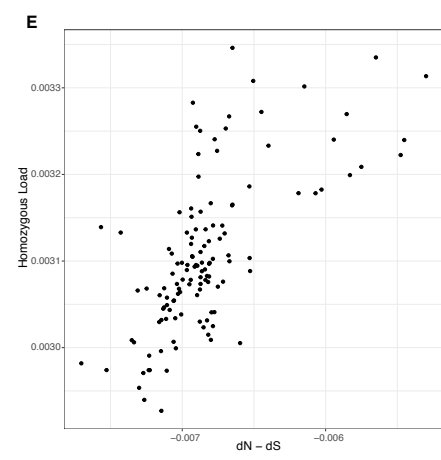
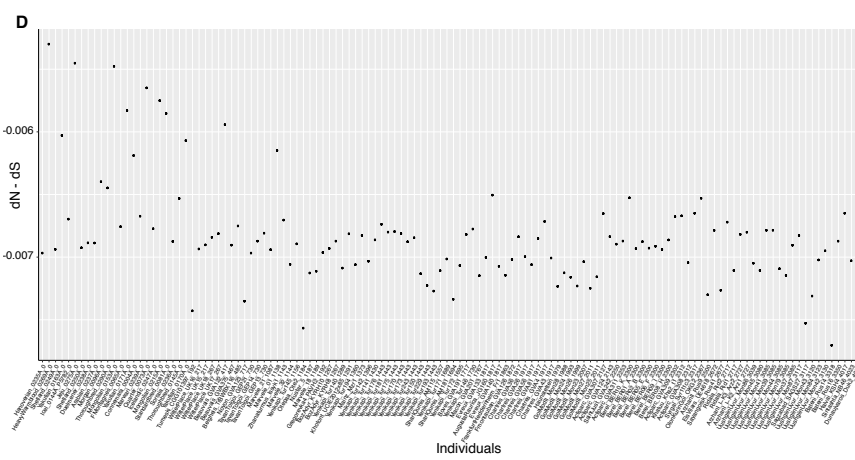
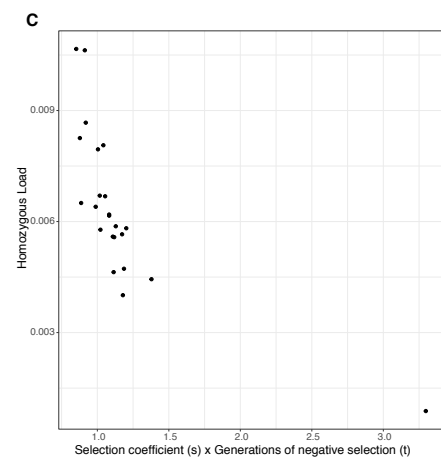
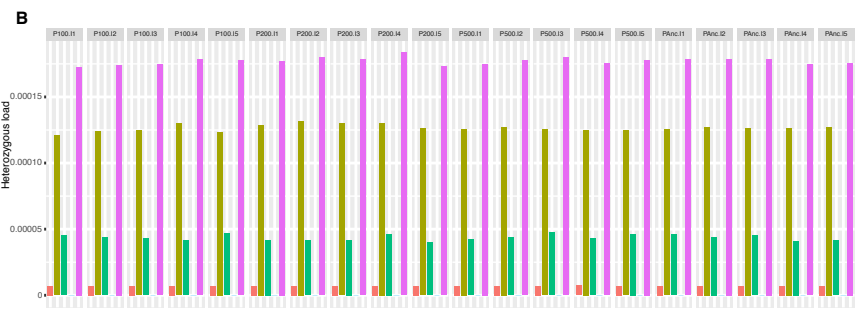
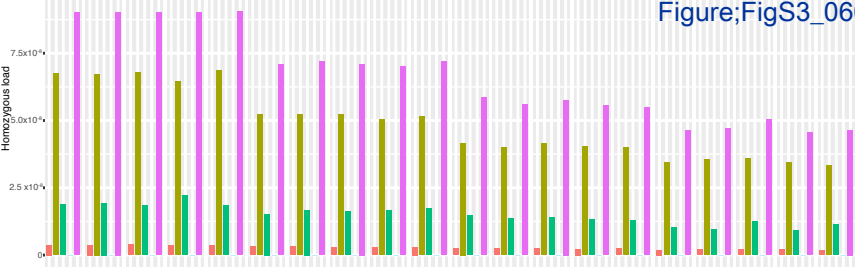
B [Click here to access/download/Supplemental Figure;FigS1_09102018.pdf](#)





Supplemental Figure 3

Click here to access/download/Supplemental Figure;FigS3_0602019.pdf



Supplemental Figure 4

[Click here to](#)

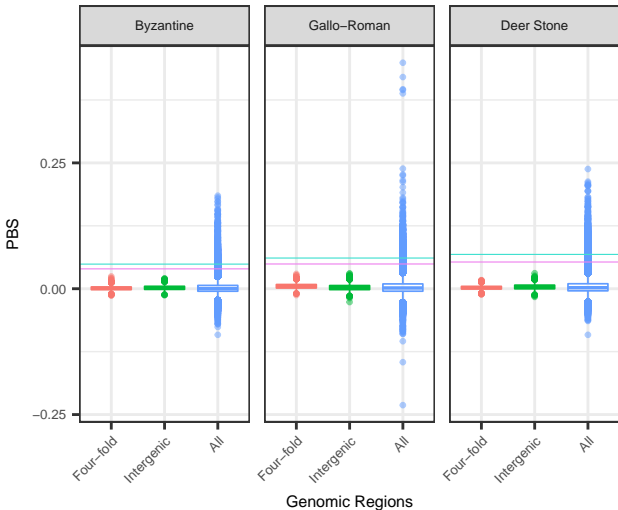


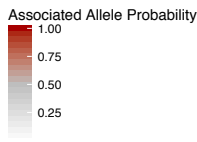
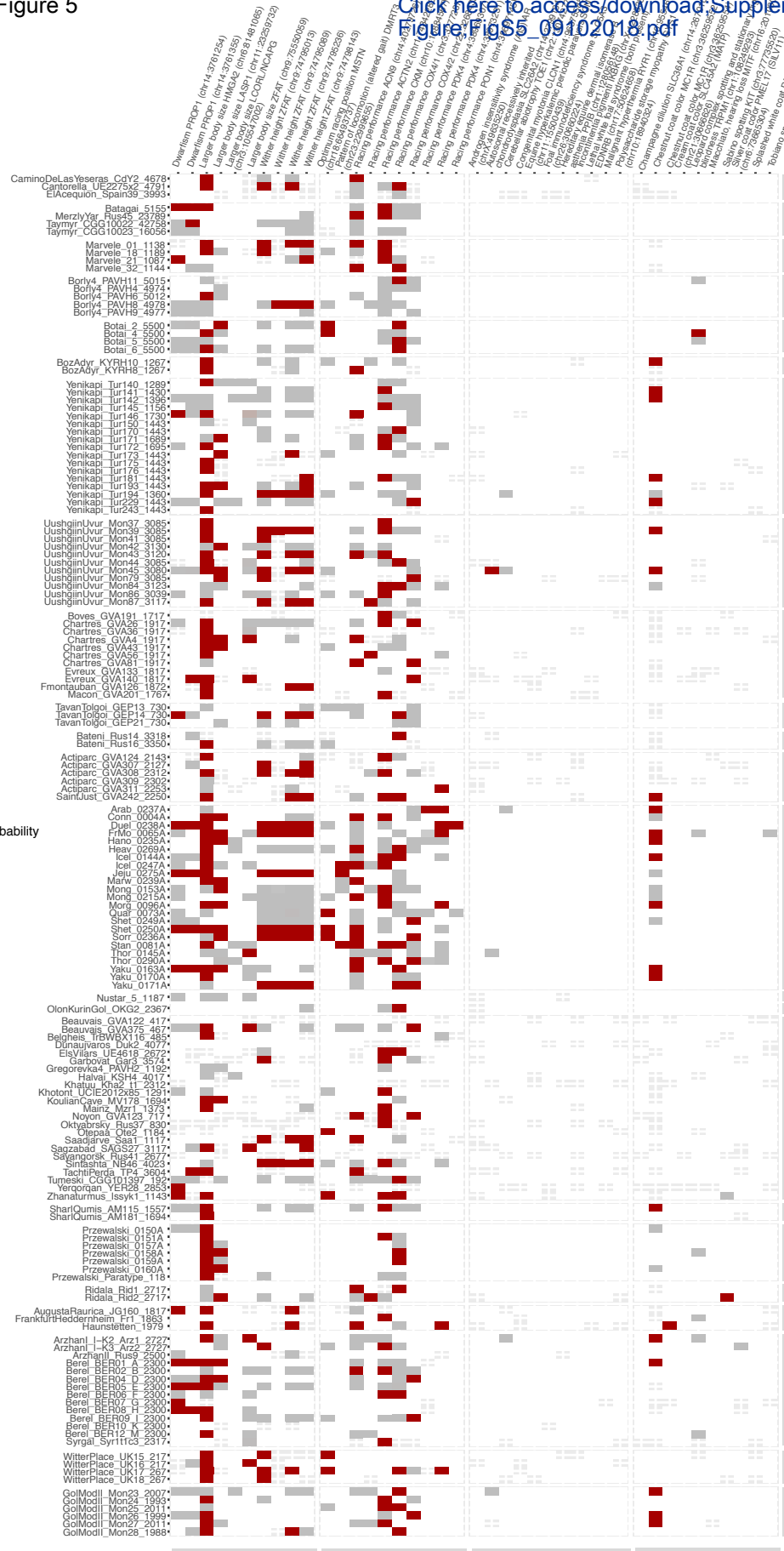
Cut-offs

— top-1000 windows

— top-100 windows

[access/download/Sup](#)





Coat Diseases Racing performances Size

Native Iberian

Archaic Russian

Aukstaiciai

Borly 4

Botai

Boz Adyr

Byzantine

Deer Stone

Gallo-Roman

Great Mongolian Empire

Karasuk

La Tène

Modern

Nustar

Olon Kurin Gol

Other ancient horses

Sassanid / Persian

Przewalski

Ridala

Roman

Pazyryk Scythian

Witter Place

Xiongnu

