Modern Northern Domestic Horses Carry Mitochondrial DNA Similar to Przewalski’s Horse

Laura Kvist 1 & Markku Niskanen 2

Accepted: 18 August 2020 / Published online: 5 September 2020
© The Author(s) 2020

Abstract
Several recent studies have suggested past gene flow between the Przewalski’s horse and modern domestic horse and questioned the wild origin of the Przewalski’s horse. Mitochondrial DNA has placed representatives of the Przewalski’s horse into three among the eighteen haplogroups detected from the modern horse. Of these, two haplogroups have so far been found exclusively in the Przewalski’s horse, while the one shared with the domestic horse includes captive individuals that have uncertain pedigrees. We recently found five domestic horse individuals of North European horse breeds to carry a mitochondrial haplogroup that was previously confined only to the Przewalski’s horse. These individuals were sequenced for 6039 bp of mitochondrial DNA and used, together with domestic and Przewalski’s horse sequences presenting all horse haplogroups, to examine the phylogenetic relationships and to date the divergence time between Przewalski’s and domestic horse clusters within this haplogroup. The divergence was dated to have likely occurred about 13,300–11,400 years ago, which coincides with the time of the Younger Dryas.

Keywords Finnhorse · Latvian horse · Divergence dating · Domestication · Haplogroup

Introduction
A recent study of the relationship of Przewalski’s horse (Equus przewalskii) genomes with ancient and modern domestic horse (Equus caballus) genomes suggested that Przewalski’s horses are actually feral descendents of domesticated horses of the Botai culture (c. 5700–5100 years ago (ya)). Modern domestic horse genomes were estimated to carry less than 3% of the ancestry of these Botai-derived horses, which however indicates small amount of gene flow from these horses to domestic horse in the past (Gaunitz et al. 2018). Further evidence of past gene flow between different equine groups comes from the finding that up to 13.7% of components of modern domestic horse can be found in the genomes of Przewalski’s horses (Der Sarkissian et al. 2015). Mitochondrial variation in domestic horse is extensive, with 17 different major haplogroups detected so far (named A-E and G-R in Achilli et al. 2012). From Przewalski’s horse, three mitochondrial haplogroups have been found (F, I, and Jprz), of which one (I) is shared with the domestic horse and two (F and Jprz) have so far been found exclusively in the Przewalski’s horse (Lippold et al. 2011; Achilli et al. 2012; Orlando et al. 2013; Der Sarkissian et al. 2015; Yoon et al. 2018). The existence of a shared mitochondrial haplogroup and close relatedness of the other Przewalski’s horse specific haplogroups with the modern domestic horse haplogroups further indicates past gene flow between these two.

We recently found evidence of four Finnhorse and one Latvian horse (Latvijas skirne) individuals out of 991 sequenced horses (mostly Finnhorses, but also including several breeds of riding horses, ponies, and harness trotters) to carry mitochondrial haplogroup F (Kvist et al. 2019), the haplogroup that was previously confined only to Przewalski’s horse. Finnhorse is a native breed that was believed to descend from ‘Northern Forest Type horses’, small sturdy horses adapted to cold climate (Peltonen and...
Saastamoinen 2007). It is closely related to other Nordic or eastern horse breeds, e.g., the Norwegian Fjord, North Swedish horse, Icelandic horse, Estonian Native horse, Mongolian horse, Yakutian horse, and Tuva horse (Petersen et al. 2013; Sild et al. 2019). The Latvian horse was formed by crossing native Latvian horses with western breeds, such as Oldenburg, Hannoverian, and Holstein horses (Ministry of Agriculture of Latvia 2003). It is assumed that the native Latvian horse was closely related to native Scandinavian breeds, e.g., the North Swedish horse and Dale Gudbrandsdal (Hendricks 1995), which are closely related also to the Finnhorse (Petersen et al. 2013).

Our aim here was to confirm the previous finding of presence of haplogroup F in modern northern horses that was based on a 631 bp fragment of the mitochondrial control region. Further, we aimed at dating the divergence forming Przewalski’s and modern domestic horse clusters within the F lineage and at inferring the divergence in relation to glacial history and horse domestication.

Material and Methods

The four Finnhorses belonging to haplogroup F were three geldings and one mare, which derived from two maternal lineages called Hilu 656PK and Elle 52810. Hilu, a red chestnut mare, was born in 1911 and was accepted into the Finnhorse studbook in 1915 in Kitee, eastern Finland. She left two offspring and this lineage has led to six mares born after year 2000. Elle was also a red chestnut mare born in 1939, owned by a farmer in the Teuva commune, in western Finland. This maternal lineage could be traced back for three more generations to a mare in Jalasjärvi, western Finland that was born in 1901. Elle was accepted into the Finnhorse studbook in 1944 in Kauhajoki, western Finland. She left four offspring, with 31 mares born after year 2000 from this lineage (Heppa-järjestelmä 2019; Sukuposti 2019). The Latvian Horse belonging to haplogroup F was a red bay Latvian horse mare that was imported to Finland in 2005. Unfortunately, we do not have a pedigree of this horse.

We sequenced 6039 bp of mitochondrial DNA from these four Finnhorses and the one Latvian horse. The regions selected for sequencing were chosen based on covering most of the informative sites in an alignment of Przewalski’s and domestic horse mitochondrial genomes. Primers used for amplification and sequencing, amplified regions of the mitochondrial DNA and PCR and sequencing protocols can be found in Supplementary material and Table S1. The obtained sequences were aligned with sequences fetched from GenBank (JN398377 – JN398457), presenting all the haplogroups A-E and G-R found in the modern domestic horse (Achilli et al. 2012) supplemented with sequences obtained from Przewalski’s horse (GenBank accession numbers KT221844, KT221845, KT368742 – KT368756, KT757761, HQ439484, JN398402, JN398403), two Przewalski’s horse x domestic horse hybrids (KT368757 and KT368758), one Yakutian Pony (HQ439467), and one Thoroughbred (Twilight KT757764) (Lippold et al. 2011; Achilli et al. 2012; Orlando et al. 2013; Der Sarkissian et al. 2015; Yoon et al. 2018; Table S2). After selection of the best substitution model (TN93 + G + I), a maximum likelihood tree (5 categories, Gamma distribution parameter = 0.8579 and proportion of invariant sites = 0.8212, including 1st, 2nd, and 3rd coding positions and tRNAs and control region as noncoding) was constructed with 1000 bootstraps in Mega X (Kumar et al. 2018).

In order to estimate the divergence time of the modern Finnhorse/Latvian horse lineage from Przewalski’s horse lineage within the haplogroup F, a timetree analysis was conducted in Mega X. The tree was inferred using the ReTime method (Tamura et al. 2018) using a) calibration to MRCA (most recent common ancestor; node separating haplogroup R from the rest A, 96520–170,400 years) or b) divergence between haplogroups F and G (14,910–32,180 years) as calibration points based on estimates from Achilli et al. (2012).

The new datasets generated during the current study are available in GenBank (https://www.ncbi.nlm.nih.gov/nuccore) by accession numbers MT476488-MT476587.

Results

In the phylogenetic tree, the four Finnhorses and the Latvian horse grouped together as a sister clade to Przewalski’s horse within lineage F. These two clades formed a strongly supported monophyletic group (bootstrap support 99%, Fig. 1). Based on time to MRCA, the divergence between Finnhorses and the Latvian Horse lineage from Przewalski lineage occurred 13,300 ya, and based on divergence between haplogroups F and G, it occurred 11,400 ya (Fig. 1).

Discussion

The wild horse was a common species in Europe, Northern Asia, and in North America until the Last Glacial Maximum (Leonardi et al. 2018). There were at least three distinct lineages of horses in Eurasia when the Last Glacial period commenced at the end of the Eemian period, 115,000 ya. These included two divergent and now extinct lineages, which lived in Iberia and Siberia, as well as the lineage that gave rise to the lineages leading to the extant domestic horse and the Przewalski’s horse. Estimates of nuclear divergences indicate that these extant horses diverged at some point between 43,800 and 35,400 ya (Fages et al. 2019). These estimates of divergence dates coincide with the formation of the so-called Khvalynian Sea, which together with the Ural Mountains likely contributed to population differentiation of ancestral populations of the European wild horse and the Przewalski’s horse.
This inland sea was formed when the Caspian Sea expanded due to meltwater from ice sheets and permafrost. The Khvalynian Sea coastline reached as far north as the southern Ural Mountains during two periods: 1) between ~35,000 ya and the onset of the Last Glacial Maximum (LGM) about ~26,500 ya and 2) between ~19,000 and ~12,800 ya (on this inland see, see Arslanov et al. 2016: fig. 1; Tudryn et al. 2016; Yanko-Hombach and Kislov 2018; Yanina et al. 2018).

The Khvalynian Sea was not a barrier for dispersal during the Last Glacial Maximum (the LGM, 26,500–19,000 ya, Clark et al. 2009) because the Caspian Sea water level dropped below the current level due to extreme aridity (Yanko-Hombach and Kislov 2018). However, this aridity could have made the region between the Caspian Sea and the Ural Mountains inhospitable. In any event, the Eurasian horse populations declined during this harsh climatic phase based on genomic evidence (Orlando 2015) and osteoarchaeological evidence (Leonardi et al. 2018). Coincidentally, divergence date estimate between haplogroups F and G (point estimate 23,550 years, Achilli et al. 2012) place places between the Caspian Sea and the Ural Mountains during two periods: 1) between ~35,000 ya and the onset of the Last Glacial Maximum (LGM) about ~26,500 ya and 2) between ~19,000 and ~12,800 ya (on this inland see, see Arslanov et al. 2016: fig. 1; Tudryn et al. 2016; Yanko-Hombach and Kislov 2018; Yanina et al. 2018).

Maternal lineages may have become differentiated via population divergence across the Bolling/Allerod-Younger Drys transition ~12,900 years ago as climate became rapidly much colder and drier, at some point during the Younger Dryas, or across the Younger Dryas-Holocene transition (marking the Pleistocene-Holocene transition) ~11,700 years ago as climate became rapidly much warmer and more humid.

Current evidence does not allow us to determine when F lineages carried by current Finnhorses and the Latvian Sports Horse arrived in the East Baltic region. Both breeds have been founded using native landraces and are believed to have some common ancestry (Hendricks 1995; Petersen et al. 2013; Sild et al. 2019). These lineages could have arrived in this region already before horse domestication, which occurred perhaps 5300–5500 years ago in the Pontic-Caspian Steppe (assuming that the Yamnaya culture indeed had domestic horses, Anthony 2007) and led to the current domestic horse (DOM2 in Gaunitz et al. 2018 and Fages et al. 2019). In this scenario, the domestic horse arriving with the Corded Ware culture about 4900 years ago derived these F lineages from the local wild horses likely still present in Estonia at that time according to Sommer et al. (2011) and Leonardi et al. (2018). Alternatively, these F lineages were present in the wild ancestors of the earliest domestic horses in the Pontic-Caspian steppe and dispersed elsewhere as the domestic horse dispersed. It is entirely possible that future ancient DNA studies will reveal that these F lineages used to be much more widespread until the relatively recent past and the lineage will be found in other modern breeds as well. However, thousands of modern horses and tens of horse breeds have already been extensively studied using mitochondrial markers. As an example, a GenBank search resulted into 2151 sequences with search words equus, caballus, mitochondrial, control, and region and to 8473 sequences with search words equus, caballus, and D-loop (search performed 12th of May 2020). Although mitochondrial studies have revealed lack of correspondence between breeds, mitochondrial lineages, and geography (e.g., Achilli et al. 2012; Hudson 2017; Hristov et al. 2017), the F lineage has not been reported in other breeds.

A discovery of F lineages among the horse of Europe is not entirely unexpected in light of a finding that the Y chromosome haplotype at present only carried by the Przewalski’s horse (Y-HT-2) was carried by some European wild horse and early domestic horse stallions (Wutke et al. 2018). More extensive genomic sampling of current horses may also lead to discoveries of additional F lineages among the domestic horse populations and even previously unknown uniparental lineages. For example, a recent study discovered that some of the Estonian Native horse stallions carry a unique Y-chromosomal haplogroup not previously described (Custaneda et al. 2019). Also, current aDNA studies of archaeological horse remains by several research groups may provide further information on the evolutionary history of the horse.

Acknowledgements. We are grateful for the horse owners providing us samples of their horses and for the help from Soile Alatalo and Hannele Parkkinen-Oinas in the lab.

Funding Open access funding provided by University of Oulu including Oulu University Hospital.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.
Sommer RS, Benecke N, Lõugas L, Nelle O, Schmölcke U (2011) Holocene survival of the wild horse in Europe: a matter of open landscape. J Quat Sci 26:805-812

Sukuposti (2019) http://www.sukuposti.net/site/index. Accessed 14 October 2019

Tamura K, Qiqing T, Kumar S (2018) Theoretical foundation of the RelTime method for estimating divergence times from variable evolutionary rates. Mol Biol Evol 35:1770-1782

Tudryn A, Leroy SAG, Toucanne S, Gilbert-Brunet E, Tucholka P, Lavrushin YA, Dafaure O, Miska S, Bayon G (2016) The Ponto-Caspian basin as a final trap for southeastern Scandinavian Ice-Sheet meltwater. Quat Sci Rev 148:29-43

Wutke S, Sandoval-Castellanos E, Benecke N, Döhle H-J, Friederich S, Gonzalez J, Hofreiter M, Lõugas L, Magnell O, Malaspinas AS, Morales-Muñiz A, Orlando L, Reissmann M, Trinks A, Ludwig A (2018) Decline of genetic diversity in ancient domestic stallions in Europe. Sci Adv 2018;4:eaap9691

Yanina T, Sorokin V, Bezrodnykh Yu, Romanyuk B (2018) Late Pleistocene climatic events reflected in the Caspian Sea geological history (based on drilling data). Quat Int 465 (Part A):130-141

Yanko-Hombach V, Kislov A (2018) Late Pleistocene – Holocene sea-level dynamics in the Caspian and Black seas: data synthesis and paradoxical interpretations. Quat Int 465:63-71

Yoon SH, Lee W, Ahn H, Caetano-Anolles K, Park KD, Kim H (2018) Origin and spread of thoroughbred racehorses inferred from complete mitochondrial genome sequences: phylogenomic and Bayesian coalescent perspectives. PLoS One 13:9:e0203917