## A new genus of horse from Pleistocene North America

Peter D. Heintzman ${ }^{\text {a,b*}}$, Grant D. Zazula ${ }^{\mathrm{c}}$, Ross D.E. MacPhee ${ }^{\mathrm{d}}$, Eric Scott ${ }^{\mathrm{e}}$, James A. Cahill ${ }^{\mathrm{a}}$, Brianna K. McHorse ${ }^{\text {f }}$, Joshua D. Kapp ${ }^{\text {a }}$, Mathias Stiller ${ }^{\text {a,g }}$, Matthew J. Wooller ${ }^{\text {h,i }}$, Ludovic Orlando ${ }^{\mathrm{j}, \mathrm{k}}$, John R. Southon ${ }^{1}$, Duane G. Froese ${ }^{\mathrm{m}}$, Beth Shapiro ${ }^{\mathrm{a}, \mathrm{n}^{*}}$

${ }^{\text {a }}$ Department of Ecology and Evolutionary Biology, University of California Santa Cruz, Santa Cruz, CA
95064
${ }^{\mathrm{b}}$ Tromsø University Museum, UiT - The Arctic University of Norway, 9037 Tromsø, Norway
${ }^{\mathrm{c}}$ Yukon Palaeontology Program, Government of Yukon, Whitehorse, YT Y1A 2C6, Canada
${ }^{\text {d }}$ Department of Mammalogy, Division of Vertebrate Zoology, American Museum of Natural History, New York, NY 10024
${ }^{\mathrm{e}}$ Dr. John D. Cooper Archaeological and Palaeontological Center, California State University Fullerton, Fullerton, CA 92831
${ }^{\mathrm{f}}$ Department of Organismal and Evolutionary Biology, Harvard University, Cambridge, MA
${ }^{\mathrm{g}}$ Department of Translational Skin Cancer Research, German Consortium for Translational Cancer Research, D-45141 Essen, Germany
${ }^{\text {h }}$ College of Fisheries and Ocean Sciences, University of Alaska Fairbanks, 905 N. Koyukuk Dr., Fairbanks AK 99775 USA
${ }^{\text {i }}$ Alaska Stable Isotope Facility, Water and Environmental Research Center, University of Alaska Fairbanks, 306 Tanana Loop, Fairbanks AK 99775 USA
${ }^{\mathrm{j}}$ Centre for GeoGenetics, Natural History Museum of Denmark, Øster Voldgad 5-7, 1350K Copenhagen, Denmark
${ }^{\mathrm{k}}$ Université de Toulouse, Université Paul Sabatier (UPS), Laboratoire AMIS, CNRS UMR 5288, Toulouse, France
${ }^{1}$ Keck-CCAMS Group, Earth System Science Department, University of California, Irvine, CA 92697
${ }^{\mathrm{m}}$ Department of Earth and Atmospheric Sciences, University of Alberta, Edmonton, AB T6G 2E3, Canada
${ }^{n}$ UCSC Genomics Institute, University of California Santa Cruz, Santa Cruz, CA 95064
*Corresponding authors: Peter D. Heintzman: peteheintzman@gmail.com; Beth Shapiro:
bashapir@ucsc.edu


#### Abstract

The extinct "New World stilt-legged", or NWSL, equids constitute a perplexing group of Pleistocene horses endemic to North America. Their slender distal limb bones resemble those of Asiatic asses, such as the Persian onager. Previous palaeogenetic studies, however, have suggested a closer relationship to caballine horses than to Asiatic asses. Here, we report complete mitochondrial and partial nuclear genomes from NWSL equids from across their geographic range. Although multiple NWSL equid species have been named, our palaeogenomic and morphometric analyses support the idea that there was only a single species of middle to late Pleistocene NWSL equid, and demonstrate that it falls outside of crown group Equus. We therefore propose a new genus, Haringtonhippus, for the sole species H. francisci. Our combined genomic and phenomic approach to resolving the systematics of extinct megafauna will allow for an improved understanding of the full extent of the terminal Pleistocene extinction event.


## Introduction

The family that includes modern horses, asses, and zebras, the Equidae, is a classic model of macroevolution. The excellent fossil record of this family clearly documents its $\sim 55$ million year evolution from dog-sized hyracotheres through many intermediate forms and extinct offshoots to present-day Equus, which comprises all living equid species (MacFadden, 1992). The downside of this excellent fossil record is that many dubious fossil equid taxa have been erected, a problem especially acute within Pleistocene Equus of North America (Macdonald \& Toohey, 1992). While numerous species are described from the fossil record, molecular data suggest that most belonged to, or were closely related to, a single, highly variable stout-legged caballine species that includes the domestic horse, E. caballus (Weinstock et al., 2005). The enigmatic and extinct "New World stilt-legged" (NWSL) forms, however, exhibit a perplexing mix of morphological characters, including slender, stilt-like distal limb bones with narrow hooves reminiscent of extant Eurasian hemionines, the Asiatic wild asses (E. hemionus, E. kiang) (Eisenmann, 1992; Eisenmann et al., 2008; Harington \& Clulow, 1973; Lundelius \& Stevens, 1970; Scott, 2004), and dentitions that have been interpreted as more consistent with either caballine horses (Lundelius \& Stevens, 1970) or hemionines (MacFadden, 1992).

On the basis of their slender distal limb bones, the NWSL equids have traditionally been considered as allied to hemionines (e.g. (Eisenmann et al., 2008; Guthrie, 2003; Scott, 2004; Skinner \& Hibbard, 1972)). Palaeogenetic analyses based on mitochondrial DNA (mtDNA) have, however, consistently placed NWSL equids closer to caballine horses (Der Sarkissian et al., 2015; Orlando et al., 2008, 2009; Vilstrup et al., 2013; Weinstock et al., 2005). The current mtDNA-based phylogenetic model therefore suggests that the stilt-legged morphology arose independently in the New and Old Worlds (Weinstock et al., 2005), and may represent
convergent adaptations to arid climates and habitats (Eisenmann, 1985). However, these models have been based on two questionable sources. The first is based on 4 short control region sequences ( $<1000$ base pairs, bp; (Weinstock et al., 2005)), a data type that can be unreliable for resolving the placement of major equid groups (Der Sarkissian et al., 2015; Orlando et al., 2009). The second consist of 2 mitochondrial genome sequences (Vilstrup et al., 2013) that are either incomplete or otherwise problematic (see Results). Given continuing uncertainty regarding the phylogenetic placement of NWSL equids-which impedes our understanding of Pleistocene equid evolution in general-we therefore sought to resolve their position using multiple mitochondrial and partial nuclear genomes from specimens representing as many parts of late Pleistocene North America as possible.

The earliest recognized NWSL equid fossils date to the late Pliocene/early Pleistocene (~2-3 million years ago, Ma) of New Mexico (Azzaroli \& Voorhies, 1993; Eisenmann, 2003; Eisenmann et al., 2008)). Middle and late Pleistocene forms tended to be smaller in stature than their early Pleistocene kin, and ranged across southern and extreme northwestern North America (i.e., eastern Beringia, which includes Alaska, USA and Yukon Territory, Canada). NWSL equids have been assigned to several named species, such as E. tau Owen 1869, E. francisci Hay 1915, E. calobatus Troxell 1915, and E. (Asinus) cf. kiang, but there is considerable confusion and disagreement regarding their taxonomy. Consequently, some researchers have chosen to refer to them collectively as Equus (Hemionus) spp. (Guthrie, 2003; Scott, 2004), or avoid a formal taxonomic designation altogether (Der Sarkissian et al., 2015; Vilstrup et al., 2013; Weinstock et al., 2005). Using our phylogenetic framework and comparisons between specimens identified by palaeogenomics and/or morphology, we attempted to determine the taxonomy of middle-late Pleistocene NWSL equids.

Radiocarbon $\left({ }^{14} \mathrm{C}\right)$ dates from Gypsum Cave, Nevada, confirm that NWSL equids persisted in areas south of the continental ice sheets during the last glacial maximum (LGM; ~26-19 thousand radiocarbon years before present (ka BP; (Clark et al., 2009)) until near the terminal Pleistocene, $\sim 13$ thousand radiocarbon years before present ( ${ }^{14} \mathrm{C}$ ka BP) (Weinstock et al., 2005), soon after which they became extinct, along with their caballine counterparts and most other coeval species of megafauna (Koch \& Barnosky, 2006). This contrasts with dates from unglaciated eastern Beringia, where NWSL equids were seemingly extirpated locally during a relatively mild interstadial interval centered on $\sim 31{ }^{14} \mathrm{C}$ ka BP (Guthrie, 2003), thus prior to the LGM (Clark et al., 2009), final loss of caballine horses (Guthrie, 2003, 2006), and arrival of humans in the region (Guthrie, 2006). The apparently discrepant extirpation chronology between NWSL equids south and north of the continental ice sheets implies that their populations responded variably to demographic pressures in different parts of their range, which is consistent with results from some other megafauna (Guthrie, 2006; Heintzman et al., 2015; Zazula et al., 2014). To further test this extinction chronology, we generated new radiocarbon dates from eastern Beringian NWSL equids.

We analyzed 26 full mitochondrial genomes and 17 partial nuclear genomes from late Pleistocene NWSL equids, which revealed that individuals from both eastern Beringia and southern North America form a single well-supported clade that falls outside the diversity of Equus and diverged from the lineage leading to Equus during the latest Miocene or early Pliocene. This novel and robust phylogenetic placement warrants the recognition of NWSL equids as a distinct genus, which we here name Haringtonhippus. After reviewing potential species names, and conducting morphometric and anatomical comparisons, we determined that francisci Hay is the most well-supported species name, and we therefore refer the analyzed

NWSL equid specimens to $H$. francisci. New radiocarbon dates revealed that $H$. francisci was extirpated in eastern Beringia $\sim 14{ }^{14} \mathrm{C}$ ka BP. In light of our analyses, we review the PlioPleistocene evolutionary history of equids, and the implications for the systematics of equids and other Pleistocene megafauna.

## Results

## Phylogeny of North American late Pleistocene and extant equids

We reconstructed whole mitochondrial genomes from 26 NWSL equids and four New World caballine Equus (two E. lambei, two E. cf. scotti). Using these and mitochondrial genomes of representatives from all extant and several late Pleistocene equids, we estimated a mitochondrial phylogeny, using a variety of outgroups (Appendix 1, Supplementary tables 1-2 and Supplementary file 1). The resulting phylogeny is mostly consistent with previous studies (Der Sarkissian et al., 2015; Vilstrup et al., 2013), including confirmation of NWSL equid monophyly (Weinstock et al., 2005). However, we recover a strongly supported placement of the NWSL equid clade outside of crown group diversity (Equus), but closer to Equus than to Hippidion (Figure 1 and Appendix 2, Figure 1-Figure supplement 1a and Supplementary tables 1-3). In contrast, previous palaeogenetic studies placed the NWSL equids within crown group Equus, closer to caballine horses than to non-caballine asses and zebras (Der Sarkissian et al., 2015; Orlando et al., 2008, 2009; Vilstrup et al., 2013; Weinstock et al., 2005). To explore possible causes for this discrepancy, we reconstructed mitochondrial genomes from previously sequenced NWSL equid specimens and used a maximum likelihood evolutionary placement algorithm (Berger et al., 2011) to place these published sequences in our phylogeny a posteriori. These analyses suggested that previous results were likely due to a combination of outgroup choice and the use of short, incomplete, or problematic mtDNA sequences (Appendix 2 and Supplementary table 4).

To confirm the mtDNA result that NWSL equids fall outside of crown group equid diversity, we sequenced and compared partial nuclear genomes from 17 NWSL equids to a caballine (horse) and a non-caballine (donkey) reference genome. After controlling for reference
genome and ancient DNA fragment length artifacts (Appendices 1-2), we examined differences in relative private transversion frequency between these genomes (Appendix 1-figure 1). We found that the relative private transversion frequency for NWSL equids was $\sim 1.4-1.5$ times greater than that for horse or donkey (Appendix 2, Supplementary table 5, Figure 1-Figure supplement 2, and Supplementary file 2). This result supports the placement of NWSL equids as sister to the horse-donkey clade (Figure 1-Figure supplement 3), the latter of which is representative of living Equus diversity (e.g. (Der Sarkissian et al., 2015; Jónsson et al., 2014)), and is therefore congruent with the mitochondrial genomic analyses.


Figure 1. Phylogeny of extant and middle-late Pleistocene equids, as inferred from the Bayesian analysis of full mitochondrial genomes. Purple node bars illustrate the $95 \%$ highest posterior density of node heights and are shown for nodes with $>0.99$ posterior probability support. The range of divergence estimates derived from the analysis of nuclear genomes is shown by the thicker lime green node bars ((Orlando et al., 2013); this study). Major nodes are labelled with
boxed numbers. The thicknesses of nodes 2 and 3 represent the range between the median nuclear and mitochondrial genomic divergence estimates. Branches are coloured based on species provenance and the most parsimonious biogeographic scenario given the data, with gray indicating ambiguity. Fossil record occurrences for major represented groups (including South American Hippidion, New World stilt-legged equids, and Old World Sussemiones) are represented by the geographically-coloured bars, with fade indicating uncertainty in the first appearance datum (after (Eisenmann et al., 2008; Forsten, 1992; O’Dea et al., 2016; Orlando et al., 2013) and references therein). The Asiatic ass species (E. kiang, E. hemionus) are not reciprocally monophyletic based on the analyzed mitochondrial genomes, and so the Asiatic ass clade is shown as 'E. kiang + hemionus'. Daggers denote extinct taxa. NW: New World.

## A



Figure 1-Figure supplement 1. An example maximum likelihood (ML) phylogeny of equid mitochondrial genomes. This topology resulted from the analysis of mtDNA data set 3 (see

Appendix 1) with all partitions and Hippidion included, and dog and ceratomorphs as outgroup (not shown). Numbers above branches are Bayesian posterior probability support values from equivalent MrBayes and BEAST analyses, with those below indicating ML bootstrap values calculated in RAxML, and are shown for major nodes. (A) Full phylogeny of the analyzed equid sequences. (B) The Haringtonhippus (NWSL equid) clade, with tips color coded by geographic origin: Beringia, blue; contiguous USA, red. Tips in bold were included in the BEAST analysis (see also Supplementary file 1).


Figure 1-Figure supplement 2. A comparison of relative private transversion frequencies between the nuclear genomes of a caballine Equus (horse, E. caballus; green), a non-caballine Equus (donkey, E. asinus; red), and 17 NWSL equids (=Haringtonhippus francisci; blue) at different read lengths, with reads divided into 10 base pair (bp) bins. Analyses are based on alignment to the horse (A) or donkey (B) genome coordinates. To account for bins with low data content, we only display comparisons with at least 200,000 observable sites.


Figure 1-Figure supplement 3. Calculation of divergence date estimates from nuclear genome data. Relative branch lengths are from Supplementary table 5. Minimum (darker blue) and maximum (lighter blue) estimates are shown for the NWSL equid branch.

## Divergence times of Hippidion, NWSL equids, and Equus

We estimated the divergence times between the lineages leading to Hippidion, the NWSL equids, and Equus. We first applied a Bayesian time-tree approach to the whole mitochondrial genome data. This gave divergence estimates for the Hippidion-NWSL/Equus split (node 1) at 5.15-7.66 Ma, consistent with (Der Sarkissian et al., 2015), the NWSL-Equus split (node 2) at 4.09-5.13 Ma, and the caballine/non-caballine Equus split (node 3) at 3.77-4.40 Ma (Figure 1 and Supplementary table 3). These estimates suggest that the NWSL-Equus mitochondrial split
occurred only $\sim 500$ thousand years (ka) prior to the caballine/non-caballine Equus split. We then estimated the NWSL-Equus divergence time using relative private transversion frequency ratios between the nuclear genomes, assuming a caballine/non-caballine Equus divergence estimate of 4-4.5 Ma (Orlando et al., 2013) and a genome-wide strict molecular clock (following (Heintzman et al., 2015)). This analysis yielded a divergence estimate of 4.87-5.69 Ma (Figure 1-Figure supplement 3), which overlaps with that obtained from the relaxed clock analysis of whole mitochondrial genome data (Figure 1). These analyses suggest that the NWSL equid and Equus clades diverged during the latest Miocene or early Pliocene (4.1-5.7 Ma; late Hemphillian or earliest Blancan).

## Systematic Palaeontology

Previous palaeontological and palaeogenetic studies have uniformly placed NWSL equids within genus Equus (Der Sarkissian et al., 2015; Harington \& Clulow, 1973; Orlando et al., 2008, 2009; Scott, 2004; Vilstrup et al., 2013; Weinstock et al., 2005). This, however, conflicts with the phylogenetic signal provided by palaeogenomic data, which strongly suggest that NWSL equids fall outside the confines of this genus. Nor is there any morphological or genetic evidence warranting the assignment of NWSL equids to an existing extinct taxon such as Hippidion. We therefore erect a new genus for NWSL equids, Haringtonhippus, as defined and delimited below:

Order: Perissodactyla, Owen 1848
Family: Equidae, Linnaeus 1758
Subfamily: Equinae, Steinmann \& Döderlein 1890
Tribe: Equini, Gray 1821

Genus: Haringtonhippus, gen. nov.

Type species. Haringtonhippus francisci Hay 1915.

Etymology. The new genus is named in honor of C. Richard Harington, who first described NWSL equids from eastern Beringia (Harington \& Clulow, 1973). 'Hippus' is from the Greek word for horse, and so Haringtonhippus is implied to mean 'Harington's horse'.

Holotype. A partial skeleton consisting of a complete cranium, mandible, and a stilt-legged third metatarsal (MTIII) (Figure 2a and Figure 2-Figure supplement 1b), which is curated at the University of Texas, Austin (Texas Memorial Museum (TMM) 34-2518). This specimen is the holotype of ' $E$ '. francisci, originally described by Hay (1915), and is from the middle Pleistocene Lissie Formation of Wharton County, Texas (Hay, 1915; Lundelius \& Stevens, 1970).).


Figure 2. Morphological analysis of extant and middle-late Pleistocene equids. (A) Crania of Haringtonhippus francisci, upper: LACM(CIT) 109/156450 from Nevada, lower: TMM 34-2518 from Texas. (B) From upper to lower, third metatarsals of: H. francisci (YG 401.268), E. lambei (YG 421.84), and E. cf. scotti (YG 198.1) from Yukon. Scale bar is 5 cm . (C) Principal component analysis of selected third metatarsals from extant and middle-late Pleistocene equids, showing clear clustering of stilt-legged (hemionine(-like) Equus and H. francisci) from stoutlegged (caballine Equus) specimens (see also Supplementary file 4). The first and second principal components explain $95 \%$ of the variance. Specimen identifications were derived from DNA (D) and/or morphology (M).


Figure 2-Figure supplement 1. The two crania assigned to H. francisci. (A) LACM(CIT)
109/156450 from Nevada, identified through mitochondrial and nuclear palaeogenomic analysis.
Upper: right side (reflected for comparison), lower: left side. (B) Part of the H. francisci holotype, TMM 34-2518 from Texas.


Figure 2-Figure supplement 2. Comparison between the limb bones of H. francisci, E. lambei, and E. cf. scotti from Yukon. (A) Third metatarsals from H. francisci (upper; YG 401.268), E. lambei (middle; YG 421.84), and E. cf. scotti (lower; YG 198.1). (B) Third metacarpals from H. francisci (upper; YG 404.663), E. lambei (middle; YG 109.6), and E. cf. scotti (lower; YG 378.15). (C) Proximal fragments of radii from H. francisci (left; YG 303.1085), and E. lambei (right; YG 303.325). (D) First phalanges from H. francisci (left; YG 130.3), E. lambei (middle; YG 404.22), and E. cf. scotti (right; YG 168.1).


Figure 2-Figure supplement 3. An example equid metacarpal from Natural Trap Cave, Wyoming (KU 47800; JK260), originally referred to Equus sp., but here identified as $H$. francisci on the basis of mitochondrial and nuclear genome data. We note the relative slenderness of this specimen, which is comparable to YG 404.663 (H. francisci) from Yukon in Figure 2-Figure supplement 2.


Figure 2-Figure supplement 4. An example femur of $H$. francisci from Gypsum Cave, Nevada (LACM(CIT)109/150708; JW277 / JK166), originally identified by Weinstock and colleagues (2005).

Referred material. On the basis of mitochondrial and nuclear genomic data, we assign the following material to Haringtonhippus: a cranium, femur, and MTIII (LACM(CIT); Nevada); three MTIIIs and three third metacarpals (MCIII) (KU; Wyoming); two radii, 12 MTIII, three MCIII, a metapodial, and a first phalanx (YG; Yukon Territory) (Figure 2-Figure supplements 14 and Supplementary file 1 ; (Weinstock et al., 2005)). This material includes 5 males and at least 6 females (Appendix 2, Supplementary table 6 and Supplementary file 3). We further assign MTIII specimens from Yukon Territory ( $\mathrm{n}=13$ ), Wyoming ( $\mathrm{n}=57$ ), and Nevada ( $\mathrm{n}=4$ ) to Haringtonhippus on the basis of morphometric analysis (Figure 2c and Supplementary file 4). We tentatively assign 19 NWSL equid metapodial specimens from Alaska (Guthrie, 2003) to Haringtonhippus, but note that morphometric and/or palaeogenomic analysis would be required to confirm this designation.

Geographic and temporal distribution. Haringtonhippus is known only from the Pleistocene of North America (Figure 1). In addition to the middle Pleistocene holotype from Texas, Haringtonhippus is confidently known from the late Pleistocene of Yukon Territory (Klondike), Wyoming (Natural Trap Cave), and Nevada (Gypsum Cave, Mineral Hill Cave), and is tentatively registered as present in Alaska (Appendix 2, Supplementary file 1, and Supplementary table 4; (Vilstrup et al., 2013; Weinstock et al., 2005)).

To investigate the last appearance date (LAD) of Haringtonhippus in eastern Beringia, we obtained new radiocarbon dates from 17 Yukon Territory fossils (Appendix 1 and Supplementary file 1). This resulted in three statistically-indistinguishable radiocarbon dates of $\sim 14.4{ }^{14} \mathrm{C}$ ka BP (derived from two independent laboratories) from a metacarpal bone (YG 401.235) of Haringtonhippus, which represents this taxon's LAD in eastern Beringia
(Supplementary file 1). The LAD for North America as a whole is based on two dates of $\sim 13.1$ ${ }^{14} \mathrm{C}$ ka BP from Gypsum Cave, Nevada (Supplementary file 1; (Weinstock et al., 2005)).

Mitogenomic diagnosis. Haringtonhippus is the sister genus to Equus (equid crown group), with Hippidion being sister to the Haringtonhippus-Equus clade (Figure 1). Haringtonhippus can be differentiated from Equus and Hippidion by 178 synapomorphic positions in the mitochondrial genome, including four insertions and 174 substitutions (Appendix 2, Supplementary table 7 and Supplementary file 5). We caution that these synapomorphies are tentative and will likely be reduced in number as a greater diversity of mitochondrial genomes for extinct equids become available.

Morphological comparisons of third metatarsals. We used morphometric analysis of caballine/stout-legged Equus and stilt-legged equids (hemionine/stilt-legged Equus,

Haringtonhippus) MTIIIs to determine how confidently these groups can be distinguished (Figure 2c). Using logistic regression on principal components, we find a strong separation that can be correctly distinguished with $98.2 \%$ accuracy (Appendix 2 and Supplementary file 6). Hemionine/stilt-legged Equus MTIIIs occupy the same morphospace as H. francisci in our analysis, although given a larger sample size, it may be possible to discriminate $E$. hemionus from the remaining stilt-legged equids. We note that Haringtonhippus seems to exhibit a negative correlation between latitude and MTIII length, and that specimens from the same latitude occupy similar morphospace regardless of whether DNA- or morphological-based identification was used (Figure 2c and Supplementary file 4).

Comments. On the basis of morphology, we assign all referred material of Haringtonhippus to the single species H. francisci Hay 1915 (Appendix 2). Comparison between the cranial anatomical features of LACM(CIT) 109/156450 and TMM 34-2518 reveal some minor differences, which can likely be ascribed to intraspecific variation (Fig 2a and Appendix 2 and Figure 2-Figure supplement 1). Further, the MTIII of TMM 34-2518 is comparable to the MTIIIs ascribed to Haringtonhippus by palaeogenomic data, and is consistent with the observed latitudinally-correlated variation in MTIII length across Haringtonhippus (Figure 2c and Appendix 2).

This action is supported indirectly by molecular evidence, namely the lack of mitochondrial phylogeographic structure and the estimated time to most recent common ancestor (tMRCA) for sampled Haringtonhippus. The mitochondrial tree topology within Haringtonhippus does not exhibit phylogeographic structure (Figure 1-Figure supplement 1b), which is consistent with sampled Haringtonhippus mitochondrial genomes belonging to the same species. Using Bayesian time-tree analysis, we estimated a tMRCA for the sampled Haringtonhippus mitochondrial genomes of $\sim 200-470$ ka BP (Figure 1 and Supplementary table 3 and Supplementary file 7). The MRCA of Haringtonhippus is therefore more recent than that of other extant equid species (such as E. asinus and E. quagga, which have a combined $95 \%$ HPD range: 410-1030 ka BP; Figure 1 and Supplementary table 3 and Supplementary file 7). Although the middle Pleistocene holotype TMM 34-2518 ( $\sim 125-780 \mathrm{ka} \mathrm{BP}$ ) may predate our Haringtonhippus mitochondrial tMRCA, this sample has no direct date and the range of possible ages falls within the tMRCA range of other extant equid species. We therefore cannot reject the hypothesis of its conspecificity with Haringtonhippus, as defined palaeogenomically. We attempted, but were unable, to recover either collagen or genomic data from TMM 34-2518
(Appendix 2), consistent with the taphonomic, stratigraphic, and geographic context of this fossil (Hay, 1915; Lundelius \& Stevens, 1970). Altogether, the molecular evidence is consistent with the assignment of $H$. francisci as the type and only species of Haringtonhippus.

## Discussion

## Reconciling the genomic and fossil records of Plio-Pleistocene equid evolution

The placement of Haringtonhippus as sister taxon to Equus is less congruent with the fossil record than previous hypotheses, which have placed Haringtonhippus within Equus, as sister to either hemionines or caballine horses (Figure 1). The earliest NWSL and caballine equids appear in the fossil record at $\sim 2-3$ and $\sim 1.9-0.7 \mathrm{Ma}$ (Eisenmann et al., 2008; Forsten, 1992), respectively, whereas our divergence estimates suggest that these lineages appeared at 4.1-5.8 and 3.8-4.5 Ma, most likely in North America. This could be due to incompleteness of the fossil record or a discordance between the timing of species divergence and the evolution of diagnostic anatomical characteristics (Forsten, 1992). Although there will likely always be problems associated with the completeness of the fossil record, the well known late Miocene and Pliocene fossil records of New World equids would suggest that this discordance is not due to missing fossils.

Pliocene Equus generally exhibits a primitive ('plesippine' in North America, 'stenonid' in the Old World) morphology that presages living zebras and asses (Forsten, 1988, 1992), with more derived caballine (stout-legged) and hemionine (stilt-legged) forms evolving in the early Pleistocene. The stilt-legged morphology appears to have evolved independently at least once in each of the Old and New Worlds, yielding the Asiatic wild asses and Haringtonhippus, respectively. We include the middle-late Pleistocene Eurasian E. hydruntinus within the Asiatic wild asses (following (Bennett et al., 2017; Burke, Eisenmann, \& Ambler, 2003; Orlando et al., 2006)), and note that the Old World sussemione E. ovodovi may represent another instance of independent stilt-legged origin, but its relation to Asiatic wild asses and other non-caballine Equus is currently unresolved (as depicted in (Der Sarkissian et al., 2015; Orlando et al., 2009;

Vilstrup et al., 2013) and Figure 1). As such, it is plausible that the more primitive Pleistocene morphological states, including Hippidion-like, only survived past the early to middle Pleistocene at lower latitudes (South America, Africa; Figure 1), and that the derived hemionine and caballine morphologies evolved from, and replaced, primitive ancestors in higher latitude North America and Eurasia during the Pleistocene, perhaps as adaptations to the extreme ecological pressures perpetuated by the advance and retreat of continental ice sheets in response to rapid climate oscillations during glacial cycles (Forsten, 1992, 1996). We note that this high latitude replacement model is consistent with the turnover observed in the Pleistocene equid regional fossil records of North America (Azzaroli, 1992; Azzaroli \& Voorhies, 1993) and Eurasia (Forsten, 1988, 1992, 1996). By contrast, in South America Hippidion co-existed with caballine horses until they both succumbed to extinction with much of the New World megafauna near the end of the Pleistocene (Forsten, 1996; Koch \& Barnosky, 2006; O’Dea et al., 2016). This model helps explain the discordance between the timings of the appearance of the caballine and hemionine morphologies in the fossil record and the divergence of lineages leading to these forms as estimated from palaeogenomic data. The model predicts that some previously described North American Pliocene and early Pleistocene primitive Equus species (e.g., E. simplicidens, E. idahoensis; (Azzaroli \& Voorhies, 1993)), or specimens thereof, may be ancestral to extant Equus and/or late Pleistocene Haringtonhippus.

## Temporal and geographic range overlap of Pleistocene equids in North America

 Three new radiocarbon dates of $\sim 14.4{ }^{14} \mathrm{C}$ ka BP from a Yukon Haringtonhippus fossil greatly extends the known temporal range of this genus in eastern Beringia. This result demonstrates, contrary to its previous LAD of $31,400 \pm 1200{ }^{14} \mathrm{C}$ years ago (AA 26780; (Guthrie, 2003)), thatHaringtonhippus survived throughout the last glacial maximum in eastern Beringia (Clark et al., 2009) and may have come into contact with humans near the end of the Pleistocene (Goebel, Waters, \& O'Rourke, 2008; Guthrie, 2006). These data suggest that populations of stilt-legged Haringtonhippus and stout-legged caballine Equus were sympatric, both north and south of the continental ice sheets, through the late Pleistocene and became extinct at roughly the same time. The near synchronous extinction of both horse groups across their entire range in North America suggests that similar causal mechanisms may have led each to their demise.

The sympatric nature of these equids raises questions of whether they managed to live within the same community without hybridizing or competing for resources. Extant members of the genus Equus vary considerably in the sequence of Prdm9, a gene involved in the speciation process, and chromosome number (karyotype) (Ryder, Epel, \& Benirschke, 1978; Steiner \& Ryder, 2013), and extant caballine and non-caballine Equus rarely produce fertile offspring (Allen \& Short, 1997; Steiner \& Ryder, 2013). It is unlikely, therefore, that the more deeply diverged Haringtonhippus and caballine Equus would have been able to hybridize. Future analysis of high coverage nuclear genomes, ideally including an outgroup such as Hippidion, will make it possible to test for admixture that may have occurred soon after the lineages leading to Haringtonhippus and Equus diverged, as occurred between the early caballine and noncaballine Equus lineages (Jónsson et al., 2014). It may also be possible to use isotopic and/or tooth mesowear analyses to assess the potential of resource partitioning between Haringtonhippus and caballine Equus in the New World.

## Fossil systematics in the palaeogenomics and proteomics era: concluding remarks

Fossils of NWSL equids have been known for more than a century, but until the present study their systematic position within Plio-Pleistocene Equidae was poorly characterized. This was not because of a lack of interest on the part of earlier workers, whose detailed anatomical studies strongly indicated that what we now call Haringtonhippus was related to Asiatic wild asses, such as Tibetan khulan and Persian onagers, rather than to caballine horses (Eisenmann et al., 2008; Guthrie, 2003; Scott, 2004; Skinner \& Hibbard, 1972). That the cues of morphology have turned out to be misleading in this case underlines a recurrent problem in systematic biology, which is how best to discriminate authentic relationships within groups, such as Neogene equids, that were prone to rampant convergence. The solution we adopted here was to utilize both palaeogenomic and morphometric information in reframing the position of Haringtonhippus, which now clearly emerges as the closest known outgroup to all living Equus.

Our success in this regard demonstrates that an approach which incorporates phenomics with molecular methods (palaeogenomic as well as palaeoproteomic, e.g. (Welker et al., 2015)) is likely to offer a means for securely detecting relationships within speciose groups that are highly diverse ecomorphologically. All methods have their limits, with taphonomic degradation being the critical one for molecular approaches. However, proteins may persist significantly longer than ancient DNA (e.g. (Rybczynski et al., 2013)), and collagen proteomics may come to play a key role in characterizing affinities, as well as the reality, of several proposed Neogene equine taxa (e.g. Dinohippus, Pliohippus, Protohippus, Calippus, and Astrohippus; (Macfadden, 1998)) whose distinctiveness and relationships are far from settled (Azzaroli \& Voorhies, 1993; Forsten, 1992). A reciprocally informative approach like the one taken here holds much promise for lessening the amount of systematic noise, due to oversplitting, that hampers our
understanding of the evolutionary biology of other major late Pleistocene megafaunal groups such as bison and mammoths (Enk et al., 2016; Froese et al., 2017). This approach is clearly capable of providing new insights into just how extensive megafaunal losses were at the end of the Pleistocene, in what might be justifiably called the opening act of the Sixth Mass Extinction in North America.

## Materials and Methods

We provide an overview of methods here; full details can be found in Appendix 1.

Sample collection and radiocarbon dating. We recovered Yukon fossil material (17
Haringtonhippus francisci, two Equus cf. scotti, and two E. lambei; Supplementary file 1) from active placer mines in the Klondike goldfields near Dawson City. We further sampled seven $H$. francisci fossils from the contiguous USA that are housed in collections at the University of Kansas Biodiversity Institute (KU; n=4), Los Angeles County Museum of Natural History (LACM(CIT); $n=2$ ), and the Texas Memorial Museum (TMM; $n=1$ ). We radiocarbon dated the Klondike fossils and the H. francisci cranium from the LACM(CIT) (Supplementary file 1).

Morphometric analysis of third metatarsals. For morphometric analysis, we took measurements of third metatarsals (MTIII) and other elements. We used a reduced data set of four MTIII variables for principal components analysis and performed logistic regression on the first three principal components, computed in R (R Development Core Team, 2008) (Supplementary file 6).

DNA extraction, library preparation, target enrichment, and sequencing. We conducted all molecular biology methods prior to indexing PCR in the dedicated palaeogenomics laboratory facilities at either the UC Santa Cruz or Pennsylvania State University. We extracted DNA from between 100 and 250 mg of bone powder following either (Rohland, Siedel, \& Hofreiter, 2010) or (Dabney, et al., 2013). We then converted DNA extracts to libraries following the Meyer and Kircher protocol (Meyer \& Kircher, 2010), as modified by (Heintzman et al., 2015) or the PSU
method of (Vilstrup et al., 2013). We enriched libraries for equid mitochondrial DNA. We then sequenced all enriched libraries and unenriched libraries from 17 samples using Illumina platforms.

Mitochondrial genome reconstruction and analysis. We prepared raw sequence data for alignment and mapped the filtered reads to the horse reference mitochondrial genome (Genbank: NC_001640.1) and a H. francisci reference mtDNA genome (Genbank: KT168321), resulting in mitogenomic coverage ranging from $5.8 \times$ to $110.7 \times$ (Supplementary file 1 ). We were unable to recover equid mtDNA from TMM 34-2518 (the francisci holotype) using this approach (Appendix 2). We supplemented our mtDNA genome sequences with 38 previously published complete equid mtDNA genomes. We constructed six alignment data sets and selected models of molecular evolution for the analyses described below (Supplementary table 8, and Supplementary files 1 and 8 ).

We tested the phylogenetic position of the NWSL equids (=H. francisci) using mtDNA data sets 1-3 and applying Bayesian (Ronquist et al., 2012) and Maximum Likelihood (ML; (Stamatakis, 2014)) analyses. We varied the outgroup, the inclusion or exclusion of the fastevolving partitions, and the inclusion or exclusion of Hippidion sequences. Due to the lack of a globally supported topology across the Bayesian and ML phylogenetic analyses, we used an Evolutionary Placement Algorithm (EPA; (Berger et al., 2011)) to determine the a posteriori likelihood of phylogenetic placements for candidate equid outgroups using mtDNA data set four. We also used the same approach to assess the placement of previously published equid sequences (Appendix 2). To infer divergence times between the four major equid groups (Hippidion, NWSL equids, caballine Equus, and non-caballine Equus), we ran Bayesian timetree
analyses (Drummond et al. 2012) using mtDNA data set five. We varied these analyses by including or excluding fast-evolving partitions, constrained the root height or not, and including or excluding the E. ovodovi sequence.

To facilitate future identification of equid mtDNA sequences, we constructed, using data set six, a list of putative synapomorphic base states, including indels and substitutions, that define the genera Hippidion, Haringtonhippus, and Equus at sites across the mtDNA genome.

## Phylogenetic inference, divergence date estimation, and sex determination from nuclear

 genomes. To test whether our mtDNA genome-based phylogenetic hypothesis truly reflects the species tree, we compared the nuclear genomes of a horse (EquCab2), donkey (Orlando et al., 2013), and the shotgun sequence data from 17 of our NWSL equid samples (Appendix 1, Appendix 1-figure 1 and Supplementary files 1-2). We applied four successive approaches, which controlled for reference genome and DNA fragment length biases (Appendix 1).We estimated the divergence between the NWSL equids and Equus (horse and donkey) by fitting the branch length, or relative private transversion frequency, ratio between horse/donkey and NWSL equids into a simple phylogenetic scenario (Figure 1-Figure supplement 3). We then multiplied the NWSL equid branch length by a previous horse-donkey divergence estimate (4.0-4.5 Ma; (Orlando et al., 2013)) to give the estimated NWSL equidEquus divergence date, following (Heintzman et al., 2015) and assuming a strict genome-wide molecular clock (Figure 1-Figure supplement 3).

We determined the sex of the 17 NWSL equid samples by comparing the relative mapping frequency of the autosomes to the X chromosome.

DNA damage analysis. We assessed the prevalence of mitochondrial and nuclear DNA damage in a subset of the equid samples using mapDamage (Jónsson et al., 2013).

Data availability. Repository details and associated metadata for curated samples can be found in Supplementary file 1. MTIII and other element measurement data are in Supplementary file 4, and the Rscript used for morphometric analysis is Supplementary file 6. MtDNA genome sequences have been deposited in Genbank under accessions KT168317-KT168336, MF134655MF134663, and an updated version of JX312727. All mtDNA genome alignments and associated XML and tree files are in Supplementary files 7-9. Raw shotgun sequence data used for the nuclear genomic analyses have been deposited in the Short Read Archive (Bioproject: PRJNA384940; SRA accessions: SRR5556842-SRR5556858).

## Appendix 1 Supplementary methods

### 1.1 Yukon sample context and identification

Pleistocene vertebrate fossils are commonly recovered at these localities, in the absence of stratigraphic context, as miners are removing frozen sediments to access underlying gold bearing gravel (Froese et al., 2009; Harington, 2011). We recovered H. francisci fossils along with other typical late Pleistocene (Rancholabrean) taxa, including caballine horses (Equus sp.), woolly mammoth (Mammuthus primigenius), steppe bison (Bison priscus), and caribou (Rangifer tarandus), which are consistent with our age estimates based on radiocarbon dating (Supplementary file 1). All Yukon fossil material consisted of limb bones that were taxonomically assigned based on their slenderness and are housed in the collections of the Yukon Government (YG).

### 1.2 Radiocarbon dating

We subsampled fossil specimens using handheld, rotating cutting tools and submitted them to either the KECK Accelerator Mass Spectrometry (AMS) Laboratory at the University of California (UC), Irvine (UCIAMS) and/or the Center for AMS (CAMS) at the Lawrence Livermore National Laboratory. We extracted collagen from the fossil subsamples using ultrafiltration (Beaumont, Beverly, Southon, \& Taylor, 2010), which was used for AMS radiocarbon dating. We were unable to recover collagen from TMM 34-2518 (francisci holotype), consistent with the probable middle Pleistocene age of this specimen (Lundelius \& Stevens, 1970). We recovered finite radiocarbon dates from all other fossils, with the exception of the two Equus cf. scotti specimens. We calibrated AMS radiocarbon dates using the IntCal13
curve (Reimer et al., 2013) in OxCal v4.2 (https://c14.arch.ox.ac.uk/oxcal/OxCal.html) and report median calibrated dates in Supplementary file 1.

### 1.3 Morphometric analysis of third metatarsals

Third metatarsal (MTIII) and other elemental measurements were either taken by G.D.Z. or E.S. or from the literature (Supplementary file 4). For morphometric analysis, we focused exclusively on MTIIIs, which exhibit notable differences in slenderness among equid groups (Figure 2Figure supplement 2a; (Weinstock et al., 2005)). Starting with a data set of 10 variables (following (Eisenmann, Alberdi, deGiuli, \& Staesche, 1988)), we compared the loadings of all variables in principal components space in order to remove redundant measurements. This reduced the data set to four variables (GL: greatest length, Pb : proximal breadth, Mb : midshaft breadth, and DABm: distal articular breadth at midline). We visualized the reduced variables using principal components analysis, computed in R (Supplementary file 5; (R Development Core Team, 2008)), and performed logistic regression on the first three principal components to test whether MTIII morphology can distinguish stilt-legged (hemionine Equus and H. francisci, $\mathrm{n}=105$ ) from stout-legged (caballine Equus, n=187) equid specimens.

### 1.4 Target enrichment and sequencing

We enriched libraries for equid mitochondrial DNA following the MyBaits v2 protocol (MY croarray, Ann Arbor, MI), with RNA bait molecules constructed from the horse reference mitochondrial genome sequence (NC_001640.1). We then sequenced the enriched libraries for $2 \times 150$ cycles on the Illumina HiSeq- 2000 platform at UC Berkeley or $2 \times 75$ cycles on the MiSeq platform at UC Santa Cruz, following the manufacturer's instructions. We produced data
for the nuclear genomic analyses by shotgun sequencing 17 of the unenriched libraries for $2 \times 75$ cycles on the MiSeq to produce $\sim 1.1-6.4$ million reads per library (Supplementary file 2).

### 1.5 Mitochondrial genome reconstruction

We initially reconstructed the mitochondrial genome for $H$. francisci specimen YG 404.663 (PH047). For sequence data enriched for the mitochondrial genome, we trimmed adapter sequences, merged paired-end reads (with a minimum overlap of 15 base pairs ( bp ) required), and removed merged reads shorter than 25 bp, using SeqPrep
(https://github.com/jstjohn/SeqPrep). We then mapped the merged and remaining unmerged reads to the horse reference mitochondrial genome sequence using the Burrows-Wheeler Aligner aln (BWA-aln v0.7.5; (Li \& Durbin, 2010)), with ancient parameters (-1 1024; (Schubert et al., 2012)). We removed reads with a mapping quality less than 20 and collapsed duplicated reads to a single sequence using SAMtools v0.1.19 rmdup (Li et al., 2009). We called consensus sequences using Geneious v8.1.7 (Biomatters, http://www.geneious.com; (Kearse et al., 2012)). We then re-mapped the reads to the same reference mitochondrial genome using the iterative assembler, MIA (Briggs et al., 2009). Consensus sequences from both alignment methods required each base position to be covered a minimum of three times, with a minimum base agreement of $67 \%$. The two consensus sequences were then combined to produce a final consensus sequence for YG 404.663 (Genbank: KT168321), which we used as the H. francisci reference mitochondrial genome sequence.

For the remaining newly analyzed 21 H . francisci, two E. cf. scotti, and two E. lambei samples, we merged and removed reads as described above. We then separately mapped the retained reads to the horse and $H$. francisci mitochondrial reference genome sequences using

MIA. Consensus sequences from MIA analyses were called as described above. The two consensus sequences were then combined to produce a final consensus sequence for each sample, with coverage ranging from $5.8 \times$ to $110.7 \times$ (Supplementary file 1 ). We also reconstructed the mitochondrial genomes for four previously published samples: YG 401.268, LACM(CIT) 109/150807, KU 62158, and KU 62055 (Supplementary file 1; (Vilstrup et al., 2013; Weinstock et al., 2005)).

### 1.6 Mitochondrial genome alignments

We supplemented our 30 new mitochondrial genome sequences with 38 previously published complete equid mitochondrial genomes, which included all extant Equus species, and extinct Hippidion, E. ovodovi, and E. cf. scotti ('equids'). We constructed six alignment data sets for the mitochondrial genome analyses: (1) equids and White rhinoceros (Ceratotherium simum; NC_001808) ( $\mathrm{n}=69$ ); (2) equids and Malayan tapir (Tapirus indicus; NC_023838) (n=69); (3) equids, six rhinos, two tapirs, and dog (Canis lupus familiaris; NC_002008) (n=77); (4) equids, six rhinos, two tapirs, ten published equid short fragments, and two published NWSL equid mitochondrial genome sequences $(\mathrm{n}=88)$; (5) a reduced equid data set $(\mathrm{n}=32)$; and (6) a full equid data set ( $\mathrm{n}=68$ ) (Supplementary file 8). For data sets three and four, we selected one representative from all rhino and tapir species for which full mitochondrial genome data are publicly available (Supplementary file 1).

For all six data sets, we first created an alignment using muscle (v3.8.31; (15)). We then manually scrutinized alignments for errors and removed a 253 bp variable number of tandem repeats (VNTR) part of the control region, corresponding to positions 16121-16373 of the horse reference mitochondrial genome. We partitioned the alignments into six partitions (three codon
positions, ribosomal-RNAs, transfer-RNAs, and control region), using the annotated horse reference mitochondrial genome in Geneious, following (Heintzman et al., 2015). We excluded the fast-evolving control region alignment for data set three, which included the highly-diverged dog sequence. For each partition, we selected models of molecular evolution using the Bayesian information criterion in jModelTest (v2.1.6; (Darriba, Taboada, Doallo, \& Posada, 2012)) (Supplementary table 8).

### 1.7 Phylogenetic analysis of mitochondrial genomes

To test the phylogenetic position of the NWSL equids, we conducted Bayesian and maximum likelihood (ML) phylogenetic analyses of data sets one, two, and three, under the partitioning scheme and selected models of molecular evolution described above. For outgroup, we selected: White rhinoceros (data set one), Malayan tapir (data set two), or dog (data set three). For each of the data sets, we varied the analyses based on (a) inclusion or exclusion of the fast-evolving partitions (third codon positions and control region, where appropriate) and (b) inclusion or exclusion of the Hippidion sequences. We ran Bayesian analyses in MrBayes (v3.2.6, (Ronquist et al., 2012)) for two parallel runs of 10 million generations, sampling every 1,000 , with the first $25 \%$ discarded as burn-in. We conducted ML analyses in RAxML (v8.2.4, (Stamatakis, 2014)), using the GTRGAMMAI model across all partitions, and selected the best of three trees. We evaluated branch support with both Bayesian posterior probability scores from MrBayes and 500 ML bootstrap replicates in RAxML.

### 1.8 Placement of outgroups and published sequences a posteriori

We used the Evolutionary Placement Algorithm (EPA) in RAxML to determine the a posteriori likelihood of phylogenetic placements for eight candidate equid outgroups (two tapirs, six rhinos) relative to the four well supported major equid groups (Hippidion, NWSL equids, caballine Equus, non-caballine Equus). We first constructed an unrooted reference tree consisting only of the equids from data set four in RAxML. We then analyzed the placements of the eight outgroups and retaining all placements up to a cumulative likelihood threshold of 0.99. We used the same approach to assess the placement of 12 previously published equid sequences derived from four NWSL equids (Vilstrup et al., 2013; Weinstock et al., 2005), five Hippidion devillei (Orlando et al., 2009), and three E. ovodovi (Orlando et al., 2009) (Supplementary table 4).

### 1.9 Divergence date estimation from mitochondrial genomes

To further investigate the topology of the four major equid groups, and to infer divergence times between them, we ran Bayesian timetree analyses in BEAST (v1.8.4; (Drummond et al., 2012)). Unlike the previous analyses, BEAST can resolve branching order in the absence of an outgroup, by using branch length and molecular clock methods. For BEAST analyses, we used data set five. We did not enforce monophyly. Where available, we used radiocarbon dates to tip date ancient samples. For two samples without available radiocarbon dates, we sampled the ages of tips. For the E. ovodovi sample (mtDNA genome: NC_018783), which was found in a cave that has been stratigraphically dated as late Pleistocene and includes other E. ovodovi remains have been dated to $\sim 45-50 \mathrm{ka}$ BP (Eisenmann \& Sergej, 2011; Orlando et al., 2009), we used the following lognormal prior (mean: $4.5 \times 10^{4}, \log \left(\right.$ stdev ): 0.766 , offset: $1.17 \times 10^{4}$ ) to ensure that $95 \%$ of the prior fell within the late Pleistocene (11.7-130 ka BP). For the E. cf. scotti
mitochondrial genome (KT757763), we used a normal prior (mean: $6.7 \times 10^{5}$, stdev: $5.64 \times 10^{4}$ ) to ensure that $95 \%$ of the prior fell within the proposed age range of this specimen (560-780 ka BP; (Orlando et al., 2013)). We further calibrated the tree using an age of 4-4.5 Ma for the root of crown group Equus (normal prior, mean $4.25 \times 10^{6}$, stdev: $1.5 \times 10^{5}$ ) (Orlando et al., 2013). To assess the impact of variables on the topology and divergence times, we either (a) included or excluded the fast-evolving partitions, (b) constrained the root height (lognormal prior: mean $1 \times 10^{7}$, stdev: 1.0) or not, and (c) included or excluded the E. ovodovi sequence, which was not directly dated. We used the models of molecular evolution estimated by jModeltest (Supplementary table 8). We estimated the substitution and clock parameters for each partition, and estimated a single tree using all partitions. We implemented the birth-death serially sampled (BDSS) tree prior. We ran two analyses for each variable combination. In each analysis, we ran the MCMC chain for 100 million generations, sampling trees and parameters every 10,000 , and discarding the first $10 \%$ as burn-in. We checked log files for convergence in Tracer (v1.6; http://tree.bio.ed.ac.uk/software/tracer/). We combined trees from the two runs for each variable combination in LogCombiner (v1.8.4) and then calculated the maximum clade credibility (MCC) tree in TreeAnnotator (v1.8.4). We report divergence dates as $95 \%$ highest posterior probability credibility intervals of node heights.

### 1.10 Mitochondrial synapomorphy analysis

We first divided data set six, which consists of all available and complete equid mitogenomic sequences, into three data sets based on the genera Hippidion, Haringtonhippus, and Equus. For each of the three genus-specific alignments, we created a strict consensus sequence, whereby sites were only called if there was $100 \%$ sequence agreement, whilst including gaps and
excluding ambiguous sites. We then compared the three genus-specific consensus sequences to determine sites where one genus exhibited a base state that is different to the other two genera, or, at five sites, where each genus has its own base state (Supplementary file 5). In this analysis, we did not make any inference regarding the ancestral state for the identified synapomorphic base states. We identified 391 putative mtDNA genome synapomorphies for Hippidion, 178 for Haringtonhippus, and 75 for Equus (Supplementary file 5).

### 1.11 Phylogenetic inference from nuclear genomes

We compared the genomes of a horse (E. caballus; EquCab2; GCA_0000002305.1) and donkey (E. asinus; Willy, $12.4 \times$; http://geogenetics.ku.dk/publications/middle-pleistocene-omics; (Orlando et al., 2013)) with shotgun sequence data from 17 of our NWSL equid samples (Supplementary files 1-2). We merged paired-end reads using SeqPrep as described above, except that we removed merged reads shorter than 30 bp . We further removed merged and remaining unmerged reads that had low sequence complexity, defined as a DUST score $>7$, using PRINSEQ-lite v0.20.4 (Schmieder \& Edwards, 2011). We used four successive approaches to minimize the impact of mapping bias introduced from ancient DNA fragment length variation and reference genome choice.

We first followed a modified version of the approach outlined in (Heintzman et al., 2015). We mapped the donkey genome to the horse genome by computationally dividing the donkey genome into 150 bp 'pseudo-reads' tiled every 75 bp , and aligned these pseudo-reads using Bowtie2-local v2.1.0 (Langmead \& Salzberg, 2012) while allowing one seed mismatch and a maximum mismatch penalty of four to better account for ancient DNA specific damage (Appendix 1-figure 1, steps 1-3). We then mapped the filtered shotgun data from each of the

NWSL equid samples to the horse genome using Bowtie2-local with the settings described above, and removed PCR duplicated reads and those with a mapping quality score of $<30$ in SAMtools. We called a pseudo-haploidized sequence for the donkey and NWSL equid alignments, by randomly picking a base with a base quality score $\geq 60$ at each position, using SAMtools mpileup. We masked positions that had a coverage not equal to $2 \times$ (donkey) or $1 \times$ (NWSL equid), and those located on scaffolds shorter than 100 kb (Appendix 1-figure 1, step 4). As the horse, donkey, and NWSL equid genome sequences were all based on the horse genome coordinates, we compared the relative transversion frequency between the donkey or NWSL equids and the horse using custom scripts. We restricted our analyses to transversions to avoid the impacts of ancient DNA damage, which can manifest as erroneous transitions from the deamination of cytosine (e.g. Appendix 2-Figures 1-2) (Dabney, Meyer, \& Pääbo, 2013). We repeated this analysis, but with the horse and NWSL equids mapped to the donkey genome (the donkey genome coordinate framework).


Appendix 1-Figure 1. An overview of the nuclear genome analysis pipeline. A first reference genome sequence (red; step 1) is divided into 150 bp pseudo-reads, tiled every 75 bp for exactly $2 \times$ genomic coverage (step 2 ). These pseudo-reads are then mapped to a second reference genome (blue; step 3), and a consensus sequence of the mapped pseudo-reads is called (step 4). Regions of the second reference genome that are not covered by the pseudo-reads are masked (step 5). For each NWSL equid sample, reads (orange) are mapped independently to the first reference consensus sequence (step 6a) and masked second reference genome (step 6b). Alignments from steps 6 a and 6 b are then merged (step 7). For alignment coordinates that have base calls for the first reference, second reference, and NWSL equid sample genomes, the relative frequencies of private transversion substitutions (yellow stars) for each genome are
calculated (step 8). The co-ordinates from the second reference genome (blue) are used for each analysis.

For the second approach and using the horse genome coordinate framework, we next masked sites in the horse reference genome that were not covered by donkey reads at a depth of $2 \times$. This resulted in the horse genome and donkey consensus sequence being masked at the same positions (Appendix 1-figure 1, step 5). We then separately mapped the filtered NWSL equid shotgun data to scaffolds longer than 100 kb for the masked horse genome and donkey consensus sequence (Appendix 1-figure 1, step 6), called NWSL equid consensus sequences, and calculated relative transversion frequencies as described above. This analysis was repeated using the donkey genome coordinate framework.

Next, for each genome coordinate framework, we combined the two alignments for each NWSL equid sample from approach two to create a union of reads mappable to both the masked coordinate genome and alternate genome consensus sequence (Appendix 1-figure 1, step 7). If a NWSL equid read mapped to different coordinates between the two references, we selected the alignment with the higher map quality score and randomly selected between mappings of equal quality. We then called NWSL equid consensus sequences as above. As this third approach allowed for simultaneous comparison of the horse, donkey, and NWSL equid sequences, we calculated relative private transversion frequencies for each sequence, at sites where all three sequences had a base call, using tri-aln-report (https://github.com/Paleogenomics/ChromCompare) (Appendix 1-figure 1, step 8).

Finally, as a fourth approach and for both genome coordinate frameworks, we repeated approach three with the exception that we divided the NWSL alignments by mapped read length.

We split the alignments into 10 bp read bins ranging from 30-39 to $120-129 \mathrm{bp}$, and discarded longer reads and paired-end reads that were unmerged by SeqPrep. We called consensus sequences and calculated relative private transversion frequencies for each sequence as described above. We only used relative private transversion frequencies from the $90-99$ to $120-129 \mathrm{bp}$ bins for divergence date estimates (Appendix 2).

### 1.12 Sex determination from nuclear genomes

We used the alignments of the 17 NWSL equids to the horse genome, from approach one described above, to infer the probable sex of these individuals. For this, we determined the number of reads mapped to each chromosome using SAMtools idxstats. For each chromosome, we then calculated the relative mapping frequency by dividing the number of mapped reads by the length of the chromosome. We then compared the relative mapping frequency between the autosomes and X-chromosome. As males and females are expected to have one and two copies of the X chromosome, respectively, and two copies of every autosome, we inferred a male if the ratio between the autosomes and X-chromosome was $0.45-0.55$ and a female if the ratio were 0.9-1.1.

### 1.13 DNA damage analysis

For a subset of nine samples, we realigned the filtered sequence data from the libraries enriched for equid mitochondrial DNA to either the $H$. francisci (for $H$. francisci samples) or horse (for $E$. lambei and E. cf. scotti samples) reference mitochondrial genome sequences using BWA-aln as described above. We also realigned the filtered unenriched sequence data to the horse reference
genome (EquCab2) for a subset of six samples using the same approach. We then analyzed patterns of DNA damage in mapDamage v2.0.5 (Jónsson et al., 2013).

## Appendix 2 Supplementary Results

### 2.1 Ancient DNA characterization

We selected a subset of samples for the analysis of DNA damage patterns. In all of these samples, we observe expected patterns of damage in both mitochondrial and nuclear DNA, including evidence of the deamination of cytosine residues at the ends of reads, depurinationinduced strand breaks, and a short mean DNA fragment length (Dabney, Meyer, \& Pääbo, 2013) (Appendix 2-Figures 1-2). We note that the sample with the greatest proportion of deaminated cytosines is E. cf. scotti (YG 198.1; Appendix 2-Figure 1v-x), which is the oldest sample in the subset (Supplementary file 1).


Appendix 2-Figure 1. Characterization of ancient mitochondrial DNA damage patterns from nine equid samples. H. francisci: (A-C) JK166 (LACM(CIT) 109/150807; Nevada), (D-F) JK207 (LACM(CIT) 109/156450; Nevada), (G-I) JK260 (KU 47800; Wyoming), (J-L) PH013 (YG 130.6; Yukon), (M-O) PH047 (YG 404.663; Yukon), (P-R) MS272 (YG 401.268; Yukon), (S-U)

MS349 (YG 130.55; Yukon); E. cf. scotti: (V-X) PH055 (YG 198.1; Yukon); E. lambei: (Y-AA) MS316 (YG 328.54; Yukon). Every third panel: (A) to (Y) DNA fragment length distributions; $(\mathrm{B})$ to $(\mathrm{Z})$ proportion of cytosines that are deaminated at fragment ends (red: cytosine $\rightarrow$ thymine; blue: guanine $\rightarrow$ adenine); and (C) to (AA) mean base frequencies immediately upstream and downstream of the 5' and 3' ends of mapped reads.


Appendix 2-Figure 2. Characterization of ancient nuclear DNA damage patterns from six $H$. francisci samples. (A-C) JK166 (LACM(CIT) 109/150807; Nevada), (D-F) JK260 (KU 47800; Wyoming), (G-I) PH013 (YG 130.6; Yukon), (J-L) PH036 (YG 76.2; Yukon), (M-O) MS349 (YG 130.55; Yukon), (P-R) MS439 (YG 401.387; Yukon). Every third panel: (A) to (P) DNA
fragment length distributions; (B) to $(\mathrm{Q})$ proportion of cytosines that are deaminated at fragment ends (red: cytosine $\rightarrow$ thymine; blue: guanine $\rightarrow$ adenine); and $(\mathrm{C})$ to $(\mathrm{R})$ mean base frequencies immediately upstream and downstream of the $5^{\prime}$ and 3 ' ends of mapped reads.

### 2.2 Resolving the phylogenetic placement of NWSL equids using mitochondrial genomes

We ran Bayesian and ML phylogenetic analyses on mtDNA genome alignment data sets 1-3, whilst varying the outgroup, including (all) or excluding (reduced) the fast-evolving partitions (see Appendix 1), and including or excluding the Hippidion sequences. In all analyses, we recover four major equid groups (Hippidion, NWSL equids(=H. francisci), caballine Equus, and non-caballine Equus) with strong statistical support (Bayesian posterior probability (BPP):
1.000; ML bootstrap: 96-100\%; Supplementary table 1), consistent with previous studies (e.g. (Der Sarkissian et al., 2015; Orlando et al., 2009)). We recover conflicting phylogenetic topologies between these four groups, however, which is dependent on the variables described above and the choice of phylogenetic algorithm (Appendix 2-Figure 3; Supplementary table 1). Across all analyses, strong statistical support (BPP: $\geq 0.99$; ML bootstrap: $\geq 95 \%$ ) is only associated with topology 1 (Appendix 2-Figure 3; Supplementary table 1), in which NWSL equids are placed outside of Equus, and Hippidion is placed outside of the NWSL equid-Equus clade. We note that the analyses with the strongest support consist of multiple outgroups (mtDNA data set three).

A

| 1 | Hippidion | NWSL | Non-caballines | Caballines |
| :---: | :---: | :---: | :---: | :---: |
| 2 | NWSL | Hippidion | Non-caballines | Caballines |
| 4 | Hippidion | Caballines | Non-caballines | NWSL |
| 5 | Hippidion | Non-caballines | Caballines | NWSL |
| 6 | Caballines | Non-caballines | Hippidion | NWSL |
| 7 | Non-caballines | Caballines | Hippidion | NWSL |



B


C

| 1 | NWSL | Non-caballines | Caballines |
| :---: | :---: | :---: | :---: |
| 2 | NWSL | Non-caballines | Caballines |
| 4 | Caballines | Non-caballines | NWSL |
| 5 | Non-caballines | Caballines | NWSL |
| 6 | Caballines | Non-caballines | NWSL |
| 7 | Non-caballines | Caballines | NWSL |

D


Appendix 2-Figure 3. Seven phylogenetic hypotheses for the four major groups of equids with sequenced mitochondrial genomes: Hippidion, the New World stilt-legged equids (=Haringtonhippus), non-caballine Equus (asses, zebras, and E. ovodovi) and caballine Equus (horses). (A) imbalanced and (B) balanced hypotheses. The hypotheses presented in (C) and (D) are identical to $(\mathrm{A})$ and $(\mathrm{B})$, except that Hippidion is excluded. Node letters are referenced in Supplementary tables 1-2. We only list combinations that were recovered by our palaeogenomic, or previous palaeogenetic, analyses

We further investigated the effect of outgroup choice by using an evolutionary placement algorithm (EPA; (Berger et al., 2011)) to place the outgroup sequences into an unrooted ML phylogeny a posteriori using the same set of variables described above. We find that the
outgroup placement likelihood is increased with the inclusion of Hippidion sequences, and that the only placements with a likelihood of $\geq 0.95$ are consistent with topology one (Appendix 2-

Figure 3; Supplementary table 2), in agreement with the Bayesian and ML phylogenetic analyses. The phylogenetic and EPA analyses demonstrate that outgroup choice can greatly impact equid phylogenetic inference and that multiple outgroups should be used for resolving relationships between major equid groups.

We lastly ran Bayesian timetree analyses in BEAST in the absence of an outgroup, whilst including or excluding the fast-evolving partitions, including or excluding the $E$. ovodovi sequence, and constraining the root prior or not. All BEAST analyses yielded a maximum clade credibility tree that is consistent with topology one (Figures 1 and S11) with BPP support for the NWSL equid-Equus and Equus clades of 0.996-1.000 (Supplementary table 3). Altogether, the phylogenetic, EPA, and timetree analyses support topology one (Appendix 2-Figure 3), with NWSL equids falling outside of Equus, and therefore the NWSL equids as a separate genus,

## Haringtonhippus.

### 2.3 Placement of previously published NWSL equid sequences

To confirm that all six previously published NWSL equid samples with available mtDNA sequence data (Vilstrup et al., 2013; Weinstock et al., 2005) belong to $H$. francisci, we either reconstructed mitochondrial genomes for these samples (JW277, JW161; (Weinstock et al., 2005)), placed the sequences into a ML phylogeny a posteriori using the EPA whilst varying the partitioning scheme and inclusion or exclusion of Hippidion (Supplementary table 4), or both. For JW277 and JW161, the mitochondrial genomes were consistent with those derived from the newly analyzed samples (Figure 1-Figure supplement 1). For three other NWSL equid
mitochondrial sequences (JW125, JW126, and JW328; (Vilstrup et al., 2013; Weinstock et al., 2005)), including one sample from Mineral Hill cave (Supplementary file 1), the EPA strongly supported a ML placement within the H. francisci clade (cumulative likelihood of 0.995-1.000). However, the EPA placed the remaining NWSL equid sample (MS272; (Vilstrup et al., 2013)) on the branch leading to H. francisci with strong support (likelihood: 1.000; Supplementary table 4). We therefore explored whether the published sequence for MS272 is problematic.

We first tested the EPA on eight other equid mitochondrial sequences (E. ovodovi, $\mathrm{n}=3$; Hippidion devillei, $\mathrm{n}=5$ ), which grouped as expected from previous analyses (likelihood: 0.9991.000; Supplementary table 4 ; (Orlando et al., 2009)). We then used our mitochondrial genome assembly pipeline to reconstruct a consensus for MS272 from the raw data used by (Vilstrup et al., 2013), which resulted in a different sequence that was consistent with other NWSL equids. To confirm this new sequence, we used the original MS272 DNA extract for library preparation, target enrichment, and sequencing. The consensus from this analysis was identical to our new sequence.

We sought to understand the origins of the problems associated with the published MS272 sequence. We first applied our synapomorphy analysis. For the called bases, we found that the published MS272 sequence contained 0/384 diagnostic bases for Hippidion, 124/164 for Haringtonhippus, and 16/70 for Equus (Supplementary file 5). We infer from this analysis that the published MS272 sequence is therefore $\sim 76 \%$ Haringtonhippus and that $\sim 23 \%$ originates from Equus. The presence of Equus synapomorphies could be explained by the fact that the enriched library for MS272 was sequenced on the same run as ancient caballine horses (Equus), thereby potentially introducing contaminating reads from barcode bleeding (Kircher, Sawyer, \& Meyer, 2012), which may have been exacerbated by alignment to the modern horse reference
mitochondrial genome with BWA-aln and consensus calling using SAMtools (Vilstrup et al., 2013). The presence of caballine horse sequence in the published MS272 mtDNA genome explains why previous phylogenetic analyses of mitochondrial genomes have recovered NWSL equids as sister to caballine Equus with strong statistical support (Der Sarkissian et al., 2015; Vilstrup et al., 2013).

### 2.4 Resolving the phylogenetic placement of NWSL equids using nuclear genomes

The horse and donkey genomes are representative of total Equus genomic diversity (Jónsson et al., 2014), and so, if NWSL equids are Equus, we should expect their genomes to be more similar to either horse or donkey than to the alternative.

Initial analyses based on approach one (see Appendix 1) were inconclusive, with some NWSL equid samples appearing to fall outside of Equus (higher relative transversion frequency between the NWSL equid and the horse or donkey than between the horse and donkey) and others inconsistently placed in the phylogeny, appearing most closely related to horse when aligned to the horse genome and most closely related donkey when aligned to the donkey genome (Supplementary file 2). We then used approaches two and three in an attempt to standardize between the horse and donkey reference genomes, and therefore reduce potential bias introduced from the reference genome. In the latter union-based approach, mapping should not be disproportionately sensitive to regions of the genome where NWSL equids are more horse- or donkey-like. These approaches, however, were not successful, but we noted that relative (private) transversion frequency for the coordinate genome and NWSL equid sequences correlated with mean DNA fragment length (Appendix 2-Figure 4 and Supplementary file 2). We therefore used approach four to control for the large variation in mean DNA fragment length
between NWSL equid sequences (Appendix 2-Figure 2 and Supplementary file 2), which is likely due to a combination of DNA preservation and differences in the DNA extraction and library preparation techniques used (Supplementary file 2). This allowed for direct comparison between the NWSL equid samples, which showed a consistent pattern across read length bins (Figure 1-Figure supplement 2, Supplementary table 3). The relative private transversion frequency for both the coordinate genome and NWSL equid sequences increase with read length until the 90-99 bp bin, at which point the coordinate genome and alternate sequence relative private transversion frequencies converge (defined as a ratio between 0.95-1.05) and the NWSL equid relative private transversion frequencies reach plateau at between $1.40-1.56 \times$ greater than that of the horse or donkey (Figure 1-Figure supplements 2-3, Supplementary table 3).

A greater relative private transversion frequency in NWSL equids, as compared to horse and donkey, is consistent with their being more diverged than the horse-donkey split (Equus) and therefore supports the hypothesis of NWSL equids as a separate genus (Haringtonhippus).


Appendix 2-Figure 4. A comparison of relative private transversion frequencies between the nuclear genomes of a caballine Equus (horse, E. caballus; green), a non-caballine Equus (donkey, E. asinus; red), and the 17 New World-stilt legged (NWSL) equid samples (=Haringtonhippus francisci; blue), using approach three (Appendix 1), with samples ordered by increasing mean mapped read length. Analyses are based on alignment to the horse (A) or donkey (B) genome coordinates.

### 2.5 Sex determination from nuclear genomes

We inferred the sex of our 17 NWSL equid samples by calculating the ratio of relative mapping frequencies between the autosomes and X-chromosome (Supplementary file 3). We find that five of our samples are male and at least eight are female (Supplementary table 6). We note that all three Gypsum Cave samples are inferred to be female, have statistically indistinguishable radiocarbon dates, and identical mtDNA genome sequences (Figure 1-Figure supplement 1b, Supplementary file 1). We therefore cannot reject the possibility that these samples derive from the same individual. Intriguingly, we further note that, across all 17 NWSL equid samples, the relative mapping frequency for chromosomes 8 and 13 is appreciably greater than the remaining autosomes (Supplementary file 3). This may suggest that duplicated regions of these chromosomes are present in NWSL equids, as compared to the horse (E. caballus).

### 2.6 Designation of a type species for Haringtonhippus

We sought to designate a type species for the NWSL equid genus, Haringtonhippus, using an existing name, in order to avoid adding to the unnecessarily extensive list of Pleistocene North American equid species names (Winans, 1985). For this, we scrutinized eight names that have
previously been assigned to NWSL equids in order of priority (date the name was first described in the literature). We rejected names that were solely based on dentitions, as these anatomical features are insufficient for delineating between equid groups (Groves \& Willoughby, 1981). The earliest named species with a valid, diagnostic holotype is francisci Hay 1915. On the basis of taxonomic priority, stratigraphic age, and cranial and metatarsal comparisons (see main results and below), we conclude that francisci Hay 1915 is the most appropriate name for

Haringtonhippus. We also note that this middle Pleistocene species is also small, like our late Pleistocene specimens.

The eight examined names were:
tau Owen, 1869: a small species erected based upon an upper cheek tooth series lacking the $\mathrm{P}^{2}$ from the Valley of Mexico. Other than small size, the species has no reliably diagnostic features. The holotype specimen has been lost, and no topotypal material is available, and so determining whether or not the species represents a NWSL equid is impossible. Eisenmann and colleagues (2008) proposed a neotype specimen for the species, consisting of a cranium (FC 673), but this is rejected here on technical grounds: 1) the proposed neotype fossil was listed as being part of a private collection, which negates its use as a neotype; 2) ICZN rules require that a neotype be "consistent with what is known of the former name-bearing type from the original description and from other sources" and derive from "as nearly as practicable from the original type locality ... and, where relevant, from the same geological horizon or host species as the original namebearing type".
semiplicatus Cope, 1893: based upon an isolated upper molar tooth from Rock Creek, Texas. The specimen has been interpreted to be derived from the same species as the holotype
metatarsal of ' $E$ '. calobatus Troxell (see below) (Azzaroli, 1995; Quinn, 1957; Sandom, Faurby, Sandel, \& Svenning, 2014).
littoralis Hay, 1913: based upon an upper cheek tooth from Peace Creek, Florida. The tooth is small, but offers no diagnostic features.
francisci Hay, 1915: Named in April of 1915 based upon a partial skeleton, including the skull, mandible, and a broken MTIII (TMM 34-2518). Confidently determined to be a NWSL equid based upon reconstruction of the right MTIII by Lundelius and Stevens (1970).
calobatus Troxell, 1915: Named in June of 1915 based upon limb bones. No holotype designated, but lectotype erected by Hibbard (1953) (YPM 13470, right MTIII). altidens Quinn, 1957: based upon a partial skeleton from Blanco Creek, Texas that exhibits elongate metapodials. Synonymized with francisci Hay by Winans (1985).
zoyatalis Mooser, 1958: based upon a partial mandible including the symphyseal region and the right dentary with p2-m3. Synonymized with francisci Hay by Winans (1985).
quinni Slaughter et al. 1962: based upon a MTIII (SMP 60578) and other referred elements from Texas. Synonymized with francisci Hay by Lundelius and Stevens (1970) and Winans (1985).

### 2.7 Anatomical comparison of the francisci holotype and Gypsum Cave crania

We compared the holotype of francisci Hay (TMM 34-2518) from Texas to the Gypsum Cave cranium (LACM(CIT) 109/156450) from Nevada, the latter of which was assigned to Haringtonhippus using palaeogenomic data (Figure 2-Figure supplement 1). Although there are minor anatomical differences between the two crania, which are outlined below, we consider these to fall within the range of intraspecific variation.

The skull from Gypsum Cave (GCS) can be distinguished from that of the francisci holotype (fHS) by its slightly larger size, and markedly longer and more slender rostrum, both absolutely and as a percentage of the skull length. The rostrum of the GCS is also absolutely narrower; the $f \mathrm{HS}$, despite being the smaller skull, is transversely broader at the $\mathrm{i} / 3$. The palatine foramina are positioned medial to the middle of the $\mathrm{M}^{2}$ in the GCS, whereas they are medial to the $\mathrm{M}^{2}-\mathrm{M}^{3}$ junction in the $f \mathrm{HS}$. Viewed laterally, the orbits of the GCS have more pronounced supraorbital ridges than those of the $f \mathrm{HS}$. The latter skull also exhibits somewhat stronger basicranial flexion than the GCS. Dentally, the GCS exhibits arcuate protocones, with strong anterior heels and marked lingual troughs in $\mathrm{P}^{3}-\mathrm{M}^{3}$; the $f \mathrm{HS}$ has smaller, triangular protocones with less pronounced anterior heels and no lingual trough or groove. These characters are not thought to result from different ontogenetic stages, since both specimens appear to be of young adults (all teeth in wear and tall in the jaw). Both the GCS and the $f \mathrm{HS}$ have relatively simple enamel patterns on the cheek teeth, with few evident plications. Not only are the observed differences between these two specimens unlikely to result from ontogeny, they also don't result from sex, since both skulls appear to be females given the absence of canine teeth. The inference of the GCS being female is further supported by palaeogenomic data (Supplementary table 6).

### 2.8 Attempt to recover DNA from the francisci holotype

We attempted to retrieve endogenous mitochondrial and nuclear DNA from the holotype of francisci Hay (TMM 34-2518), to directly link this anatomically-derived species name with our palaeogenomically-derived genus name Haringtonhippus, but were unsuccessful.

After sequencing a library enriched for equid mitochondrial DNA (see Appendix 1), we could only align 11 reads to the horse reference mitochondrial genome sequence with BWA.

Using the basic local alignment search tool (BLASTn), we show that these reads are $100 \%$ match to human and therefore likely originate from contamination. We repeated this approach using MIA and aligned 166 reads, which were concentrated in 20 regions of the mitochondrial genome. We identified these sequences as human ( $n=18,96-100 \%$ identity $)$, cow $(n=1,100 \%)$, or Aves $(\mathrm{n}=1,100 \%)$, consistent with the absence of endogenous mitochondrial DNA in this sample.

We further generated $\sim 800,000$ reads from the unenriched library for TMM 34-2518, and followed a modified metagenomic approach, outlined in (Graham et al., 2016), to assess if any endogenous DNA was present. We mapped the reads to the horse reference genome (EquCab2), using the BWA-aln settings of (Graham et al., 2016), of which 538 reads aligned. We then compared these aligned reads to the BLASTn database. None of the reads uniquely hit Equidae or had a higher score to Equidae than non-Equidae, whereas 492 of the reads either uniquely hit non-Equidae or had a higher score to non-Equidae than Equidae. These results are consistent with either a complete lack, or an ultra-low occurrence, of endogenous DNA in TMM 34-2518.

### 2.9 Morphometric analysis of third metatarsals

Stilt- and stout-legged equids can be distinguished with high accuracy ( $98.2 \%$; logistic regression) on the basis of third metatarsal (MTIII) morphology (Figure 2c, and Supplementary files 3 and 5), which has the potential to easily and confidently distinguish candidates from either group prior to more costly genetic testing. We note that future genetic analysis of ambiguous specimens, that cross the 'middle ground' between stilt- and stout-legged regions of morphospace, could open the possibility of a simple length-vs-width definition for these two morphotypes. Furthermore, we can highlight potential misidentifications, such as the two putative E. lambei specimens that fall within stilt-legged morphospace (Figure 2c), which could
then be tested by genetic analysis. Intriguingly, an Old World E. ovodovi (stilt-legged; MT no. 6; (Eisenmann \& Sergej, 2011))) and New World E. scotti (stout-legged; CMN 29867) specimen directly overlap in a stout-legged region of morphospace (Figure 2c), which could indicate that either this $E$. ovodovi specimen was misidentified or that this species straddles the delineation between stilt- and stout-legged morphologies.
H. francisci occupies a region of morphospace distinct from caballine/stout-legged Equus, but overlaps considerably with hemionine/stilt-legged Equus (Figure 2c). The holotype of H. francisci (TMM 34-2518) is very pronounced in its slenderness; it has a greater MTIII length than most other H. francisci but slightly smaller width/breadth measurements. This holotype is surpassed in these dimensions only by the quinni Slaughter et al. holotype, which has itself previously been synonymized with francisci Hay (Lundelius \& Stevens, 1970; Winans, 1985). This suggests a potentially larger range of MTIII morphology for H. francisci than exhibited by the presently assigned specimens. We observe that this diversity may be influenced by geography, with $H$. francisci specimens from high-latitude Beringia having shorter MTIIIs relative to those from the lower-latitude contiguous USA.

We note that two New World caballine Equus from Yukon, E. cf. scotti and E. lambei, appear to separate in morphospace (Figure 2c), primarily by MTIII length, supporting the potential delineation of these two taxa using MTIII morphology alone.

## Acknowledgements

We thank the Klondike placer gold mining community of Yukon for their support and providing access to their mines from which many of our Haringtonhippus fossils were recovered. We thank Matt Brown and Chris Sagebiel of the Texas Vertebrate Palaeontology Collections at the University of Texas, Austin for access to a portion of TMM 34-2518, and also thank Sam McLeod, Vanessa Rhue, and Aimee Montenegro at the Los Angeles County Museum for access to the Gypsum Cave material for consumptive sampling. Thanks to Brent Breithaupt (Bureau of Land Management) for permitting the sampling of fossils from Natural Trap Cave that were originally recovered by Larry Martin, Miles Gilbert, and colleagues, and are presently curated by the University of Kansas Biodiversity Institute. We thank Chris Beard and David Burnham (University of Kansas) for facilitating access to these fossils. Thanks to Tom Guilderson, Andrew Fields, Dan Chang, and Samuel Vohr for technical assistance. Thanks to Greger Larson for providing the base map in Figure 1. We thank two anonymous reviewers whose comments improved this manuscript. This work used the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley, supported by NIH S10 Instrumentation Grants S10RR029668 and S10RR027303. PDH, JAC, MS, and BS were supported by NSF grants PLR-1417036 and 09090456. LO was supported by the Danish Council for Independent Research Natural Sciences (Grant 4002-00152B); the Danish National Research Foundation (Grant); the "Chaires d'Attractivit. 2014" IDEX, University of Toulouse, France (OURASI), and the European Research Council (ERC-CoG-2015-681605).

## Competing interests

The authors declare no competing interests.

## References

Allen, W. R., \& Short, R. V. (1997). Interspecific and extraspecific pregnancies in equids: anything goes. The Journal of Heredity, 88(5), 384-392.

Azzaroli, A. (1992). Ascent and decline of monodactyl equids: a case for prehistoric overkill. Annales Zoologici Fennici, 28(3-4), 151-163.

Azzaroli, A. (1995). A synopsis of the Quaternary species of Equus in North America. Bolletino Della Societa Palaeontologica Italiana, 34(2), 205-221.

Azzaroli, A., \& Voorhies, M. R. (1993). The Genus Equus in North America. The Blancan species. Palaeontographia Italica, 80, 175-198.

Beaumont, W., Beverly, R., Southon, J., \& Taylor, R. E. (2010). Bone preparation at the KCCAMS laboratory. Nuclear Instruments \& Methods in Physics Research. Section B, Beam Interactions with Materials and Atoms, 268(7-8), 906-909.

Bennett, E. A., Champlot, S., Peters, J., Arbuckle, B. S., Guimaraes, S., Pruvost, M., ... Geigl, E.-M. (2017). Taming the Late Quaternary phylogeography of the Eurasiatic wild ass through ancient and modern DNA. PLoS One, 12(4), e0174216.

Berger, S. A., Krompass, D., \& Stamatakis, A. (2011). Performance, accuracy, and Web server for evolutionary placement of short sequence reads under maximum likelihood. Systematic Biology, 60(3), 291-302.

Briggs, A. W., Good, J. M., Green, R. E., Krause, J., Maricic, T., Stenzel, U., ... Pääbo, S. (2009). Targeted retrieval and analysis of five Neandertal mtDNA genomes. Science, 325(5938), 318-321.

Burke, A., Eisenmann, V., \& Ambler, G. K. (2003). The systematic position of Equus hydruntinus, an extinct species of Pleistocene equid $\uparrow$. Quaternary Research, 59(3), 459469.

Clark, P. U., Dyke, A. S., Shakun, J. D., Carlson, A. E., Clark, J., Wohlfarth, B., ... McCabe, A. M. (2009). The Last Glacial Maximum. Science, 325(5941), 710-714.

Dabney, J., Knapp, M., Glocke, I., Gansauge, M.-T., Weihmann, A., Nickel, B., ... Meyer, M. (2013). Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. Proceedings of the National Academy of Sciences of the United States of America, 110(39), 15758-15763.

Dabney, J., Meyer, M., \& Pääbo, S. (2013). Ancient DNA damage. Cold Spring Harbor Perspectives in Biology, 5(7). https://doi.org/10.1101/cshperspect.a012567

Darriba, D., Taboada, G. L., Doallo, R., \& Posada, D. (2012). jModelTest 2: more models, new heuristics and parallel computing. Nature Methods, 9(8), 772.

Der Sarkissian, C., Vilstrup, J. T., Schubert, M., Seguin-Orlando, A., Eme, D., Weinstock, J., ... Orlando, L. (2015). Mitochondrial genomes reveal the extinct Hippidion as an outgroup to all living equids. Biology Letters, 11(3), 20141058-20141058.

Drummond, A. J., Suchard, M. A., Xie, D., \& Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. Molecular Biology and Evolution, 29(8), 1969-1973.

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research, 32(5), 1792-1797.

Eisenmann, V. (1985). Indications paléoécologiques fournies par les Equus (Mammalia, Perissodactyla) plio-pléistocènes d'Afrique. In L'Environnement des Hominidés au PlioPléistocène (pp. 57-79).

Eisenmann, V. (1992). Origins, dispersals, and migrations of Equus (Mammalia, Perissodactyla). CFS Courier Forschungsinstitut Senckenberg, 153, 161-170.

Eisenmann, V. (2003). Gigantic Horses. In A. Petculescu \& E. Stiuca (Eds.), Advances in Palaeontology "Hen to Panta", Papers in Honour of C. Radulescu and P. M. Samson (pp. 31-40). Bucharest.

Eisenmann, V., Alberdi, M. T., deGiuli, C., \& Staesche, U. (1988). Volume 1: methodology. In M. Woodburne \& P. Sondaar (Eds.), Studying Fossil Horses (pp. 1-71). E. J. Brill, Leiden. Eisenmann, V., Howe, J., \& Pichardo, M. (2008). Old world hemiones and new world slender species (Mammalia, Equidae). Palaeovertebrata, 36(1-4), 159-233.

Eisenmann, V., \& Sergej, V. (2011). Unexpected finding of a new Equus species (Mammalia, Perissodactyla) belonging to a supposedly extinct subgenus in late Pleistocene deposits of Khakassia (Southwestern Siberia). Geodiversitas, 33(3), 519-530.

Enk, J., Devault, A., Widga, C., Saunders, J., Szpak, P., Southon, J., ... Poinar, H. (2016). Mammuthus population dynamics in Late Pleistocene North America: divergence, phylogeography, and introgression. Frontiers in Ecology and Evolution, 4. https://doi.org/10.3389/fevo.2016.00042

Forsten, A. (1988). Middle Pleistocene replacement of stenonid horses by caballoid horses ecological implications. Palaeogeography, Palaeoclimatology, Palaeoecology, 65(1-2), 2333.

Forsten, A. (1992). Mitochondrial-DNA time-table and the evolution of Equus: comparison of molecular and palaeontological evidence. Annales Zoologici Fennici, 28(3-4), 301-309.

Forsten, A. (1996). Climate and the evolution of Equus (Perissodactyla, Equidae) in the PlioPleistocene of Eurasia. Acta Zoologica Cracoviensia, 39(1), 161-166.

Froese, D. G., Stiller, M., Heintzman, P. D., Reyes, A. V., Zazula, G. D., Soares, A. E. R., ... Shapiro, B. (2017). New fossil and genomic evidence constrains the timing of bison arrival in North America. Proceedings of the National Academy of Sciences of the United States of America, 114(13), 3457-3462.

Froese, D. G., Zazula, G. D., Westgate, J. A., Preece, S. J., Sanborn, P. T., Reyes, A. V., \& Pearce, N. J. G. (2009). The Klondike goldfields and Pleistocene environments of Beringia. GSA Today: A Publication of the Geological Society of America, 19(8), 4.

Goebel, T., Waters, M. R., \& O'Rourke, D. H. (2008). The late Pleistocene dispersal of modern humans in the Americas. Science, 319(5869), 1497-1502.

Graham, R. W., Belmecheri, S., Choy, K., Culleton, B. J., Davies, L. J., Froese, D., ... Wooller, M. J. (2016). Timing and causes of mid-Holocene mammoth extinction on St. Paul Island, Alaska. Proceedings of the National Academy of Sciences of the United States of America, 113(33), 9310-9314.

Groves, C. P., \& Willoughby, D. P. (1981). Studies on the taxonomy and phylogeny of the genus Equus. 1. Subgeneric classification of the recent species. Mammalia, 45(3). https://doi.org/10.1515/mamm.1981.45.3.321

Guthrie, R. D. (2003). Rapid body size decline in Alaskan Pleistocene horses before extinction. Nature, 426(6963), 169-171.

Guthrie, R. D. (2006). New carbon dates link climatic change with human colonization and Pleistocene extinctions. Nature, 441(7090), 207-209.

Harington, C. R. (2011). Pleistocene vertebrates of the Yukon Territory. Quaternary Science Reviews, 30(17-18), 2341-2354.

Harington, C. R., \& Clulow, F. V. (1973). Pleistocene Mammals from Gold Run Creek, Yukon Territory. Canadian Journal of Earth Sciences, 10(5), 697-759.

Hay, O. P. (1915). Contributions to the knowledge of the mammals of the Pleistocene of North America. Proceedings of the United States National Museum, 48(2086), 515-575.

Heintzman, P. D., Zazula, G. D., Cahill, J. A., Reyes, A. V., MacPhee, R. D. E., \& Shapiro, B. (2015). Genomic data from extinct North American Camelops revise camel evolutionary history. Molecular Biology and Evolution, 32(9), 2433-2440.

Hibbard, C. W. (1953). Equus (Asinus) calobatus Troxell and associated vertebrates from the Pleistocene of Kansas. Transactions of the Kansas Academy of Science. Kansas Academy of Science, 56(1), 111.

Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P. L. F., \& Orlando, L. (2013). mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. Bioinformatics , 29(13), 1682-1684.

Jónsson, H., Schubert, M., Seguin-Orlando, A., Ginolhac, A., Petersen, L., Fumagalli, M., ... Orlando, L. (2014). Speciation with gene flow in equids despite extensive chromosomal plasticity. Proceedings of the National Academy of Sciences of the United States of America, 111(52), 18655-18660.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., ... Drummond, A. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics, 28(12), 1647-1649.

Kircher, M., Sawyer, S., \& Meyer, M. (2012). Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. Nucleic Acids Research, 40(1), e3.

Koch, P. L., \& Barnosky, A. D. (2006). Late Quaternary extinctions: state of the debate. Annual Review of Ecology, Evolution, and Systematics, 37(1), 215-250.

Langmead, B., \& Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. Nature Methods, 9(4), 357-359.

Li, H., \& Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics , 26(5), 589-595.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., ... 1000 Genome Project Data Processing Subgroup. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics, 25(16), 2078-2079.

Lundelius, E. L., \& Stevens, M. S. (1970). Equus francisci Hay, a Small Stilt-Legged Horse, Middle Pleistocene of Texas. Journal of Palaeontology, 44(1), 148-153.

Macdonald, M. L., \& Toohey, L. M. (1992). The Species, Genera, and Tribes of the Living and Extinct Horses of the World 1758-1966: From the Work of Morris F. Skinner.

MacFadden, B. J. (1992). Fossil horses: systematics, palaeobiology, and evolution of the family Equidae. Cambridge University Press.

Macfadden, B. J. (1998). Equidae. In Evolution of Tertiary Mammals of North America (pp. 537559).

Meyer, M., \& Kircher, M. (2010). Illumina sequencing library preparation for highly multiplexed target capture and sequencing. Cold Spring Harbor Protocols, 2010(6), db.prot5448.

O’Dea, A., Lessios, H. A., Coates, A. G., Eytan, R. I., Restrepo-Moreno, S. A., Cione, A. L., ... Jackson, J. B. C. (2016). Formation of the Isthmus of Panama. Science Advances, 2(8), e1600883.

Orlando, L., Ginolhac, A., Zhang, G., Froese, D., Albrechtsen, A., Stiller, M., ... Willerslev, E. (2013). Recalibrating Equus evolution using the genome sequence of an early Middle Pleistocene horse. Nature, 499(7456), 74-78.

Orlando, L., Male, D., Alberdi, M. T., Prado, J. L., Prieto, A., Cooper, A., \& Hänni, C. (2008). Ancient DNA clarifies the evolutionary history of American Late Pleistocene equids. Journal of Molecular Evolution, 66(5), 533-538.

Orlando, L., Mashkour, M., Burke, A., Douady, C. J., Eisenmann, V., \& Hänni, C. (2006). Geographic distribution of an extinct equid (Equus hydruntinus: Mammalia, Equidae) revealed by morphological and genetical analyses of fossils. Molecular Ecology, 15(8), 2083-2093.

Orlando, L., Metcalf, J. L., Alberdi, M. T., Telles-Antunes, M., Bonjean, D., Otte, M., ... Cooper, A. (2009). Revising the recent evolutionary history of equids using ancient DNA. Proceedings of the National Academy of Sciences of the United States of America, 106(51), 21754-21759.

Quinn, J. H. (1957). Pleistocene Equidae of Texas. Bureau of Economic Geology, University of Texas, Report of Investigations, 33, 1-51.

R Development Core Team. (2008). R: A language and environment for statistical computing. Retrieved from R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Reimer, P. J., Bard, E., Bayliss, A., Beck, J. W., Blackwell, P. G., Bronk Ramsey, C., ... van der Plicht, J. (2013). IntCal13 and Marine13 radiocarbon age calibration curves 0-50,000 years cal BP. Radiocarbon, 55(4), 1869-1887.

Rohland, N., Siedel, H., \& Hofreiter, M. (2010). A rapid column-based ancient DNA extraction method for increased sample throughput. Molecular Ecology Resources, 10(4), 677-683. Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D. L., Darling, A., Höhna, S., ... Huelsenbeck, J. P. (2012). MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Systematic Biology, 61(3), 539-542.

Rybczynski, N., Gosse, J. C., Harington, C. R., Wogelius, R. A., Hidy, A. J., \& Buckley, M. (2013). Mid-Pliocene warm-period deposits in the High Arctic yield insight into camel evolution. Nature Communications, 4, 1550.

Ryder, O. A., Epel, N. C., \& Benirschke, K. (1978). Chromosome banding studies of the Equidae. Cytogenetics and Cell Genetics, 20(1-6), 332-350.

Sandom, C., Faurby, S., Sandel, B., \& Svenning, J.-C. (2014). Global late Quaternary megafauna extinctions linked to humans, not climate change. Proceedings. Biological Sciences / The Royal Society, 281(1787). https://doi.org/10.1098/rspb.2013.3254

Schmieder, R., \& Edwards, R. (2011). Quality control and preprocessing of metagenomic datasets. Bioinformatics , 27(6), 863-864.

Schubert, M., Ginolhac, A., Lindgreen, S., Thompson, J. F., Al-Rasheid, K. A. S., Willerslev, E., ... Orlando, L. (2012). Improving ancient DNA read mapping against modern reference genomes. BMC Genomics, 13, 178.

Scott, E. (2004). Pliocene and Pleistocene Horses from Porcupine Cave. In Biodiversity Response to Climate Change in the Middle Pleistocene (pp. 264-279).

Skinner, M. F., \& Hibbard, C. W. (1972). Early Pleistocene pre-glacial and glacial rocks and faunas of North-Central Nebraska. Bulletin of the American Museum of Natural History, 148, 1-148.

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics, 30(9), 1312-1313.

Steiner, C. C., \& Ryder, O. A. (2013). Characterization of $\operatorname{Prdm} 9$ in equids and sterility in mules. PloS One, 8(4), e61746.

Vilstrup, J. T., Seguin-Orlando, A., Stiller, M., Ginolhac, A., Raghavan, M., Nielsen, S. C. A., ... Orlando, L. (2013). Mitochondrial phylogenomics of modern and ancient equids. PloS One, 8(2), e55950.

Weinstock, J., Willerslev, E., Sher, A., Tong, W., Ho, S. Y. W., Rubenstein, D., ... Cooper, A. (2005). Evolution, systematics, and phylogeography of Pleistocene horses in the new world: a molecular perspective. PLoS Biology, 3(8), e241.

Welker, F., Collins, M. J., Thomas, J. A., Wadsley, M., Brace, S., Cappellini, E., ... MacPhee, R. D. E. (2015). Ancient proteins resolve the evolutionary history of Darwin's South American ungulates. Nature, 522(7554), 81-84.

Winans, M. C. (1985). Revision of North American fossil species of the genus Equus (Mammalia: Perissodactyla: Equidae). Dissertation, Univ. of Texas, Austin. 264 pp.

Zazula, G. D., MacPhee, R. D. E., Metcalfe, J. Z., Reyes, A. V., Brock, F., Druckenmiller, P. S., ... Southon, J. R. (2014). American mastodon extirpation in the Arctic and Subarctic predates human colonization and terminal Pleistocene climate change. Proceedings of the National Academy of Sciences of the United States of America, 111(52), 18460-18465.

## Supplementary tables 1-8

Supplementary table 1. Topological shape and support values for the best supported trees from the Bayesian and Maximum likelihood (ML) analyses of mtDNA data sets 1-3, including either the all or reduced partition sets, and with Hippidion sequences either included or excluded.

| Outgroup | Partitions | Hippidion? | Tips | MrBayes |  |  |  |  |  |  | RAxML |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  | Topology | Support |  |  |  |  |  | Topology | Support |  |  |  |  |  |
|  |  |  |  |  | Node A | Node B | Hippidion | NWSL | NCs | Caballines |  | Node A | Node B | Hippidion | NWSL | NCs | Caballines |
| White rhino <br> (Data set 1) | All | Excluded | 63 | 1/2/3 | 0.996* | N/A | N/A | 1.000 | 1.000 | 1.000 | 1/2/3 | 71* | N/A | N/A | 100 | 99 | 100 |
|  |  | Included | 69 | 2 | 0.751 | 1.000* | 1.000 | 1.000 | 1.000 | 1.000 | 1 | 64* | 96* | 100 | 100 | 100 | 100 |
|  | Reduced | Excluded | 63 | $1 / 2 / 3$ | 1.000* | N/A | N/A | 1.000 | 1.000 | 1.000 | $1 / 2 / 3$ | 100* | N/A | N/A | 99 | 100 | 100 |
|  |  | Included | 69 | 2 | 0.948 | 1.000* | 1.000 | 1.000 | 1.000 | 1.000 | 2 | 73 | 98* | 100 | 99 | 100 | 100 |
| Malayan tapir <br> (Data set 2) | All | Excluded | 63 | $5 / 7$ | 0.971 | N/A | N/A | 1.000 | 1.000 | 1.000 | $5 / 7$ | 87 | N/A | N/A | 100 | 99 | 99 |
|  |  | Included | 69 | 6 | 0.808 | 0.867 | 1.000 | 1.000 | 1.000 | 1.000 | 6 | 55 | 63 | 100 | 100 | 100 | 100 |
|  | Reduced | Excluded | 63 | $1 / 2 / 3$ | 0.675* | N/A | N/A | 1.000 | 1.000 | 1.000 | 4 / 6 | 28 | N/A | N/A | 100 | 96 | 98 |
|  |  | Included | 69 | 3 | 0.685 | 0.864* | 1.000 | 1.000 | 1.000 | 1.000 | 3 | 70 | 69 | 100 | 100 | 100 | 100 |
| Dog+ceratomorphs <br> (Data set 3) | All | Excluded | 71 | $1 / 2 / 3$ | 0.598* | N/A | N/A | 1.000 | 1.000 | 1.000 | 4 / 6 | 59 | N/A | N/A | 100 | 100 | 100 |
|  |  | Included | 77 | 1 | 1.000* | 1.000* | 1.000 | 1.000 | 1.000 | 1.000 | 1 | 94* | 96* | 100 | 100 | 100 | 100 |
|  | Reduced | Excluded | 71 | $1 / 2 / 3$ | 0.999* | N/A | N/A | 1.000 | 1.000 | 1.000 | $1 / 2 / 3$ | 97* | N/A | N/A | 100 | 100 | 100 |
|  |  | Included | 77 | 1 | 1.000* | 1.000* | 1.000 | 1.000 | 1.000 | 1.000 | 1 | 99* | 100* | 100 | 100 | 100 | 100 |

Topology numbers and node letters refer to those outlined in Appendix 2-figure 3. Bayesian posterior probability support of $>0.99$ and ML bootstrap support of $>95 \%$ are in bold for nodes A and B. *support for nodes that are consistent with topology 1 in Appendix 2figure 3. NCs: non-caballines.

Supplementary table 2. The a posteriori phylogenetic placement likelihood for eight ceratomorph (rhino and tapir) outgroups using a ML evolutionary placement algorithm, whilst varying the partition set used (all or reduced), and either including or excluding Hippidion sequences.

| Hippidion? | Outgroup | All |  |  |  |  | Reduced |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | 1 | 2 | 3 | 6 | 7 | 1 | 2 | 3 | 6 | 7 |
| Included | Tapirus terrestris (AJ428947) | 0.457 | 0.317 | 0.205 | 0.018 |  | 0.410 | 0.391 | 0.199 |  |  |
| Included | Tapirus indicus (NC023838) | 0.276 | 0.105 | 0.225 | 0.389 |  | 0.536 | 0.298 | 0.166 |  |  |
| Included | Coelodonta antiquitatis (NC012681) | 0.998 |  |  |  |  | 0.411 | 0.554 | 0.035 |  |  |
| Included | Dicerorhinus sumatrensis (NC012684) | 0.981 |  | 0.009 |  |  | 0.983 | 0.015 |  |  |  |
| Included | Rhinoceros unicornis (NC001779) | 0.998 |  |  |  |  | 0.998 |  |  |  |  |
| Included | Rhinoceros sondaicus (NC012683) | 0.989 | 0.006 |  |  |  | 0.895 | 0.102 |  |  |  |
| Included | Ceratotherium simum (NC001808) | 0.448 | 0.499 | 0.052 |  |  | 0.296 | 0.704 |  |  |  |
| Included | Diceros bicornis (NC012682) | 0.918 | 0.065 | 0.018 |  |  | 0.996 |  |  |  |  |
|  | Topology |  | 1/2/3 |  | 4 / 6 | $5 / 7$ |  | 1/2/3 |  | $4 / 6$ | 5/7 |
| Excluded | Tapirus terrestris (AJ428947) |  | 0.550 |  | 0.312 | 0.138 |  | 0.987 |  |  | 0.012 |
| Excluded | Tapirus indicus (NC023838) |  | 0.050 |  | 0.908 | 0.041 |  | 0.995 |  |  |  |
| Excluded | Coelodonta antiquitatis (NC012681) |  | 0.248 |  | 0.451 | 0.301 |  | 1.000 |  |  |  |
| Excluded | Dicerorhinus sumatrensis (NC012684) |  | 0.155 |  | 0.553 | 0.291 |  | 1.000 |  |  |  |
| Excluded | Rhinoceros unicornis (NC001779) |  | 0.530 |  | 0.334 | 0.136 |  | 1.000 |  |  |  |
| Excluded | Rhinoceros sondaicus (NC012683) |  | 0.733 |  | 0.196 | 0.071 |  | 1.000 |  |  |  |
| Excluded | Ceratotherium simum (NC001808) |  | 0.949 |  | 0.018 | 0.033 |  | 1.000 |  |  |  |
| Excluded | Diceros bicornis (NC012682) |  | 0.851 |  | 0.073 | 0.075 |  | 1.000 |  |  |  |

The best supported placement for each analysis has been given a gray background, and likelihoods $>0.95$ are in bold. Topology numbers refer to those outlined in Appendix 2-figure 3. Genbank accession numbers are given in parentheses after outgroup names.

Supplementary table 3. Bayesian time tree analysis results, with support and estimated divergence times for major nodes, and the tMRCAs for Haringtonhippus, E. asinus, and E. quagga summarized.

| E. ovodovi? | Partition | Root prior | Hippidion- <br> Haringtonhippus/Equus <br> Node height |  | Haringtonhippus-Equus |  |  | Equus |  |  | Haringtonhippus |  |  | E. asinus |  |  | E. quagga |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  |  | Posterior | Node height |  | Posterior | Node height |  | Posterior | Node height |  | Posterior | Node height |  | Posterior | Node height |  |
|  |  |  | Median | 95\% HPD |  | Median | 95\% HPD |  | Median | 95\% HPD |  | Median | 95\% HPD |  | Median | 95\% HPD |  | Median | 95\% HPD |
| Included | All | Constrained | 6.19 | 5.22-7.23 | 0.999 | 4.59 | 4.18-5.06 | 1.000 | 4.10 | 3.81-4.39 | 1.000 | 0.264 | 0.198-0.348 | 1.000 | 0.720 | 0.556-0.916 | 1.000 | 0.599 | 0.449-0.770 |
|  |  | Not constrained | 6.20 | 5.20-7.24 | 1.000 | 4.59 | 4.17-5.04 | 1.000 | 4.10 | 3.81-4.38 | 1.000 | 0.263 | 0.196-0.342 | 1.000 | 0.719 | 0.549-0.906 | 1.000 | 0.600 | 0.446-0.770 |
|  | Reduced | Constrained | 6.30 | 5.15-7.50 | 0.997 | 4.58 | 4.11-5.11 | 0.998 | 4.08 | 3.77-4.36 | 1.000 | 0.334 | 0.224-0.463 | 1.000 | 0.755 | 0.544-0.992 | 1.000 | 0.605 | 0.411-0.816 |
|  |  | Not constrained | 6.35 | 5.21-7.60 | 0.996 | 4.59 | 4.09-5.09 | 0.996 | 4.08 | 3.79-4.38 | 1.000 | 0.335 | 0.228-0.465 | 1.000 | 0.755 | 0.550-0.998 | 1.000 | 0.605 | 0.412-0.821 |
| Excluded | All | Constrained | 6.21 | 5.22-7.21 | 1.000 | 4.58 | 4.16-5.04 | 1.000 | 4.10 | 3.81-4.39 | 1.000 | 0.269 | 0.200-0.352 | 1.000 | 0.735 | 0.561-0.924 | 1.000 | 0.608 | 0.459-0.777 |
|  |  | Not constrained | 6.29 | 5.33-7.33 | 1.000 | 4.60 | 4.17-5.05 | 1.000 | 4.11 | 3.81-4.40 | 1.000 | 0.273 | 0.200-0.354 | 1.000 | 0.742 | 0.560-0.921 | 1.000 | 0.618 | 0.463-0.791 |
|  | Reduced | Constrained | 6.44 | 5.36-7.63 | 0.999 | 4.60 | 4.12-5.11 | 0.998 | 4.08 | 3.78-4.36 | 1.000 | 0.338 | 0.227-0.466 | 1.000 | 0.778 | 0.570-1.020 | 1.000 | 0.609 | 0.431-0.851 |
|  |  | Not constrained | 6.49 | 5.38-7.66 | 0.999 | 4.61 | 4.13-5.13 | 0.999 | 4.09 | 3.79-4.38 | 1.000 | 0.339 | 0.228-0.467 | 1.000 | 0.781 | 0.574-1.027 | 1.000 | 0.614 | 0.423-0.819 |
|  |  | Range | 6.19-6.49 | 5.15-7.66 | 0.996-1.000 | 4.58-4.61 | 4.09-5.13 | 0.996-1.000 | 4.08-4.11 | 3.77-4.40 | 1.000 | 0.263-0.339 | 0.196-0.467 | 1.000 | 0.719-0.781 | 0.544-1.027 | 1.000 | 0.599-0.618 | 0.441-0.851 |

All analyses supported topology 1 in Appendix 2-figure 3. HPD: highest posterior density.

Supplementary table 4. The a posteriori phylogenetic placement likelihood for 12 published equid mitochondrial sequences using the ML evolutionary placement algorithm, whilst varying the partition set used (all or reduced), and either including or excluding Hippidion sequences.

| Hippidion? | Published sample | All |  |  |  | Reduced |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | Sister to <br> E. ovodovi | Sister to <br> Hippidion | Within NWSL | Sister to NWSL | Within NWSL | Sister to NWSL |
| Included | E. ovodovi (ACAD2305) | 1.000 |  |  |  | N/A | N/A |
| Included | E. ovodovi (ACAD2302) | 1.000 |  |  |  | N/A | N/A |
| Included | E. ovodovi (ACAD2303) | 1.000 |  |  |  | N/A | N/A |
| Included | H. devillei (ACAD3615) |  | 1.000 |  |  | N/A | N/A |
| Included | H. devillei (ACAD3625) |  | 1.000 |  |  | N/A | N/A |
| Included | H. devillei (ACAD3627) |  | 1.000 |  |  | N/A | N/A |
| Included | H. devillei (ACAD3628) |  | 0.999 |  |  | N/A | N/A |
| Included | H. devillei (ACAD3629) |  | 0.999 |  |  | N/A | N/A |
| Included | NWSL equid (JW125) |  |  | 0.996 |  | N/A | N/A |
| Included | NWSL equid (JW126) |  |  | 0.999 |  | N/A | N/A |
| Included | NWSL equid (JW328; JX312726) |  |  | 1.000 |  | 0.996 |  |
| Included | NWSL equid (MS272; JX312727) |  |  |  | 1.000 |  | 1.000 |
| Excluded | E. ovodovi (ACAD2305) | 1.000 |  |  |  | N/A | N/A |
| Excluded | E. ovodovi (ACAD2302) | 1.000 |  |  |  | N/A | N/A |
| Excluded | E. ovodovi (ACAD2303) | 1.000 |  |  |  | N/A | N/A |
| Excluded | NWSL equid (JW125) |  |  | 0.996 |  | N/A | N/A |
| Excluded | NWSL equid (JW126) |  |  | 0.999 |  | N/A | N/A |
| Excluded | NWSL equid (JW328; JX312726) |  |  | 1.000 |  | 0.995 |  |
| Excluded | NWSL equid (MS272; JX312727) |  |  |  | 1.000 |  | 1.000 |

For mitochondrial genome sequences, sample names and Genbank accession numbers are given in parentheses after the species names. For short mitochondrial DNA fragments, sample names are given in parentheses after the species names.

1309 Supplementary table 5. Summary of nuclear genome data from all 17 NWSL equids pooled together and analyzed using approach four.

| Genome coordinates | Fragment length bin (bp) | Count of all sites | Private transversion frequency |  |  |  |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  | Absolute |  |  | Relative |  |  | Combined relative frequencies |  |  | Ratio between relative frequencies |  |  |
|  |  |  | NWSL | Donkey | Horse | NWSL | Donkey | Horse | NWSL+ <br> Donkey | NWSL+ Horse | Donkey+ <br> Horse | NWSL: <br> Donkey | NWSL: <br> Horse | Donkey: <br> Horse |
| Horse-based | 30-39 | 49318650 | 64465 | 83016 | 53488 | $1.307 \mathrm{E}-03$ | $1.683 \mathrm{E}-03$ | $1.085 \mathrm{E}-03$ | $2.990 \mathrm{E}-03$ | 2.392E-03 | $2.768 \mathrm{E}-03$ | 0.777 | 1.205 | 1.552 |
|  | 40-49 | 59781463 | 100232 | 100055 | 73823 | $1.677 \mathrm{E}-03$ | $1.674 \mathrm{E}-03$ | $1.235 \mathrm{E}-03$ | $3.350 \mathrm{E}-03$ | $2.912 \mathrm{E}-03$ | $2.909 \mathrm{E}-03$ | 1.002 | 1.358 | 1.355 |
|  | 50-59 | 57405192 | 107455 | 96797 | 75746 | $1.872 \mathrm{E}-03$ | $1.686 \mathrm{E}-03$ | $1.319 \mathrm{E}-03$ | $3.558 \mathrm{E}-03$ | 3.191E-03 | $3.006 \mathrm{E}-03$ | 1.110 | 1.419 | 1.278 |
|  | 60-69 | 49392344 | 97377 | 82944 | 68033 | $1.971 \mathrm{E}-03$ | $1.679 \mathrm{E}-03$ | $1.377 \mathrm{E}-03$ | $3.651 \mathrm{E}-03$ | $3.349 \mathrm{E}-03$ | $3.057 \mathrm{E}-03$ | 1.174 | 1.431 | 1.219 |
|  | 70-79 | 41109339 | 84198 | 68807 | 58380 | $2.048 \mathrm{E}-03$ | $1.674 \mathrm{E}-03$ | $1.420 \mathrm{E}-03$ | $3.722 \mathrm{E}-03$ | $3.468 \mathrm{E}-03$ | $3.094 \mathrm{E}-03$ | 1.224 | 1.442 | 1.179 |
|  | 80-89 | 28683294 | 65569 | 47845 | 44893 | $2.286 \mathrm{E}-03$ | $1.668 \mathrm{E}-03$ | $1.565 \mathrm{E}-03$ | $3.954 \mathrm{E}-03$ | $3.851 \mathrm{E}-03$ | $3.233 \mathrm{E}-03$ | 1.370 | 1.461 | 1.066 |
|  | 90-99 | 17754307 | 43342 | 29760 | 29359 | $2.441 \mathrm{E}-03$ | $1.676 \mathrm{E}-03$ | $1.654 \mathrm{E}-03$ | $4.117 \mathrm{E}-03$ | $4.095 \mathrm{E}-03$ | $3.330 \mathrm{E}-03$ | 1.456 | 1.476 | 1.014 |
|  | 100-109 | 10556561 | 26655 | 17725 | 17424 | $2.525 \mathrm{E}-03$ | $1.679 \mathrm{E}-03$ | $1.651 \mathrm{E}-03$ | $4.204 \mathrm{E}-03$ | 4.176E-03 | $3.330 \mathrm{E}-03$ | 1.504 | 1.530 | 1.017 |
|  | 110-119 | 6384882 | 16336 | 10713 | 10562 | $2.559 \mathrm{E}-03$ | $1.678 \mathrm{E}-03$ | $1.654 \mathrm{E}-03$ | $4.236 \mathrm{E}-03$ | $4.213 \mathrm{E}-03$ | $3.332 \mathrm{E}-03$ | 1.525 | 1.547 | 1.014 |
|  | 120-129 | 3664296 | 9435 | 6159 | 6052 | $2.575 \mathrm{E}-03$ | $1.681 \mathrm{E}-03$ | $1.652 \mathrm{E}-03$ | $4.256 \mathrm{E}-03$ | $4.226 \mathrm{E}-03$ | $3.332 \mathrm{E}-03$ | 1.532 | 1.559 | 1.018 |
|  | Total/Mean: | 38360046 | 95768 | 64357 | 63397 | $2.525 \mathrm{E}-03$ | $1.678 \mathrm{E}-03$ | $1.652 \mathrm{E}-03$ | $4.203 \mathrm{E}-03$ | $4.177 \mathrm{E}-03$ | $3.331 \mathrm{E}-03$ | 1.504 | 1.528 | 1.016 |
| Donkey-based | 30-39 | 28276643 | 38149 | 30750 | 47684 | $1.349 \mathrm{E}-03$ | $1.087 \mathrm{E}-03$ | $1.686 \mathrm{E}-03$ | $2.437 \mathrm{E}-03$ | $3.035 \mathrm{E}-03$ | $2.774 \mathrm{E}-03$ | 1.241 | 0.800 | 0.645 |
|  | 40-49 | 34035949 | 56874 | 41946 | 57300 | $1.671 \mathrm{E}-03$ | $1.232 \mathrm{E}-03$ | $1.684 \mathrm{E}-03$ | $2.903 \mathrm{E}-03$ | 3.355E-03 | $2.916 \mathrm{E}-03$ | 1.356 | 0.993 | 0.732 |
|  | 50-59 | 32495044 | 60172 | 43517 | 54413 | $1.852 \mathrm{E}-03$ | $1.339 \mathrm{E}-03$ | $1.675 \mathrm{E}-03$ | $3.191 \mathrm{E}-03$ | $3.526 \mathrm{E}-03$ | $3.014 \mathrm{E}-03$ | 1.383 | 1.106 | 0.800 |
|  | 60-69 | 27941039 | 54058 | 38665 | 46445 | $1.935 \mathrm{E}-03$ | $1.384 \mathrm{E}-03$ | $1.662 \mathrm{E}-03$ | $3.319 \mathrm{E}-03$ | $3.597 \mathrm{E}-03$ | $3.046 \mathrm{E}-03$ | 1.398 | 1.164 | 0.832 |
|  | 70-79 | 23282014 | 46578 | 33146 | 38794 | $2.001 \mathrm{E}-03$ | $1.424 \mathrm{E}-03$ | $1.666 \mathrm{E}-03$ | $3.424 \mathrm{E}-03$ | $3.667 \mathrm{E}-03$ | $3.090 \mathrm{E}-03$ | 1.405 | 1.201 | 0.854 |
|  | 80-89 | 16178161 | 36031 | 25345 | 27071 | $2.227 \mathrm{E}-03$ | $1.567 \mathrm{E}-03$ | $1.673 \mathrm{E}-03$ | $3.794 \mathrm{E}-03$ | $3.900 \mathrm{E}-03$ | $3.240 \mathrm{E}-03$ | 1.422 | 1.331 | 0.936 |
|  | 90-99 | 9992386 | 23517 | 16457 | 16802 | $2.353 \mathrm{E}-03$ | $1.647 \mathrm{E}-03$ | $1.681 \mathrm{E}-03$ | $4.000 \mathrm{E}-03$ | $4.035 \mathrm{E}-03$ | $3.328 \mathrm{E}-03$ | 1.429 | 1.400 | 0.979 |
|  | 100-109 | 5915994 | 14259 | 9723 | 10014 | $2.410 \mathrm{E}-03$ | $1.644 \mathrm{E}-03$ | $1.693 \mathrm{E}-03$ | $4.054 \mathrm{E}-03$ | $4.103 \mathrm{E}-03$ | $3.336 \mathrm{E}-03$ | 1.467 | 1.424 | 0.971 |
|  | 110-119 | 3576444 | 8713 | 5943 | 5996 | $2.436 \mathrm{E}-03$ | $1.662 \mathrm{E}-03$ | $1.677 \mathrm{E}-03$ | $4.098 \mathrm{E}-03$ | $4.113 \mathrm{E}-03$ | $3.338 \mathrm{E}-03$ | 1.466 | 1.453 | 0.991 |
|  | 120-129 | 2058710 | 5031 | 3467 | 3414 | $2.444 \mathrm{E}-03$ | $1.684 \mathrm{E}-03$ | $1.658 \mathrm{E}-03$ | $4.128 \mathrm{E}-03$ | $4.102 \mathrm{E}-03$ | $3.342 \mathrm{E}-03$ | 1.451 | 1.474 | 1.016 |
|  | Total/Mean: | 21543534 | 51520 | 35590 | 36226 | $2.411 \mathrm{E}-03$ | $1.659 \mathrm{E}-03$ | $1.677 \mathrm{E}-03$ | $4.070 \mathrm{E}-03$ | $4.088 \mathrm{E}-03$ | $3.336 \mathrm{E}-03$ | 1.453 | 1.437 | 0.989 |

Minimum and maximum NWSL:Equus ratios between relative frequencies are in bold, and are used for the divergence estimates in
1312 Figure 1-Figure supplement 3. Total and mean values are for the four longest bins only (90-99 to 120-129 bp). Mean values equally

1314 Supplementary table 6. Sex determination analysis of 17 NWSL equids.

|  | Sample | Museum accession | Chromosome ratio | Inferred sex |
| :--- | :--- | :--- | :---: | :---: |
| AF037 | YG 402.235 | 0.48 | male |  |
| JK166 | LACM(CIT) 109/150807 | 0.93 | female |  |
| JK167 | LACM(CIT) 109/149291 | 0.91 | female |  |
| JK207 | LACM(CIT) 109/156450 | 0.92 | female |  |
| JK260 | KU 47800 | 0.95 | female |  |
| JK276 | KU 53678 | 0.91 | female |  |
| MS341 | YG 303.1085 | 0.50 | male |  |
| MS349 | YG 130.55 | 0.48 | male |  |
|  | MS439 | YG 401.387 | 0.98 | female |
|  | PH008 | YG 404.205 | 0.90 | female |
|  | PH013 | YG 130.6 | 0.87 | probable female |
|  | PH014 | YG 303.371 | 0.46 | male |
|  | PH015 | YG 404.662 | 0.44 | male |
| PH021 | YG 29.169 | 0.83 | probable female |  |
|  | PH023 | YG 160.8 | 0.91 | female |
|  | PH036 | YG 76.2 | 0.81 | probable female |
|  | PH047 | YG 404.663 | 0.88 | probable female |

1316 Chromosome ratio is the relative mapping frequency ratio between all autosomes and the Xchromosome. Males are inferred if the ratio is $0.45-0.55$ and females if the ratio is $0.9-1.1$.

Supplementary table 7. Summary of the number and type of synapomorphic bases for each of the three examined equid genera.

| Substitution | Hippidion | Haringtonhippus | Equus |
| :--- | ---: | ---: | ---: |
| Transition | 338 | 147 | 66 |
| Transversion | 43 | 22 | 4 |
| Insertion | 2 | 4 | 0 |
| Deletion | 3 | 0 | 0 |
| Total* | 391 | 178 | 75 |

A full list of these substitutions, and their position relative to the E. caballus reference mitochondrial genome (NC_001640), can be found in Supplementary file 5. *total includes a further five synapomorphic sites that have unique states in each genus.

1327 Supplementary table 8. Selected models of molecular evolution for partitions of the first five mtDNA genome alignment data sets.

| Data set | Coding1 |  | Coding2 |  | Coding 3 |  | rRNAs |  | tRNAs |  | CR |  | Total length |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Length | Model | Length | Model | Length | Model | Length | Model | Length | Model | Length | Model | All | Reduced |
| 1. White rhino out | 3803 | GTR+I+G | 3803 | HKY+I+G | 3803 | GTR+I+G | 2579 | GTR+I+G | 1529 | HKY+I+G | 1066 | HKY*+I+G | 16583 | 11714 |
| 2. Malayan tapir out | 3803 | GTR+I+G | 3803 | HKY+I+G | 3803 | GTR+I+G | 2585 | GTR+I+G | 1530 | HKY+I+G | 1065 | HKY*+G | 16589 | 11721 |
| 3. Dog+ceratomorphs out | 3803 | GTR+I+G | 3803 | HKY+I+G | 3803 | GTR+I+G | 2615 | GTR+I+G | 1540 | HKY+I+G | N/A | N/A | 15564 | 11761 |
| 4. EPA | 3803 | GTR+I+G | 3803 | TrN+I+G | 3803 | GTR+I+G | 2601 | GTR+I+G | 1534 | HKY+I+G | 1118 | HKY $+\mathrm{I}+\mathrm{G}$ | 16662 | 11741 |
| 5. Equids | 3802 | TrN+I+G | 3802 | TrN+I+G | 3802 | GTR+G | 2571 | TrN+I+G | 1528 | HKY+I | 971 | HKY+G | 16476 | 11703 |

All lengths are in base pairs. Reduced length excludes the Coding3 and CR partitions. For all RAxML analyses the GTR model was implemented. *The TrN model was selected, but this cannot be implemented in MrBayes and so the HKY model was used. EPA: evolutionary placement algorithm; CR: control region.

## Supplementary files

Supplementary file 1. Metadata for all samples used in the mitochondrial genomic analyses.

Supplementary file 2. Statistics from the phylogenetic inference analyses of nuclear genomes using all four approaches. (A) Read mapping statistics. (B) Relative transversion frequencies for approaches 1-3. (C) Relative private transversion frequencies for approach 4.

Supplementary file 3. Data from the sex determination analyses.

Supplementary file 4. Measurement data for (A) equid third metatarsals, which were used in the morphometrics analysis, and (B) other NWSL equid elements.

Supplementary file 5. A compilation of all 634 putative synapomorphic sites in the mitochondrial genome for Hippidion, Haringtonhippus, and Equus (A), with a comparison to the published MS272 mitochondrial genome sequence at the 140 sites with a base state that matches one of the three genera (B).

Supplementary file 6. Rscript used for the morphometric analysis.

Supplementary file 7. TREE files of all mitochondrial trees generated by this study.

Supplementary file 8. NEXUS files of all six mitochondrial genome alignment data sets.

Supplementary file 9. XML files for all BEAST analyses.

