Received: 11 July 2023 | Revised: 17 May 2024 | Accepted: 14 June 2024

DOI: 10.1111/mec.17440

ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

The first complete genome of the extinct European wild ass (*Equus hemionus hydruntinus*)

Mustafa Özkan[1] | Kanat Gürün[1] | Eren Yüncü[1] | Kıvılcım Başak Vural[1] | Gözde Atağ[1] | Ali Akbaba[2,3] | Fatma Rabia Fidan[1,4] | Ekin Sağlıcan[5] | Ezgi N. Altınışık[6] | Dilek Koptekin[5] | Kamilla Pawłowska[7] | Ian Hodder[8] | Sarah E. Adcock[9] | Benjamin S. Arbuckle[10] | Sharon R. Steadman[11] | Gregory McMahon[12] | Yılmaz Selim Erdal[6] | C. Can Bilgin[1] | İnci Togan[1] | Eva-Maria Geigl[13] | Anders Götherström[14] | Thierry Grange[13] | Füsun Özer[5] | Mehmet Somel[1]

[1]Department of Biological Sciences, Middle East Technical University, Ankara, Turkey

[2]Department of Anthropology, Ankara University, Ankara, Turkey

[3]Alparslan University, Muş, Turkey

[4]Cancer Dynamics Laboratory, The Francis Crick Institute, London, UK

[5]Department of Health Informatics, Middle East Technical University, Ankara, Turkey

[6]Department of Anthropology, Human_G Laboratory, Hacettepe University, Ankara, Turkey

[7]Department of Palaeoenvironmental Research, Adam Mickiewicz University, Poznań, Poland

[8]Department of Anthropology, Stanford University, Stanford, California, USA

[9]Institute for the Study of the Ancient World, New York University, New York, New York, USA

[10]Department of Anthropology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA

[11]Department of Sociology/Anthropology, SUNY Cortland, Cortland, New York, USA

[12]Classics, Humanities and Italian Studies Department, University of New Hampshire, Durham, New Hampshire, USA

[13]Institut Jacques Monod, CNRS, Université de Paris, Paris, France

[14]Department of Archaeology and Classical Studies, Stockholm University, Stockholm, Sweden

Correspondence
Mustafa Özkan and Mehmet Somel, Department of Biological Sciences, Middle East Technical University, Ankara, Turkey.
Email: mustafa.ozkan.sci@gmail.com and somel.mehmet@googlemail.com

Füsun Özer, Department of Health Informatics, Middle East Technical University, Ankara, Turkey.
Email: fusun.ozer@gmail.com

Thierry Grange, Institut Jacques Monod, CNRS, Université de Paris, Paris, France.
Email: thierry.grange@ijm.fr

Abstract

We present palaeogenomes of three morphologically unidentified Anatolian equids dating to the first millennium BCE, sequenced to a coverage of 0.6–6.4×. Mitochondrial DNA haplotypes of the Anatolian individuals clustered with those of *Equus hydruntinus* (or *Equus hemionus hydruntinus*), the extinct European wild ass, secular name 'hydruntine'. Further, the Anatolian wild ass whole genome profiles fell outside the genomic diversity of other extant and past Asiatic wild ass (*E. hemionus*) lineages. These observations suggest that the three Anatolian wild asses represent hydruntines, making them the latest recorded survivors of this lineage, about a millennium

Eva-Maria Geigl, Anders Götherstrom and Thierry Grange contributed equally.

Funding information
European Research Council Consolidator Grant H2020, Grant/Award Number: 772390; European Commission Horizon 2020 TWINNING Programme, Grant/Award Number: 952317; Türkiye Bilimsel ve Teknolojik Araştırma Kurumu, Grant/Award Number: 117Z991

Handling Editor: Andrew DeWoody

later than the latest observations in the zooarchaeological record. Our mitogenomic and genomic analyses indicate that *E. h. hydruntinus* was a clade belonging to ancient and present-day *E. hemionus* lineages that radiated possibly between 0.6 and 0.8 Mya. We also find evidence consistent with recent gene flow between hydruntines and Middle Eastern wild asses. Analyses of genome-wide heterozygosity and runs of homozygosity suggest that the Anatolian wild ass population may have lost genetic diversity by the mid-first millennium BCE, a possible sign of its eventual demise.

KEYWORDS
ancient DNA, Asiatic wild ass, demography, *Equus hemionus hydruntinus*, population genetics, taxonomy

1 | INTRODUCTION

Since its palaeontological description more than a century ago (Regalia, 1907), the 'European wild ass', *Equus hydruntinus* (or *Equus hemionus hydruntinus*), has remained an enigmatic taxon (Bennett et al., 2017; Boulbes & van Asperen, 2019; Burke et al., 2003; Geigl & Grange, 2012; Orlando et al., 2006). This species, which hereinafter we will name the hydruntine, was a gracile non-caballine equid, once roamed open habitats in Europe and Southwest Asia (Figure 1a) and featured in Upper Palaeolithic cave art (Bennett et al., 2017; BernÁldez-Sánchez & García-Viñas, 2019; Cleyet-Merle & Madelaine, 1991) and on Neolithic pottery (Bennett et al., 2017). Its history in the fossil record starts with the Late Middle or Late Pleistocene and ends within the Middle/Late Holocene, when it goes extinct (Boulbes & van Asperen, 2019; Crees & Turvey, 2014; Geigl & Grange, 2012).

Early research in the 20th century comparing *E. hydruntinus* remains with those of other equids identified resemblances to diverse taxa, including the African asses (*E. asinus*) and the zebra (*E. zebra*), the Asiatic asses (*E. hemionus*), or the extinct stenonine horses, leaving its phylogenetic position disputed for many decades (Azzaroli, 1991; Davis, 1980; Eisenmann & Baylac, 2000; Eisenmann & Mashkour, 2000; Forsten & Ziegler, 1995; Stehlin & Graziosi, 1935). Within the last two decades, osteological studies on new fossil findings concluded that *E. hydruntinus* was systematically closer to extant Asian asses, that is, hemiones, than to African asses or zebras (Burke et al., 2003; Eisenmann & Mashkour, 1999; Orlando et al., 2006) (Figure 1a). Ancient DNA analyses of mitochondrial DNA (mtDNA) supported this conclusion: partial and complete mtDNA sequences from *E. hydruntinus* and extant hemiones (the kiang of Tibet; the kulan of Mongolia; the kulan of Turkmenistan; the onager of Iran) were found to cluster together to the exclusion of other equids (Bennett et al., 2017; Catalano et al., 2020; Orlando et al., 2006, 2009). Moreover, these mtDNA studies suggested that the hemione – hydruntine division may represent taxonomic over-splitting (Bennett et al., 2017; Orlando et al., 2006, 2009). Bennett et al. (2017) pointed out that in their mtDNA analyses hemiones and hydruntines did not appear reciprocally monophyletic. This could be explained by rapid diversification of the *E. hemionus* lineages, including hydruntines, creating an unresolved radiation node (Model 1 in Figure 1b). Accordingly, hydruntines could also be considered a subspecies of *E. hemionus*, *E. h. hydruntinus*, similar to the kulan (*E. h. kulan*) and onager (*E. h. onager*), considered *E. hemionus* subspecies by the IUCN (Kaczensky et al., 2015).

However, the phylogenetic patterns described were only based on partial mtDNA sequences. Thus, it is possible that analyses of full mtDNA and of nuclear loci would reveal different patterns, for example, an early hydruntine–hemione split (Model 2 in Figure 1b). In line with the early-split idea, osteological analyses have suggested that *E. hydruntinus* carried a number of unique adaptations distinct from other hemiones, such as a short and wide muzzle adapted to relatively cold environmental conditions (Boulbes & van Asperen, 2019; van Asperen, 2012). Given the equivocal evidence, there have been calls for in-depth morphometric analyses (Twiss et al., 2017) and for the analysis of nuclear genomic data to resolve the issue (Boulbes & van Asperen, 2019; Crees & Turvey, 2014).

Another controversy surrounding the hydruntine involves its extinction dynamics. Crees and Turvey (2014) studied Holocene zooarchaeological records of hydruntines along with palaeovegetation data, suggesting that during the Holocene, the hydruntine range was highly fragmented and restricted to regions with relatively open habitats, such as the Danube basin and the Anatolian steppe. The authors predicted its extinction in the Danube region by the third millennium BCE, and in Iran and South Caucasus possibly within the first millennium BCE. Other scholars have suggested hydruntines in Iran and in Anatolia could have gone extinct during or before the second millennium BCE (Mashkour et al., 1999), attributing its extinction to a combination of factors including increased aridity associated with the 4.2-ka event, competition with livestock for pasture resources and hunting (Guimaraes et al., 2020). Nevertheless, due to the relative rarity of hydruntines in the zooarchaeological record (compared to, e.g. red deer) and also due to difficulties in morphological identification (Geigl & Grange, 2012; Twiss et al., 2017), the timing of the hydruntine extinction remained largely uncertain (Boulbes & van Asperen, 2019; Crees & Turvey, 2014; Nores et al., 2015).

Here we present the first full genomic data genetically attributable to the hydruntine, obtained from three Anatolian equids from first millennium BCE Central Anatolia. The analysis of these palaeogenomes allows us to resolve questions on the phylogenetics, demographic history and extinction dynamics of this taxon.

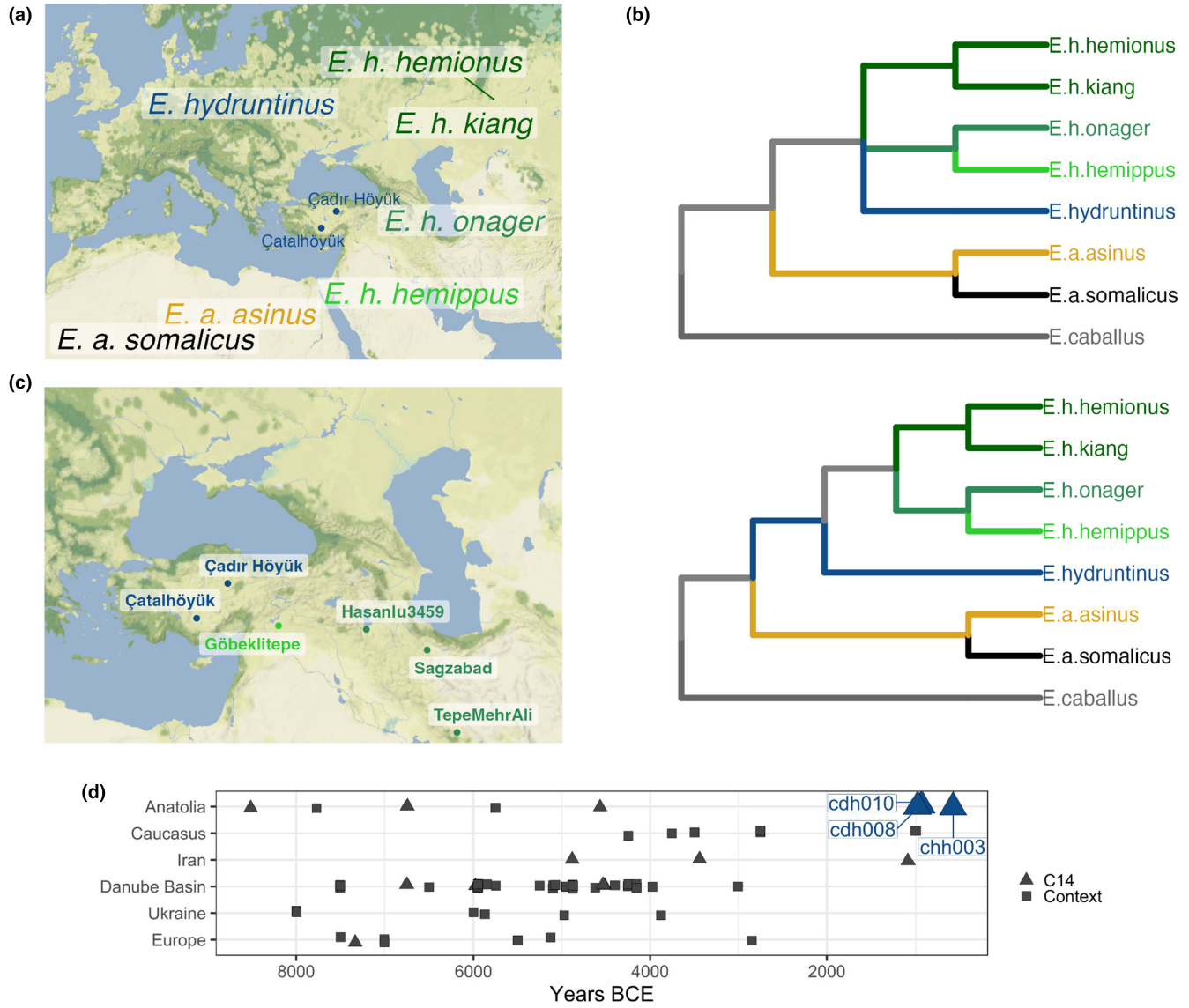


FIGURE 1 (a) The map shows the geographical locations of Çadır Höyük and Çatalhöyük (the excavation sites where the wild ass remains analysed in this study were recovered), as well as the approximate dispersal areas of extant and extinct ass taxa. We avoid showing distinct distribution ranges, especially for *E. h. hydruntinus*, due to scarcity of hydruntine remains in zooarchaeological assemblages and the difficulty of morphological distinction between hydruntines and other Eurasian wild asses (see also Bennett et al., 2017; Crees & Turvey, 2014). (b) Models summarising two hypotheses regarding the taxonomic position of European wild ass relative to other equid lineages. (c) The locations of three ancient hemione individuals and one hemippe individual the partial genomes of which were previously published (Bennett et al., 2022; Fages et al., 2019). (d) Timeline of dated *E. hydruntinus* reportings with samples reported in this study (adapted from Crees & Turvey, 2014).

2 | MATERIALS AND METHODS

2.1 | Archaeological samples

We studied zooarchaeological skeletal samples that were morphologically identified as belonging to equids, excavated in the archaeological sites of Çatalhöyük and Çadır Höyük in modern-day Turkey. Çatalhöyük is a major Ceramic Neolithic period site in Central Anatolia, but its upper layers have also yielded remains dating to the Bronze and Iron Ages, and later periods (Hordecki, 2020; Pawłowska, 2020). Çadır Höyük has demonstrated continuous occupation from the Middle Chalcolithic to the Byzantine Era (early fifth millennium BCE to 14th century CE) (Ross et al., 2019; Steadman, Hackley, et al., 2019; Steadman, McMahon, et al., 2019). We genetically analysed 11 equid samples from Çatalhöyük and four from Çadır Höyük. See the Appendix S1 for further information on the sites and the archaeological material.

2.2 | Radiocarbon dating

We radiocarbon dated all three equid samples that showed Eurasian wild ass-related genetic profiles (see Section 2.7 below). For each sample, approximately 3 g of bone or tooth material were cut using

non-carbon-based discs attached to a Dremel tool and sent to the TÜBİTAK-MAM AMS facility (Gebze, Turkey) for carbon-14 dating. The dates were calibrated using the R package '*Bchron*' using the IntCal20 curve (Reimer et al., 2020).

2.3 | Ancient DNA isolation and sequencing

All pre-PCR experiments were conducted in the Middle East Technical University (METU) Ancient DNA Clean Room, in a dedicated laboratory for aDNA research, located in a different building from the post-PCR laboratory. Samples were subjected to a standard ancient DNA isolation protocol (Dabney et al., 2013), with minor modifications. In brief, surfaces of archaeological samples were cleaned with damp paper cloth, 100–200 mg of bone or tooth samples were cut using discs attached to a Dremel, were pulverised with mortar and pestle, and transferred into 2 mL screw-top tubes. Each sample's powder was treated with 1 mL extraction buffer (0.45 M EDTA and 0.25 mg/mL Proteinase K) in a 37°C rotating incubator for 18 h. Tubes were centrifuged and supernatants were added to reservoirs containing 13 mL of binding buffer (5 M guanidine hydrochloride, 40% (vol/vol) Isopropanol, 0.05% Tween-20 and 90 mM sodium acetate) to bind DNA fragments to Qiagen spin columns. DNA fragments were washed twice with a washing buffer (Qiagen) and fragments were collected after two consecutive elution steps into 50 μL Qiagen EB buffer. Blunt-end ancient DNA libraries were prepared using the Meyer and Kircher protocol (Meyer & Kircher, 2010) with a single indexing approach. All experiments on the Çadır Höyük and Çatalhöyük samples were conducted on different days with fresh reagents.

Libraries were sequenced on the Illumina HiSeq 4000 platform at SciLife, Stockholm, Sweden. Using 47 thousand–79 million reads (median 20 million) obtained in initial low-coverage sequencing (Table S1), we identified three equid samples, two from Çadır Höyük (cdh008, phalanx; cdh010, calcaneus) and one from Çatalhöyük (chh003, tooth), that displayed a wild ass-related genetic signature on the *Zonkey* pipeline (Schubert et al., 2017) (see Section 2.7). We further sequenced cdh008 three times, and cdh010 and chh003 one more time at the same facility; we thus generated 2501 million, 316 million and 265 million reads in total per sample respectively (Table S2).

2.4 | Data preprocessing

The data were processed following published workflows (Günther et al., 2015; Kılınç et al., 2016); the workflow was applied to all samples screening libraries genetically, and also to the libraries of the three samples (cdh008, cdh010 and chh003) that were deep-sequenced. Raw sequencing reads were demultiplexed, adapter sequences were removed and paired-end sequencing reads were collapsed using *AdapterRemoval* (v. 2.3.1), requiring a minimum of an 11 bp overlap between pairs (Schubert et al., 2016). Collapsed .*fastq* files were aligned to the horse reference genome (EquCab2.0; including autosomes, X chromosome, and the mitochondrial genome) (Wade et al., 2009) using the *aln* module of *BWA* software (v. 0.7.15) (Li & Durbin, 2009) with the parameters '*-n 0.01 -o 2'* and seed disabled (*"-l 16500"*). All alignment files of the same individual were merged using *SAMtools* (v. 1.9) *merge* (Li et al., 2009). PCR-duplicates were removed using *FilterUniqueSAMCons.py* (Meyer & Kircher, 2010); reads with length <35 bp were discarded. Reads with a mismatch to fragment length ratio >0.1 and with mapping quality MAPQ <30 were also removed. The average genome coverage for each sample was calculated using filtered reads with the '*genomeCoverageBed*' tool within *bedtools2* (Quinlan & Hall, 2010).

2.5 | Authentication of genetic data and trimming

We studied the authenticity of the data using *PMDtools* (v. 0.60) (Skoglund et al., 2014). Post-mortem damage (PMD) profiles (Briggs et al., 2007) for each sample were generated with the *PMDtools* '*--deamination*' parameter. Deamination-related cytosine to thymine transitions at 5′-end positions were 0.32, 0.29 and 0.36, while average read lengths were 74, 75 and 67 bp for cdh008, cdh010 and chh003 respectively (Figures S1 and S2, Table S2), supporting authenticity (Pedersen et al., 2014). Reads were trimmed 10 bp from both ends using the '*trimBAM*' command of the *bamUtil* software (Jun et al., 2015) to remove PMD.

2.6 | Complete mitogenome sequences

We first generated a complete mitogenome sequence of our highest coverage sample cdh008 (mtDNA coverage: 140.04×) as follows. The *.fastq* reads were mapped against a linearised reference kiang mitogenome (NC_020433.1) where the first 200 bp were duplicated at the end. The mapped reads were then remapped on the original circular kiang mitogenome using the circular mapper implemented in Geneious Prime 2023.2 (Dotmatics, Boston, MA). Visual inspection of the mapped reads within the Geneious browser revealed a sequence insertion within the Control Region that was absent from the kiang reference, with a distinct mismatch pattern of the reads overlapping both sides of the missing region. To characterise this region, a first consensus cdh008 mitogenome sequence was produced and manually edited to introduce the few starting bases identified on the browser alongside a stretch of Ns in-between. After end duplication, the *.fastq* sequences were remapped on this novel sequence using the local aligner *BWA mem*. The mapped reads were then remapped using the circular mapper on the first circularised cdh008 consensus. After removal of the soft-clipping annotations, the missing bases could be identified and reintroduced into a revised version of the cdh008 consensus, used for a final mapping step using the global aligner *BWA aln*, with a seed length of 18 nucleotides (*–l 18*) followed by remapping with the circular aligner as above. The insertion consisted mainly of repetitions (possibly 14 copies for cdh008) of a 12-bp motif:

RCACCTGTGCAC. We thus obtained a consensus sequence for cdh008 without missing bases or ambiguities and complete mitogenome sequences for cdh010 and chh003 with <5 ambiguities (R) within repeated motifs, which were deposited to the NCBI Genbank (accessions: OP448588, PP101609 and PP101610).

2.7 | Taxonomy and sex determination

We used the *Zonkey* pipeline implemented in *PaleoMix* (v. 1.2.14) (Schubert et al., 2017) for investigating the taxonomic identity of cdh008, cdh010 and chh003, using 158–164 thousand SNPs per sample on the *Zonkey* reference panel. All three samples were clustered together with the *E. h. onager* and *E. kiang* (Asian wild ass) group in PCA and TreeMix, confirming their Eurasian wild ass status. Sex estimation was also performed using *Zonkey*, based on autosomal versus X chromosome coverages. For chh003, X chromosome coverage was approximately half the autosomal coverage, suggesting this sample was male, while cdh008 and cdh010 showed similar autosomal and X chromosomal coverage, suggesting they were female.

2.8 | Phylogenetic analyses of mtDNA data

We used *MEGA X* (Kumar et al., 2018) to construct maximum likelihood (ML) and neighbour-joining (NJ) trees. A 361-bp fragment of the mtDNA D-loop, as well as a 16,790-bp consensus sequence of the whole mitochondria, were used as reference sequences (Genbank). For the ML analysis using 138 published D-loop sequences (Table S3), the substitution model was chosen as Hasegawa–Kishino–Yano (HKY) with a Gamma-distributed rate (no invariant sites) using *jModeltest* (v.0.1.1) (Guindon & Gascuel, 2003; Posada, 2008). A maximum composite likelihood model was used for NJ tree construction. Nodal support was evaluated by 1000 bootstraps. To construct the median-joining (MJ) network, 121 partial D-loop samples were trimmed to 249 bp based on the longest shared fragment. The MJ network was constructed with *NETWORK* v.5 maximum parsimony post-processing (http://www.fluxus-engineering.com).

For analyses of the full mitogenome sequences, in addition to the three hydruntine sequences obtained as described above, we used the following published sequences: a hydruntine from Sicily (*E. h. hydruntinus*) (Catalano et al., 2020), three extinct hemippes (*E. h. hemippus*) (Bennett et al., 2022), as well as five donkeys (*E. a. asinus*), one Somalian wild ass (*E. a. somalicus*), eight kiangs and Mongolian kulans, two Iranian wild asses (*E. h. onager*), one Turkmen kulan (*E. h. kulan*) (described in Bennett et al., 2022), and the reference horse mitogenome as outgroup (total 25 sequences). *MEGA X* estimated the best substitution model as HKY with a Gamma-distributed rate and invariant sites. ML analyses were performed using *PHYML* v2.2.4 (Guindon & Gascuel, 2003) with 100 bootstraps. Tip dating using Bayesian inference was performed with *BEAST* (v. 1.8.4) (Drummond et al., 2012). For the three Anatolian samples produced in this study, as well as one *E. hydruntinus* sample (Catalano et al., 2020) C14 dates were calibrated using Calib (v. 8) with the IntCal20 curve (Reimer et al., 2020; Stuiver & Reimer, 1993), whereas for the three *E. hemionus hemippus* samples, estimations from their corresponding study were used (Bennett et al., 2022). The equid root age was calibrated using a lognormal prior ensuring a median value of 4.3 Mya with a 95% confidence interval (95CI) encompassing the most extreme values of the 95CI identified in (Vilstrup et al., 2013) (lognormal priors: *mean* = '*15.27*' *stdev* = '*0.1*'). An extended Bayesian skyline plot model was used with a lognormal effective population size prior with a median of 6000 and with a wide 95CI (*mean* = '*11.0*' *stdev* = '*1.0*' *offset* = '*3.4*'). A strict molecular clock was used with a clock rate prior to ensure a median value of 2.0E-8 with a 95CI encompassing the most extreme values of the 95CI identified in Vilstrup et al. (2013) (*mean* = '*–17.7*' *stdev* = '*0.7*'). The analysis was run six times for 500,000,000 steps, logging every 50,000 states and discarding the first 10% states as burn-in, resulting in effective sample sizes >9000 for all traces in each independent run, and >50,000 for the combined runs. Robustness was verified by analysing each log and maximum clade credibility tree with median node height using LogAnalyser and TreeAnnotator (Drummond et al., 2012). All six log and tree files after the removal of the 10% burn-in were combined using LogCombiner (Drummond et al., 2012). Convergences were visualised using *Tracer* (v. 1.6) (Rambaut, 2014) and the best tree was drawn using *FigTree* (v. 1.4.3) (Rambaut, 2014). Final posterior values are available in Table S9.

2.9 | SNP panel construction

We constructed a SNP panel using 12 published modern-day genomes (median coverage 10×) (Huang et al., 2015; Jónsson et al., 2014; Renaud et al., 2018; Wang et al., 2020): six domesticated donkeys (*E. a. asinus*) and six Asian wild asses (three *E. h. hemionus*, one *E. h. onager* and two *E. h. kiang*) (Table S3). The data were mapped to the horse reference EquCab2.0 using the *BWA* (v0.7.15) *mem* module with the default parameters. Reads were sorted using the *SAMtools sort* (v.1.9) and PCR duplicates were removed using *Picard MarkDuplicates* (https://broadinstitute.github.io/picard/). A mapping quality filter of <20 was applied using *SAMtools* (v.1.9). Multiple libraries from the same individual were merged with *SAMtools* merge (v.1.9). Two horse genomes were included as outgroup (Der Sarkissian et al., 2015; Orlando et al., 2013) (Table S3). We performed de novo SNP calling using *angsd* (v. 0.934) (Korneliussen et al., 2014) with the *SAMtools* model, implemented as '*-GL 1*' on the donkey and wild ass data separately. We limited de novo SNP calls to positions with depth ≥4, and required non-missing genotypes across all individuals. A total of 31,592,013 and 38,029,882 SNPs were thus called from donkeys and wild asses, respectively, and merged into a union. Limiting these to autosomal positions, applying a minor allele

frequency filter >0.1, a Hardy–Weinberg filter of '1e-30' and keeping only biallelic transversion positions yielded 2,146,416 SNPs, which we refer to as the main SNP panel.

To confirm results from this main SNP panel, we used a published equid variation SNP panel (Bennett et al., 2022), derived from the *Zonkey* panel and established using nine genomes of diverse equids (Schubert et al., 2017) and filtered to eliminate SNPs in repeat regions (Bennett et al., 2022). We further filtered this set to keep positions polymorphic in asinus and hemionus and used only transversions, leaving 2,408,064 SNPs. We refer to this as the secondary SNP panel.

2.10 | Genotyping and dataset construction

We genotyped all individuals using either of the SNP panels. Modern-day individuals were genotyped using *SAMtools (v. 1.9) mpileup* in diploid fashion. Published ancient genomes included three ancient onagers 'Sagzabad', 'TepeHasanlu3459' and 'TepeMehrAli' (Fages et al., 2019), and three ancient hemippes 'GT64', 'Hm_1864' and 'Hm_1892' (Bennett et al., 2022) (note that Hm_1892 is coded NMW1308 in mtDNA analyses); we filtered their reads for mapping quality >30 and genotyped them with our three genomes using *sequenceTools* (v. 1.4.0.5) (github.com/stschiff/sequenceTools) with the option '-randomHaploid'. The *pileupCaller* module of this program was run on ancient samples, genotyping the SNP panel positions in pseudohaploid mode (choosing one random read per SNP). Genotypes of modern-day and ancient individuals were merged using *PLINK* (v. 1.90) (Purcell et al., 2007) and *EIGENSOFT* (v. 7.2.1) (Patterson et al., 2006) to create the final dataset for the two SNP panels (the main and secondary datasets). These were used in all population genetic analyses except CDS-based phylogenetic tree construction and for selection scans (Appendix S1).

2.11 | f_3-statistics

We performed outgroup-f_3 statistics on the autosomal dataset using *popstats* (Skoglund et al., 2015) with '*--not23*' and '*--f3 vanilla*' parameters, using the *PLINK* (v. 1.90) *.tped/.tfam* file format, using the horse as outgroup (Table S3). A heatmap graph was generated from the f_3 results using *heatmap.2* of the R package *gplots* (https://cran.r-project.org/web/packages/gplots/index.html).

2.12 | Multidimensional scaling

A genetic distance matrix was generated by subtracting pairwise outgroup-f_3 values from 1. This matrix was used in multidimensional scaling analysis by running the '*cmdscale*' command in R with parameters '*eig = True*' and '*k = 2*'. The first two dimensions were visualised in R.

2.13 | *D*-statistics

We performed *D*-tests using either the main or secondary datasets, in *EIGENSTRAT* file format (*.geno/.ind/.snp*), with the '*qpDstat*' module of *AdmixTools* (v. 7.0) (Patterson et al., 2012) with default parameters. The analysis was run on all possible trio combinations on the individual level, using horse as an outgroup. We limited the analyses to pairs with sufficient numbers of overlapping SNPs (minimum of 5000 SNPs, e.g. thus excluding Hm_1892). Results with absolute *Z*-score >3 were defined as nominally significant. In several figures (e.g. Figure 4), we use boxplots of individual-based *D*-statistics grouped based on geography (e.g. Gobi/Tibet, or Anatolia); the colours of the boxplots indicate the proportion of individual-based tests with $Z > 3$. We preferred individual-based *D*-tests instead of a priori defining populations, as this approach allows studying consistent behaviour among individual genomes within a group. Results were visualised with *ggplot2* (Wickham, 2016).

2.14 | Autosomal CDS tree

Following (Chen et al., 2021), we created a phylogenetic tree of equid lineages based on autosomal protein coding sequences (CDS) to facilitate alignment. We used CDS annotation from NCBI (Refseq Accession Number GCF_000002305.2) and chose the longer transcript among overlapping genes. CDS of 15 ass/donkey genomes (Table S3) and one horse (outgroup, ENA accession SAMEA3498888) were extracted from *.bam* files by *angsd* (v. 0.934) (Korneliussen et al., 2014) with the base with the highest effective depth ('*-doFasta 3*' option), and requiring a base quality of >30 and mapping quality of >30 ('*-minQ 30 -minMapQ 30*'). Missing nucleotides were removed from resulting fasta files and sequences were re-aligned for each chromosome. A final CDS dataset was constructed by merging all chromosomes and filtering all transitions. CDS selection, removing missing nucleotides, re-aligning, merging and transition filtering were performed with custom Python scripts (filter_gene.py, trim_py, merge_py, and remove_transitions.py respectively) (see Data Availability Statement). An ML tree was generated with this CDS dataset using *RAxML*v. 8.2.12 (Stamatakis, 2014) as follows: a preliminary tree was constructed using the GTRCAT model, 200 bootstrapping replicates were applied on the preliminary tree using the GTRGMMA model, and a final tree was generated by overlaying the calculated bootstraps on the preliminary tree. The final tree was visualised by R packages *ape* v. 5.6.2 (Paradis & Schliep, 2019) and *phytools* v. 1.0.3 (Revell, 2012).

2.15 | Divergence time estimation

In order to estimate the divergence time between the hydruntines and Asian wild assess, we used the F(A|B) statistic (Green

TABLE 1 Genome information and radiocarbon dates of the three newly sequenced ancient Anatolian equids. Genome coverages are provided as average read depth across all positions on the whole genome or the mitochondrial genome, after removing duplicate reads. '5' Damage' and '3' Damage' represent postmortem damage-induced substitution frequencies (C->T and G->A), given as fractions of C or G nucleotides at the 5′ and 3′ ends, respectively. "SNP count" refers to SNPs called among a total of 2,146,416 variable autosomal transversion positions identified among 12 *E.asinus* and *E.heminonus* genomes (Section 2). Radiocarbon (C14) date intervals (95%), calibrated using IntCal20 (Reimer et al., 2020) curve, are given as calibrated years Before Present (BP). The raw radiocarbon dates and laboratory codes are available in Appendix S1: Table S2.

Sample	Site	C14 dates (cal BP)	Genetic sex	Genomic coverage	mtDNA coverage	5' damage	3' damage	mtDNA Haplo- group	Transversion SNP count
cdh008	Çadır Höyük	3004–2779	Female	6.38×	140.04×	0.32	0.31	*H1*	1,879,550
cdh010	Çadır Höyük	2999–2850	Female	0.72×	18.24×	0.29	0.29	*H1*	800,057
chh003	Çatal- höyük	2698–2356	Male	0.57×	41.39×	0.36	0.36	*H1*	654,521

et al., 2010), which calculates the probability of observing a derived allele in individual A (with low genome quality), at the sites where individual B is heterozygous. In our case, individual A corresponds to the hydruntine cdh008, and individual B corresponds to the kiang sample Kia2 (Wang et al., 2020), representing the Asian lineage. In total, $n = 2{,}626{,}614$ heterozygous positions were in Kia2, and alleles were polarised using EquCab2.0. We genotype cdh008 at these positions with *SAMtools* (v. 1.9) *mpileup*, with read and base quality >30. F(A|B) was calculated by choosing a random allele per site. This probability was then computed for different divergence times and effective population sizes (N_e) using the formula $e^{-T/(2Ne)}/3$ (Mualim et al., 2021) or with simulations using *msprime* (v. 1.0) (Baumdicker et al., 2022), with a generation time of 8 years, a mutation rate of 7.242×10^{-9} and a recombination rate of 1 cM/1Mbp (Orlando et al., 2013). We simulated 100 chunks of 1 Mb and calculated the mean F(A|B). Using the formula $H = 4N_e\mu$, we estimated N_e for Kia2 as 62,319.3. Expected F(A|B) values were calculated for divergence times ranging from 100 kya and 2.1 Mya, and for effective population sizes ranging from 5 k to 80 k. We determined the expected F(A|B) values, which span the observed F(A|B) value to find the likely time of population split. We also estimated split times of *E. h. onager* and *E. a. asinus* each from *E. kiang* with F(A|B). We repeated the F(A|B) calculation using only transversions ($n = 811{,}814$) to limit the chance of homoplasy, which resulted in a highly similar estimate (0.149, compared to 0.145 using all SNPs). The scripts are shared in the "Data Availability Statement" section.

2.16 | Runs of homozygosity

ROH can be called efficiently from >5× genomes (Ceballos et al., 2021). We estimated ROH for cdh008 and published equid ass/donkey genomes (Table S3), downsampling all *.bam*files to 6.38× to match the lowest coverage case, cdh008. We calculated ROH with *PLINK* (v. 1.90) with the parameters '*--chr-set 34 --homozyg -homozyg-snp 30 -homozyg-kb 300 -homozyg-density 30 -homozyg-window-snp 30 -homozyg-gap 1000 -homozyg-window-het 3 -homozyg-window-missing 5 -homozyg-window-threshold 0.05*'. The resulting data were merged into a single file and visualised in R.

2.17 | Genome-wide heterozygosity estimation

We downsampled *.bam* files from published equid ass/donkey genomes by downsampling these to 6.38× coverage, that is, the coverage of the Anatolian cdh008. Using these and cdh008, we called de novo variants in each genome at all positions with 5–13× depth using *SAMtools* (v. 1.9), filtered these sites to only keep transversions (to avoid the PMD effect), and determined the proportion of heterozygous loci across the transversion.

2.18 | Masking against reference bias

Ancient genomes and/or genomes mapped to divergent reference genomes can be subject to reference biases (Günther & Nettelblad, 2019), while masking by changing polymorphic sites to N can mitigate this bias (Koptekin et al., 2023). We accordingly masked EquCab2.0 at 2,146,416 polymorphic sites in the main SNP dataset, repeated alignment, and estimated heterozygosity among a selected set of individuals (Sp-5, Kia2, Ona and cdh008). This procedure does not involve masking hydruntine-specific derived positions, but this is not expected to account for the stark heterozygosity differences we observe among lineages.

3 | RESULTS

We extracted ancient DNA from 15 equid tooth and bone samples excavated in two Central Anatolian sites, 11 from Çatalhöyük and four from Çadır Höyük (Figure 1c). We generated genome-wide data from these using shotgun sequencing (Table S1) and obtained 0.02%–11.35% (median = 0.07%) of endogenous DNA by mapping to the horse reference genome (EquCab2.0). We chose three samples (cdh008, cdh010 and chh003) with >5% endogenous DNA and authenticity signals (Section 2) for further sequencing (Table 1, Figure S1). We thus produced three genomes with nuclear coverages of 6.38×, 0.72× and 0.57× (Table 1, Table S2). Radiocarbon dating of the skeletal material placed all three individuals within the early/mid-first millennium (Iron Age) in Anatolia (Table 1, Figure 1d). We also constructed mitogenome sequences of these three individuals (Section 2).

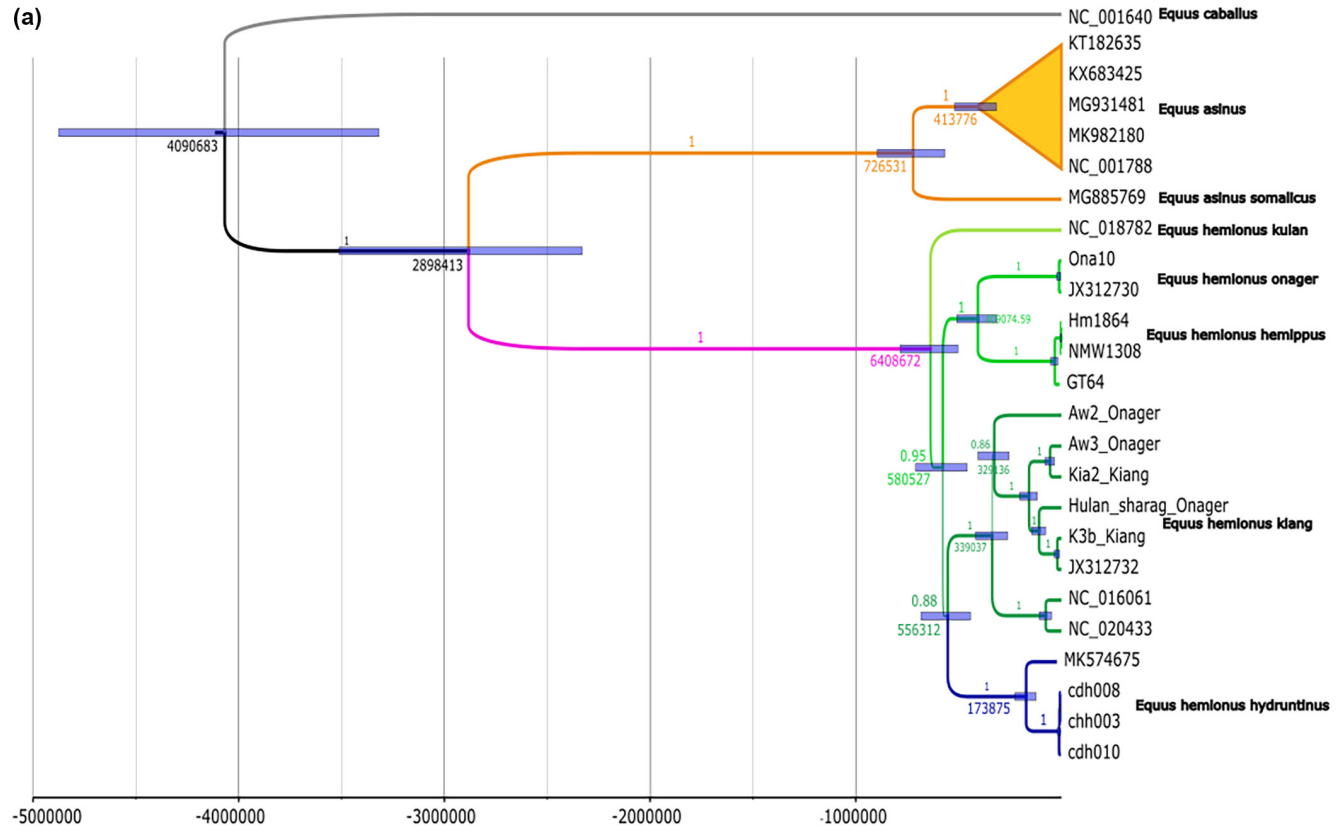
3.1 | Genetic analyses assign the iron age Anatolian wild ass individuals to hydruntines

During the Bronze and Iron Ages, horses and the domestic donkey were common in Anatolia, while hydruntines and hemiones were also present, the latter restricted to Southeast Turkey (Table S1 in Appendix S1 see references therein). Due to a lack of sufficient diagnostic characteristics, the three specimens were identified osteologically only to the level of the genus *Equus* (Table S2 in Appendix S1). As a first step towards characterising the three Anatolian individuals, we used the *Zonkey* pipeline (Schubert et al., 2017), which classified all three individuals as 'Asian wild ass-related' (Table S1).

We then compared mtDNA data from these three Anatolian equids with published partial mtDNA sequences from 82 hemiones (Bennett et al., 2017; Huang et al., 2015; Orlando et al., 2009; Wang et al., 2020), 21 morphologically-identified hydruntine individuals (Bennett et al., 2017; Catalano et al., 2020; Orlando et al., 2009; Schubert et al., 2016) and seven other equids (Bennett et al., 2017; Jónsson et al., 2014; Wang et al., 2020). We used a 249-bp-long fragment shared across all samples to construct a haplogroup network; we also used a 361 bp-long D-loop fragment for *BEAST* analysis (Table S3). In both analyses, mtDNA sequences of the three Anatolian equids clustered with published hemione–hydruntine sequences with 100% support (Figures S3 and S4). Moreover, the three new sequences were a sub-branch of the *H1* haplotype clade (as defined by Bennett et al. (2017)). *H1* is the most prevalent mtDNA haplotype among the 19 hydruntine individuals in this dataset, appears exclusive to hydruntine, and was already detected among Anatolian hydruntines from c.7850–2500 BCE (Bennett et al., 2017; Guimaraes et al., 2020) (Figure S4). Together, these results indicate that the mitochondrial lineage of the three Anatolian equids belonged to the wild asses of Eurasia, and was more closely related to hydruntines than to hemiones.

We investigated this further using complete mtDNA sequences, taking advantage of a recently published near complete mtDNA sequence of a morphologically identified *E. h. hydruntinus/E. hydruntinus* specimen from Sicily (Catalano et al., 2020), along with full mtDNA sequences from nine ancient and modern-day equids (Section 2, Table S3). Again, in both the Bayesian and ML trees (Figure 2a, Figure S5) the Anatolian wild asses clustered with the published hydruntine with full support. The mitochondrial phylogenies further suggested that the Eurasian wild asses radiated into four groups within a relatively narrow time range: (a) the hydruntine group encompassing European and Anatolian individuals, (b) the Iranian onagers and the Syrian hemippes, (c) the Gobi kulans and the Tibetan kiangs and (d) the Central Asian kulans represented by a single individual. The order of these radiation events is unclear, however, and is estimated by *BEAST* analyses to have occurred within a relatively short time frame of 300,000 years, between ~0.8 and 0.5 Mya.

The above analyses suggest that the three Anatolian wild asses studied here belonged to the hydruntine lineage. However, mtDNA phylogenies may be incongruent with the overall phylogenetic history of a species due to drift, selection or incomplete lineage sorting (e.g. Toews & Brelsford, 2012). We hence asked whether the whole genome data also support hydruntine ancestry for the three Anatolian wild asses. We compiled a list of ~2 million autosomal transversion SNPs using published asinus and hemionus genomes (Section 2.9). We genotyped the three Anatolian wild ass genomes, 13 published modern-day ass/donkey genomes, as well as three ancient onagers from Iran and three ancient hemippes from the Middle East (Table 1, Figure 1c, Table S3). We then constructed an MDS plot of genomic diversity among these genomes with the outgroup f_3-statistic (Section 2) (Figure 3, Figure S6, Table S4). This revealed a unique position for the three Anatolian wild asses, distinct from the two other Asiatic wild ass groups included (Gobi kulans/kiangs, and the onagers/hemippes). Because the hydruntine is the only other wild ass lineage identified in the fossil record to have lived in Holocene Anatolia, together with the mtDNA evidence, our results strongly suggest that the three Anatolian wild asses belonged to the extinct hydruntine lineage.



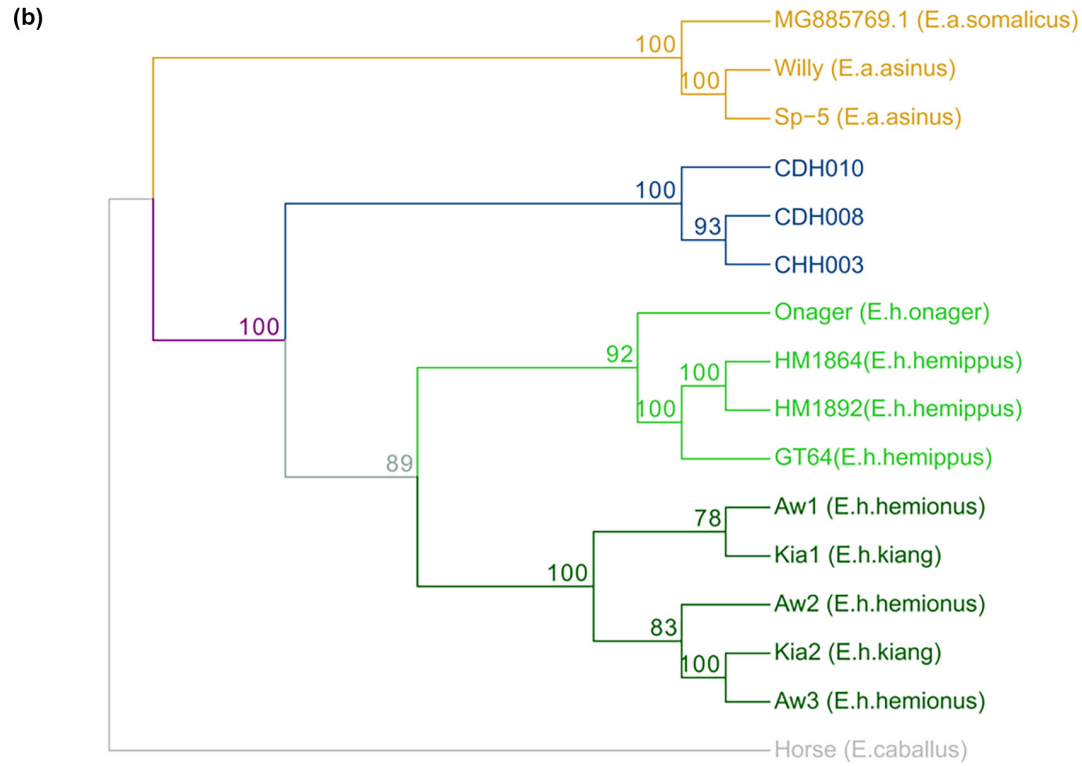


FIGURE 2 (a) A Bayesian phylogenetic tree constructed using a 16,790-bp long consensus sequence of the whole mitochondrial DNA among various equid lineages. Numbers on the internal branches show posterior probability support in percentages and the numbers on the nodes show the estimations for divergence times in years. The purple node bars show the 95% confidence intervals for divergence times. The colour codes correspond to those in Figure 1. (b) A maximum likelihood tree was constructed using concatenated protein coding sequences (29,384,180 bp) across the 15 genomes using only transversion sites. Numbers on the internal nodes show bootstrap support. External nodes indicate the genome/individual IDs, with species names in parentheses. A version showing branch lengths can be found in Figure S9.

3.2 | Whole genome analyses support hemiones and hydruntines as sister clades

Because the hydruntine has yet been genetically described only through mtDNA, we leveraged upon the new Anatolian genomes to investigate evolutionary relationships among asses. Based on partial mtDNA sequences, it had been previously suggested that hydruntines were most closely related to hemiones, to the exclusion of other equids (Bennett et al., 2017; Catalano et al., 2020; Orlando et al., 2006, 2009). Our results above are fully consistent with this hypothesis, since the Anatolian and modern-day Asiatic wild ass individuals cluster together in the mtDNA analyses to the exclusion of African asses (Figure 2, Figures S3 and S4). Anatolian wild asses are also closer to hemiones in the MDS- and heatmap-based summaries of nuclear genomic variation using outgroup-f_3 scores (Figure 3; Figures S6 and S7, Table S4). We further tested this pattern with *D*-statistics, employing the c.2 million autosomal transversion SNP panel (Section 2) (Table S5). *D*-tests of form *D(Horse, Anatolia; Africa, Non-Anatolia)* and *D(Horse, Non-Anatolia; Africa, Anatolia)*, where *Non-Anatolia* refers to any Asian hemione, kiang, onager and hemippe individual, separately in each comparison, were all significantly positive ($Z > 3$; Section 2), in line with the notion that Asian and the hydruntines represent sister taxa (Figure 4; Table S5). We replicated the same results using a secondary dataset of c.2 million transversions (Schubert et al., 2017) (Section 2) (Figure S8, Table S6).

3.3 | Hemiones and hydruntines are reciprocally monophyletic

Earlier analyses of partial mtDNA sequences had suggested that the Asian hemione versus hydruntine divide could be artificial, with some hemiones appearing closer to hydruntines than to other hemiones (Bennett et al., 2017; Orlando et al., 2009). Meanwhile, the identification of hydruntines as one of at least three hemione subgroups in MDS analyses (Figure 3, Figure S6) prompted us to evaluate the relationships between these three subgroups. We tested this question further through a number of approaches.

We first constructed an ML tree using whole genome data of 13 equids and using concatenated protein-coding sequences, limited to transversions (Section 2). This indicated reciprocal monophyly between the Asiatic wild asses (9 genomes representing the kiang, hemione, onager and hemippe) and the hydruntine clade (Anatolian wild asses) (Figure 2b, Figure S9; compare with Figure 2a and Figure S5). The Asian wild asses themselves further split into two groups, the Middle Eastern hemippe and onager, and the Gobi/Tibet cluster comprising the kiang and kulan.

We further used *D*-tests to clarify relationships between Asian and Anatolian wild asses. We tested monophyly of the Anatolian wild ass genomes with *D(Horse, AnatoliaX; AnatoliaY, Non-Anatolia)*, where *AnatoliaX* and *AnatoliaY* refer to different Anatolian wild asses, and *Non-Anatolia* refers to any individual hemione, kiang, onager or hemippe (Section 2, Table S5). The three Anatolian wild asses formed their own clade to the exclusion of all Asiatic wild asses, with all 60 relevant comparisons being significant in this direction ($Z > 3$; Section 2) (Figure 4, Table S5). We also conducted reciprocal tests of the form *D(Horse, Non-AnatoliaZ; AnatoliaX, Non-AnatoliaW)*, where *AnatoliaX* refers to any Anatolian wild ass genome, while *Non-AnatoliaZ* and *Non-AnatoliaW* refer to genomes of different hemione, kiang, onager or hemippe individuals. In all comparisons, each hemione individual (modern-day or ancient) showed higher affinity to other hemiones over Anatolian individuals ($Z > 3$) (Figure 4, Table S5). The evidence thus supports reciprocal monophyly between the Anatolian and Asian wild asses.

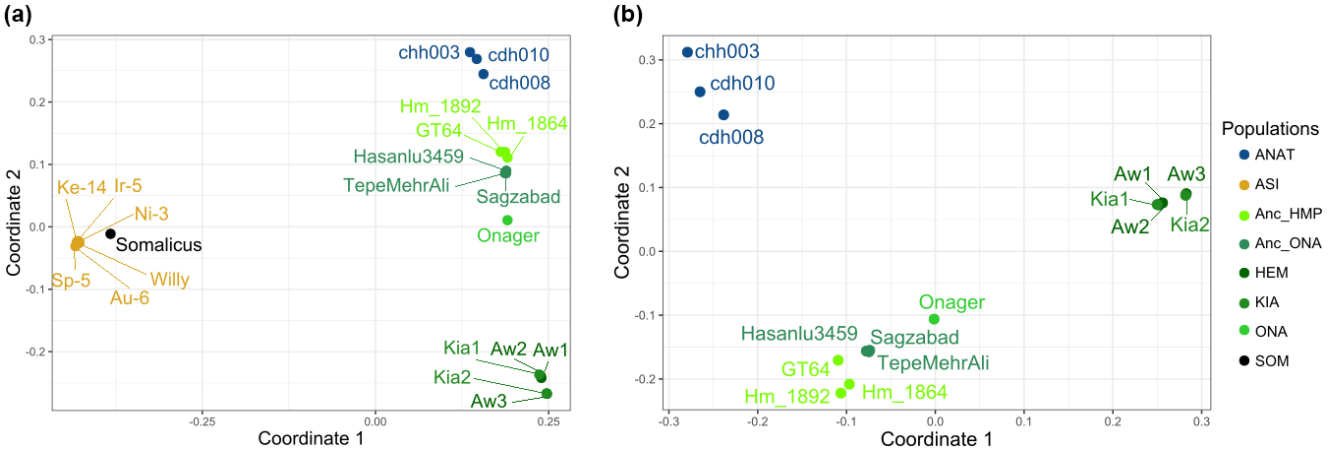


FIGURE 3 Multidimensional scaling plot of (a) donkey and wild asses and (b) only wild asses generated using a distance matrix based on (1-outgroup f_3) statistics calculated from the autosomal variation dataset consisting of 2,146,461 sites. ANAT: Anatolian samples reported in this article; Anc_HMP: Ancient *E. h. hemippus*; Anc_ONA: Ancient *E. h. onager*; ASI: *E. asinus*; HEM: *E. hemionus*; KIA: *E. h. kiang*; ONA: *E. h. onager*; SOM: *E. somalicus*.

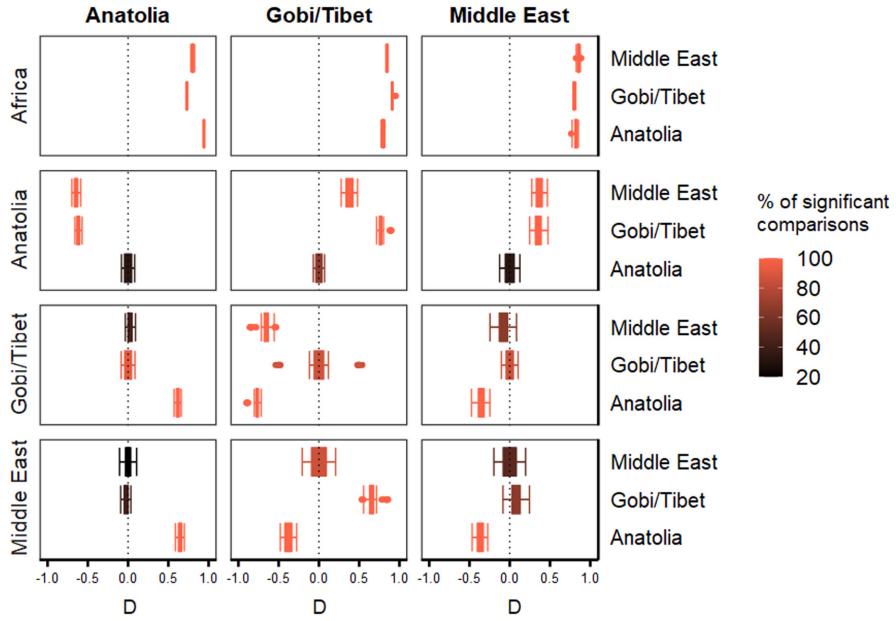


FIGURE 4 Boxplots showing *D*-statistics calculated between wild ass genomes from different regional groups, using autosomal SNPs from the main dataset. The statistics are calculated as *D(Horse, Test; Pop1, Pop2)*, with *Test* shown on the top, and *Pop1* and *Pop2* on the left and right, respectively. All tests are performed using individual genomes chosen from regional groups. Anatolia: cdh008, cdh010 and chh003; Gobi/Tibet: Aw1, Aw2, Aw3, Kia1 and Kia2; Middle East: GT64 and Hm_1864, Onager, Hasanlu3459, Sagzabad; Africa: Asinus, Somalicus. Each boxplot thus shows a set of *D*-statistics involving the regional populations shown. The colour gradient from pink to black represents the fraction of *D*-tests in that comparison that are nominally significant ($Z > 3$). For instance, *D(Horse, Anatolia; Middle East, Middle East)* involves 30 unique tests, only 20% of which reach $Z > 3$. In comparison, *D(Horse, Anatolia; Anatolia, Middle East)* also involves 30 unique tests, only 100% of which reach $Z > 3$, marking that the three Anatolians cluster genetically. See Table S5 for the *D*-test results. See Figure S8 for the same figure drawn using the secondary dataset.

3.4 | Evidence for gene flow between hydruntines and Middle Eastern wild asses

We next asked whether the Anatolian wild asses might be symmetrically related to all Asian wild ass lineages. Following our earlier observations (Figures 2 and 3), here we divided Asiatic wild asses into Gobi/Tibet and Middle East clusters. We then performed *D*-tests of form *D(Horse, Anatolia; MiddleEast, Gobi/Tibet)*, where *MiddleEast* represents Iranian onager and Syrian hemippe individuals while *Gobi/Tibet* consists of Gobi kulans and Tibetan kiangs. We found affinities of all three Anatolian wild asses towards the modern-day Iranian onager and a museum specimen of hemippe, Hm_1864 over other Asian wild asses, including other onagers and hemippes (27 of 30 comparisons with $Z > 0$; 23 of 30 comparisons with $Z > 3$) (Figure 4, Figure S10, Table S5). This observation is even more pronounced (30 of 30 comparisons with $Z > 3$) in comparisons using the secondary dataset (Figure S11, Table S6). This would be compatible with gene flow between Middle Eastern hemiones and hydruntine populations in Southwest Asia.

We investigated this further by testing *D(Horse, Anatolia; MiddleEastX, MiddleEastY)* where we compared the genetic affinity of Anatolian hydruntines between the onager and/or hemippe individuals. Using the main variation panel we observed slight affinity towards the 19th century hemippe over the other hemippe and onagers (12 of 12 with $Z > 0$, 5 of 12 with $Z > 3$; Figure S12). Using the secondary dataset we found a higher affinity to the modern-day onager over all other Middle East wild ass genomes (12 of 12 with $Z > 3$; Figure S13). Although we cannot yet fully determine the exact nature of the possible admixture, we hypothesise that gene flow from hydruntine to the onager and hemippe lineages could explain the observed patterns (see Section 4).

3.5 | The timing of the hydruntine and Asian wild ass split

The above analyses suggest that hydruntines split from other Eurasian wild asses relatively early in their history, although they might have admixed with Middle Eastern wild asses lineages in more recent times as reflected in the autosomal data. Hence we can estimate the hydruntine-Asiatic wild ass split time using mtDNA and also autosomal comparisons between hydruntines and Gobi/Tibet wild assess. As discussed above, the mtDNA data indicated a split involving the various Eurasiatic lineages between 0.8–0.5 Mya. We further used the F(A|B) method for estimating the split time between hydruntines and kiang from the Gobi/Tibet region using autosomal DNA, using both a theoretical estimate and also simulations (Section 2). The intersection of the observed and expected statistics

suggests a most likely population split between 0.6 and 1.2 Mya (Figure 5). This overlaps the mtDNA split time confidence interval between 0.6 and 0.8 Mya.

Given an early split and distinct phenotypes, we tested for possible hydruntine-specific coding sequence changes that could indicate adaptations unique to this lineage (see Appendix S1: '*Selection and Relaxation Scans on Coding Sequences*'). We first used the *PAML codeml* framework (Yang, 2007), but did not identify hydruntine-specific protein coding changes that reached genome-wide significance (Appendix S1, Table S7). Second, we devised an alternative approach, the Pairwise McDonald–Kreitman statistic, where we combined the McDonald–Kreitman test (Stoletzki & Eyre-Walker, 2011) with the population branch statistic (Yi et al., 2010). This approach identified 155 genes that deviated significantly from the rest of the coding genes in terms of excess differentiation on the hydruntine lineage ($p < .01$ after Benjamini–Hochberg correction). However, this gene set did not show functional enrichment relative to the rest of the genome (Appendix S1, Table S8, Figure S15).

3.6 | Severely depleted genetic diversity in the Anatolian wild ass signals population decline

To our knowledge, these three Anatolian individuals are the latest directly dated hydruntines known, if not the latest recorded hydruntines to date (Crees & Turvey, 2014; Guimaraes et al., 2020). Zooarchaeological evidence has suggested that by the 3rd millennium BCE the European population had already gone extinct, while the population was possibly dwindling or lost in Southwest Asia, including in Anatolia, Iran and the Caucasus (Boulbes & van Asperen, 2019; Crees & Turvey, 2014). We thus asked whether we might detect signatures of population decline in our genetic data.

We first estimated ROH, where short ROH can indicate past inbreeding caused by low population size (Ceballos et al., 2018). We called ROH in the highest coverage Anatolian wild ass genome (cdh008) and compared these ROH estimates with those from genomes of other asses down-sampled to the same coverage (Section 2). African asses had high ROH loads (Figure 6a), as expected, since these populations are known to have experienced recent bottlenecks (Jónsson et al., 2014; Renaud et al., 2018). Kiangs, kulans and onagers carried relatively low ROH loads, implying higher diversity than their African counterparts. In turn, cdh008 carried a high number of short runs with complementing long runs, highly similar to the African genomes. The fraction of regions within ROH (FROH) is also high in cdh008 (0.011%) as opposed to other wild asses (0.001%–0.008%, median: 0.004%) save Somali wild ass (0.021%).

We investigated this further by estimating heterozygosity from de novo-called transversion variants across the same genomes. The hemiones had the highest heterozygosity, followed by donkeys and the Somali wild ass. The cdh008 genome harboured the lowest heterozygosity (0.021), even below the African group known to have undergone several bottlenecks (Figure 6b). We confirmed this result was not an artefact of reference bias in cdh008 genotypes by replicating the result using a reference genome masked at polymorphic sites (Section 2) (Figure S14). Combined with ROH analysis, these observations suggest that the Anatolian population was in strong decline already by the first millennium BCE, which is also supported by its rare occurrence in the archaeological record.

Finally, we asked whether a small effective population size, as indicated by the high ROH load, may have led to elevated non-synonymous mutation load in the hydruntine lineage due to relaxed selection (Rogers & Slatkin, 2017). We compared *PAML*-estimated *dN/dS* ratios between cdh008, Asian, and African genomes (Section 2), using only genes under purifying selection ($dN/dS < 1$). This revealed slightly but significantly lower values in cdh008 than the other two genomes (Figure S16).

4 | DISCUSSION

Our analyses of palaeogenomic data from three Anatolian wild asses radiocarbon dated to the first millennium BCE showed that all three carried a putative hydruntine-specific mitochondrial haplotype (*H1*; Bennett et al., 2017), their mtDNA sequences clustered with those of morphologically-identified hydruntine (*E. hydruntinus* or *E. hemionus hydruntinus*) sequences with full support, and all three genome profiles fell outside the diversity of known African or Asiatic ass genomes. Together, the evidence strongly indicates that the three Anatolian wild ass individuals belonged to the extinct hydruntine *E. hydruntinus/E. h. hydruntinus*, no genomic data of which has been published hitherto.

Although it has been termed the 'European wild ass', according to archaeozoological data the hydruntine range included Southwest Asia and possibly even North Africa (reviewed in Boulbes & van Asperen, 2019). In Anatolia, hydruntines were not uncommon in zooarchaeological assemblages until the Bronze Age (Arbuckle, 2013; Arbuckle & Öztan, 2018; Bennett et al., 2017), although their presence largely depends on the archaeological site and subsistence strategies (Martin & Russell, 2012). At Neolithic Çatalhöyük, for example, small and medium-sized equids (*E. hydruntinus/hemionus*) are evidenced by bones in daily and social contexts (Pawłowska, 2020) and also in art, although relatively rare; for instance, they appear in one set of wall paintings, while phalanges were used as raw material for making idols, one of which is known as the first bone figurine from the site (Martin & Russell, 2012; Pawłowska & Barański, 2020).

Until now, the last hydruntine detected on a genetic basis in Anatolia was dated to ~2200 BCE (Guimaraes et al., 2020). In Iran, osteologically identified specimens were dated to the second millennium BCE (Mashkour et al., 1999). Our three Anatolian wild asses, dated to early/middle first millennium BCE, are thus the last hydruntines yet identified.

Our study reveals a number of novel insights into the phylogenic history of hydruntines, represented by the three Anatolian genomes. First, using genomic evidence we confirm that Asiatic hemiones and

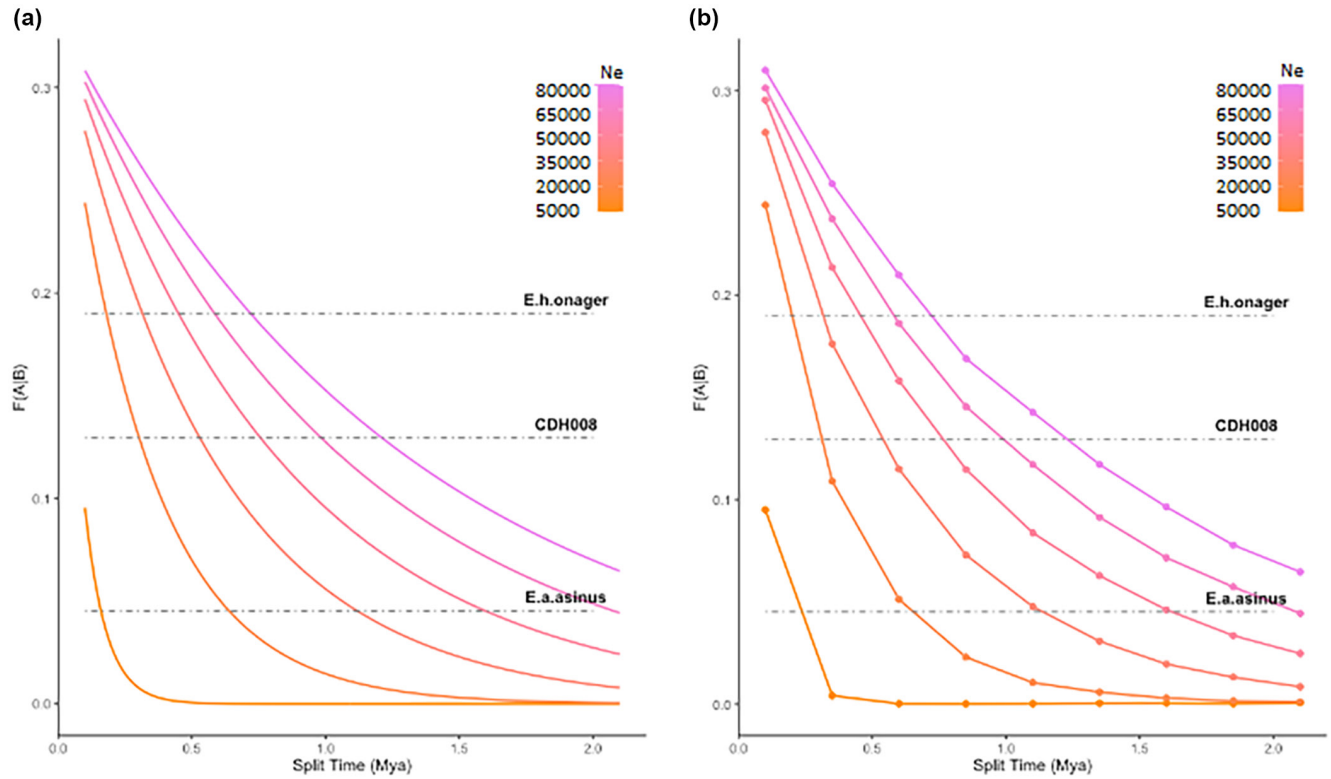


FIGURE 5 Population split time estimation of Anatolian wild ass (cdh008), *E. a. asinus* and *E. h. onager* from Asian wild ass based on the F(A|B) statistic. F(A|B) is the probability of observing a derived allele in the Anatolian wild ass, *E. a. asinus* or *E. h. onager* at the positions where the Asian wild ass genome is heterozygous (Section 2). Observed and expected F(A|B) values, calculated using the formula $e^{-T/(2N)/3}$ in panel (a), and using population genetic simulations in panel (b). We explored the space of divergence times between 100 kya to 2 Mya, and population sizes ranging from 5 k to 80 k.

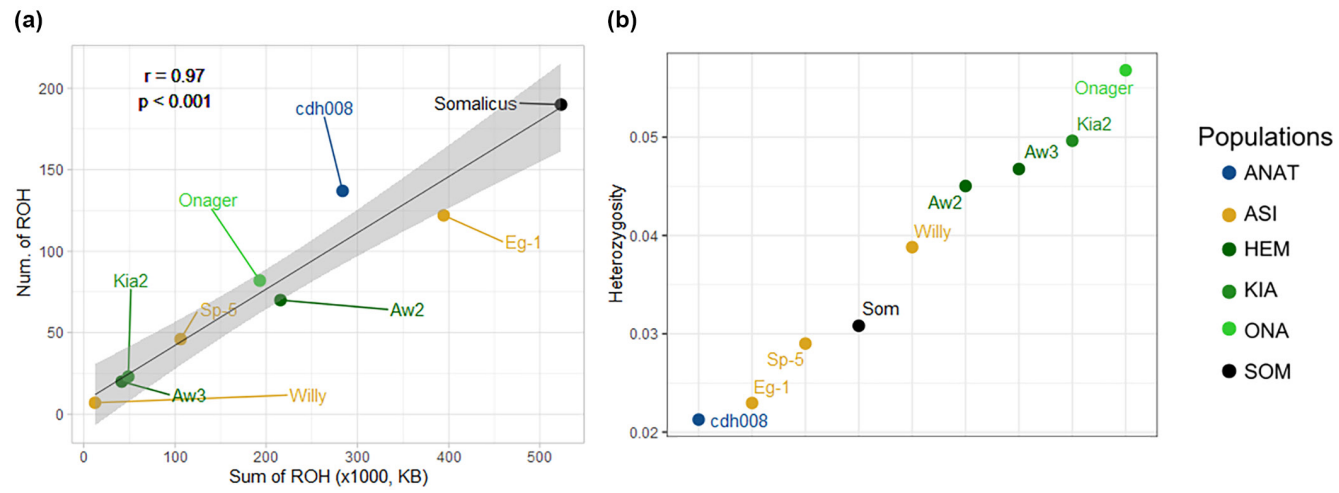


FIGURE 6 (a) Total number versus the total length (sum) of runs of homozygosity (ROH) tracks over 1.5 Mb. The Spearman correlation coefficient and *p*-value are shown in the inset. (b) Heterozygosity calculated across 2,146,416 transversion sites.

hydruntines were sister taxa to the exclusion of African asses, as suggested earlier based on osteological data (Burke et al., 2003; Eisenmann & Mashkour, 1999; Orlando et al., 2006) and mtDNA analyses (Bennett et al., 2017; Catalano et al., 2020; Orlando et al., 2006, 2009). Thus, we may speak of Late Pleistocene and Holocene Eurasiatic wild asses as a broadly coherent evolutionary group.

Our mitogenome phylogeny places hydruntines within Asiatic hemione diversity, splitting after the Central Asian kulan, although with limited statistical support (Figure 1a). In contrast, our nuclear

genome phylogeny indicates hydruntines were the earliest diverging branch of Eurasiatic wild asses, with full support (Figure 1b). The difference highlights difficulties in resolving epitomies even with full mitogenomes, likely due to the unique and non-recombining nature of this marker. Meanwhile, both mtDNA and nuclear genome analyses converge in, suggesting that hydruntines split from Asiatic hemiones within a time frame of 0.5–1 Mya, most likely between 0.6 and 0.8 Mya. This would be compatible with the first fossil records of *E. hydruntinus* in Europe c.600 kya (Boulbes & van Asperen, 2019).

Our mitochondrial and nuclear analyses also indicate the radiation of at least three different lineages within a relatively short time frame: one encompassing the hydruntines, one encompassing the Mongolian kulans and the Tibetan kiangs, and another including the Iranian onagers and Syrian hemippes. This suggests that geographical separation may have been the main driving force behind the divergence of these lineages. It was proposed previously that the European hydruntine may have split from the Central Asiatic wild ass populations and evolved independently in Europe before colonising Anatolia in the Late Pleistocene/Early Holocene (Bennett et al., 2017). In our *D*-tests, hydruntines indeed appear distinct from all other Asiatic wild asses tested, including hemiones from Iran and Syria (as well as those in the Caucasus; Bennett et al., 2017).

This supports the notion that the hydruntine lineage initially evolved in isolation from Middle Eastern hemiones, in Europe, where it evolved its unique phenotypes such as a short muzzle, and may thus deserve an independent taxonomic status from other hemiones, as previously proposed (Boulbes & van Asperen, 2019; van Asperen, 2012). However, because (i) our genome dataset does not include the full diversity of hemiones (e.g. the Indian khur and Central Asiatic wild ass genomes are not yet sequenced), (ii) we lack the information about full hydruntine diversity through ages and its geographic range and (iii) our Anatolian samples represent a population that appears to have undergone a severe bottleneck, the relationships between hydruntines and Asiatic hemiones remains an open question.

We do not know when the hydruntine expanded into Anatolia from Europe. The emergence of a recent sympatric zone may account for the traces of gene flow we detect between hydruntines and Middle Eastern wild asses. Such gene flow can be inferred from the asymmetric relationships between hydruntines and Middle Eastern vs. Gobi/Tibet hemiones, with hydruntines being on average closer to the former (Figures S10 and S11, Tables S5 and S6). More interestingly, the hydruntine has a higher affinity to the modern onager and the 19th century hemippe genomes than to Iron Age onagers or a Neolithic hemippe respectively (although the result is variable between the SNP panels) (Figures S12 and S13, Tables S5 and S6). These observations imply that gene flow was directional, occurring from hydruntines into Middle Eastern hemiones in recent times. This is because gene flow from the onagers into hydruntines would not lead to asymmetric relationships with more recent onagers and hemippes. Alternatively, population structure within both onagers and hemippes could be an explanation, albeit being less parsimonious, as it would require the same type of structure independently maintained in both onagers and hemippes.

Gene flow between hydruntines and hemiones may not be surprising, given evidence for the widespread introgression among equids (Jónsson et al., 2014), and the possibility that hydruntines may have lived in sympatry with these Middle Eastern hemiones in Southwest Asia during the Holocene (Crees & Turvey, 2014; Eisenmann & Mashkour, 1999; Orlando et al., 2006). Putative gene flow between hydruntines and Middle Eastern hemiones could be considered a reason for a taxonomic grouping of hydruntines within *E. hemionus*. At the same time, the limited extent of these admixtures may suggest that these populations were distinct enough to not fully interbreed, supporting the idea that the hydruntine had already experienced a large portion of the speciation process separating them from Middle Eastern wild asses. Similar examples of partial admixture between sister taxa have been reported for various groups of organisms (Burbrink et al., 2022; Jónsson et al., 2014).

Our results also provide hints on possible extinction dynamics of the hydruntine in Southwest Asia. Our ROH and heterozygosity analyses on the first millennium BCE genome from Çadır Höyük suggest low diversity, on a par with the critically endangered African wild ass (Moehlman & Kebede, 2014). Our analysis is admittedly limited by being based on a single genome of <10× coverage; nevertheless, excess of short ROH is a particular signature created by small long-term effective population size (Ceballos et al., 2018) and our results may thus be considered informative about population-level dynamics. Low diversity could have increased the risk of extinction (Newman & Pilson, 1997). Low effective population size itself could have various non-exclusive causes, including habitat fragmentation as a result of human activity and/or climatic events, or human predation (Bennett et al., 2017; Boulbes & van Asperen, 2019; Cai et al., 2021; Crees & Turvey, 2014; Spassov & Iliev, 2002). It is notable that the remains of the three individuals studied here were recovered from residential or midden contexts (Table S2 in Appendix S1) and were likely consumed by humans. Meanwhile, the lack of an elevated *dN*/*dS* signal in the cdh008 genome is notable, given high levels of damaging variant loads in populations subject to strong bottlenecks (e.g. Bosse et al., 2019; Robinson et al., 2019; Rogers & Slatkin, 2017). One possibility is that the presumed reduction in the hydruntine population size may have been too recent and rapid to visibly impact functional variation. The genomic analysis of additional hydruntine individuals from the Holocene may clarify the exact dynamics of its extinction.

A final question remaining is how much longer hydruntine populations survived in Southwest Asia, beyond the individuals we sampled from Anatolia. The trend of higher affinity of hydruntines to the modern-day than to Iron Age onagers implies that hydruntines may have admixed with onagers after the Iron Age (Figure S12). Interestingly, classical authors such as the Roman historian Strabo (*Geography*, Book XII), Varro (*De Re Rustica*, Book II), and Pliny the Elder (*Natural History*) mention the presence of 'wild asses' (*onagri*)

in the vicinity of ancient Late Tatta (modern-day Tuz Gölü in Central Anatolia), and also in the Anatolian dry grasslands of Lycaonia, Garsauira and Bagadania. Therefore, it seems possible that these mentions may be referencing hydruntine populations surviving on the central Anatolian plateau (in the Konya-Ereğli plain, Cappadocia, and/or Kayseri) into the first millennium AD. Our results underscore the need for further archaeogenomics-based taxonomic assignment of non-caballine equid material from Southwest Asia. This would help fully fathom the extinction dynamics of the hydruntine.

AUTHOR CONTRIBUTIONS

M.Ö., F.Ö. and M.S. conceived and designed the study and experiments. K.P., I.A., S.E.A., B.S.A., S.R.S. and G.M. prepared and provided zooarchaeological material. M.Ö. performed molecular biology laboratory experiments with support from A.A. and F.Ö., M.Ö., T.G., K.G., E.Y., K.B.V., G.A., F.R.F., E.S., E.N.A. and D.K. analysed data under the supervision of F.Ö., C.C.B., E.M.G., T.G., A.G., and M.S., Y.S.E., İ.T.. M.Ö., T.G., F.Ö. and M.S. wrote the manuscript with contributions from K.G., E.M.G., E.Y., G.A., F.R.F., E.S., I.H., B.S.A. All authors read and approved the manuscript.

ACKNOWLEDGEMENTS

The authors would like to thank Dr Derya Baykara, for helpful comments; Dr Mikolaj Lisowski and Dr Jesse Wolfhagen for their contributions to sample selection; the Turkish Ministry of Culture and Tourism, Yozgat Museum and Konya Museum for permissions to work on archaeological samples; Sevgi Yorulmaz, Dr Duygu Deniz Kazancı and Dr Aslıhan Ilgaz for their assistance, and all METU CompEvo research team members for discussion and support.

FUNDING INFORMATION

This study was funded by a Scientific and Technological Research Council of Türkiye (TÜBİTAK) 1001 Grant Programme (no. 117Z991 to F.Ö.), the European Commission Horizon 2020 TWINNING Programme (no. 952317 "NEOMATRIX" to M.S.) and European Research Council Consolidator Grant H2020 ERC (no. 772390 "NEOGENE" to M.S.).

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no conflicts of interest.

DATA AVAILABILITY STATEMENT

Genetic data generated from 15 Anatolian ancient equid samples were deposited to the European Nucleotide Archive ENA under project code PRJEB52847 as '*.fastq*' files. The mitochondrial DNA sequences were deposited to the NCBI Genbank with accession no: OP448588, PP101609 and PP101610. The scripts incorporated into phylogenetic and demographic analyses were uploaded to https://github.com/CompEvoMetu/Hydruntine2023. The scripts for forming sequences for three lineages and running PAML analysis are available in https://github.com/quetzalekin/equus. The scripts related to selection analysis, namely, used for obtaining the polymorphism in the representative gene set, counting D_N, D_S, P_N and P_S and calculating PMK can be accessed at https://github.com/rabiafidan/equus.

BENEFIT-SHARING STATEMENT

Benefits from this research accrue from the sharing of our data and results on public databases as described above.

ORCID

Mustafa Özkan https://orcid.org/0000-0002-7520-4572
Kanat Gürün https://orcid.org/0000-0002-0433-2593
Eren Yüncü https://orcid.org/0000-0002-8194-0277
Kıvılcım Başak Vural https://orcid.org/0000-0003-3964-3065
Gözde Atağ https://orcid.org/0000-0001-6173-3126
Ali Akbaba https://orcid.org/0000-0001-6755-6546
Fatma Rabia Fidan https://orcid.org/0000-0001-5877-7945
Ekin Sağlıcan https://orcid.org/0000-0001-8646-1163
Ezgi N. Altınışık https://orcid.org/0000-0003-0653-4292
Dilek Koptekin https://orcid.org/0000-0003-2664-5774
Kamilla Pawłowska https://orcid.org/0000-0002-4103-4501
Benjamin S. Arbuckle https://orcid.org/0000-0002-5445-5516
Yılmaz Selim Erdal https://orcid.org/0000-0001-8143-8159
C. Can Bilgin https://orcid.org/0000-0001-9284-307X
Eva-Maria Geigl https://orcid.org/0000-0001-6376-2094
Anders Götherström https://orcid.org/0000-0001-8579-1304
Thierry Grange https://orcid.org/0000-0001-8700-3092
Füsun Özer https://orcid.org/0000-0003-0443-5805
Mehmet Somel https://orcid.org/0000-0002-3138-1307

REFERENCES

Arbuckle, B. S. (2013). Zooarchaeology at Acemhöyük 2013. *Anadolu*, *39*, 55–68. https://doi.org/10.1501/andl_0000000403

Arbuckle, B. S., & Öztan, A. (2018). Horse and hemione hunting at Late Neolithic/chalcolithic Köşk Höyük, central Turkey. In C. Çakýrlar, J. Chahoud, R. Berthon, & S. P. Birch (Eds.), *Archaeozoology of the near east XII* (pp. 41–58). JSTOR. http://www.jstor.org/stable/j.ctvggx2m4.9

Azzaroli, A. (1991). Ascent and decline of monodactyl equids: A case for prehistoric overkill. *Annales Zoologici Fennici*, *28*(3/4), 151–163.

Baumdicker, F., Bisschop, G., Goldstein, D., Gower, G., Ragsdale, A. P., Tsambos, G., Zhu, S., Eldon, B., Ellerman, E. C., Galloway, J. G., Gladstein, A. L., Gorjanc, G., Guo, B., Jeffery, B., Kretzschumar, W. W., Lohse, K., Matschiner, M., Nelson, D., Pope, N. S., … Kelleher, J. (2022). Efficient ancestry and mutation simulation with msprime 1.0. *Genetics*, *220*(3), iyab229. https://doi.org/10.1093/genetics/iyab229

Bennett, E. A., Champlot, S., Peters, J., Arbuckle, B. S., Guimaraes, S., Pruvost, M., Bar-David, S., Davis, S. J. M., Gautier, M., Kaczensky, P., Kuehn, R., Mashkour, M., Morales-Muñiz, A., Pucher, E., Tournepiche, J.-F., Uerpmann, H.-P., Bălăşescu, A., Germonpré, M., Gündem, C. Y., … Geigl, E.-M. (2017). Taming the late Quaternary phylogeography of the Eurasiatic wild ass through ancient and modern DNA. *PLoS One*, *12*(4), e0174216. https://doi.org/10.1371/journal.pone.0174216

Bennett, E. A., Weber, J., Bendhafer, W., Champlot, S., Peters, J., Schwartz, G. M., Grange, T., & Geigl, E.-M. (2022). The genetic identity of the earliest human-made hybrid animals, the kungas of Syro-Mesopotamia. *Science Advances*, *8*, eabm0218. https://doi.org/10.1126/sciadv.abm0218

BernÁldez-Sánchez, E., & García-Viñas, E. (2019). The equids represented in cave art and current horses: A proposal to determine morphological differences and similarities. *Anthropozoologica*, *54*(1), 1–12. https://doi.org/10.5252/anthropozoologica2019v54a1

Bosse, M., Megens, J., Derks, M. F. L., de Cara, Á. M. R., & Groenen, M. A. M. (2019). Deleterious alleles in the context of domestication, inbreeding, and selection. *Evolutionary Applications*, *12*(1), 6–17. https://doi.org/10.1111/eva.12691

Boulbes, N., & van Asperen, E. N. (2019). Biostratigraphy and palaeoecology of European Equus. *Frontiers in Ecology and Evolution*, *7*, 301. https://doi.org/10.3389/fevo.2019.00301

Briggs, A. W., Stenzel, U., Johnson, P. L. F., Green, R. E., Kelso, J., Prüfer, K., Meyer, M., Krause, J., Ronan, M. T., Lachmann, M., & Pääbo, S. (2007). Patterns of damage in genomic DNA sequences from a neandertal. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(37), 14616–14621. https://doi.org/10.1073/pnas.0704665104

Burbrink, F. T., Bernstein, J. M., Kuhn, A., Gehara, M., & Ruane, S. (2022). Ecological divergence and the history of gene flow in the Nearctic Milksnakes (*Lampropeltis triangulum* complex). *Systematic Biology*, *71*(4), 839–858. https://doi.org/10.1093/sysbio/syab093

Burke, A., Eisenmann, V., & Ambler, G. K. (2003). The systematic position of *Equus hydruntinus*, an extinct species of Pleistocene equid. *Quaternary Research*, *59*(3), 459–469. https://doi.org/10.1016/S0033-5894(03)00059-0

Cai, D., Zhu, S., Gong, M., Zhang, N., Wen, J., Liang, Q., Sun, W., Shao, X., Guo, Y., Cai, Y., Zheng, Z., Zhang, W., Hu, S., Wang, X., Tian, H., Li, Y., Liu, W., Yang, M., Yang, J., … Jiang, Y. (2021). *Ancient genomes redate the extinction of Sussemionus, a subgenus of Equus, to late Holocene* (p. 2021.09.13.460072). https://doi.org/10.1101/2021.09.13.460072

Catalano, G., Modi, A., Mangano, G., Sineo, L., Lari, M., & Bonfiglio, L. (2020). A mitogenome sequence of an *Equus hydruntinus* specimen from Late Quaternary site of San Teodoro Cave (Sicily, Italy). *Quaternary Science Reviews*, *236*, 106280. https://doi.org/10.1016/j.quascirev.2020.106280

Ceballos, F. C., Gürün, K., Altınışık, N. E., Gemici, H. C., Karamurat, C., Koptekin, D., Vural, K. B., Mapelli, I., Sağlıcan, E., Sürer, E., Erdal, Y. S., Götherström, A., Özer, F., Atakuman, Ç., & Somel, M. (2021). Human inbreeding has decreased in time through the Holocene. *Current Biology*, *31*(17), 3925–3934.e8. https://doi.org/10.1016/j.cub.2021.06.027

Ceballos, F. C., Joshi, P. K., Clark, D. W., Ramsay, M., & Wilson, J. F. (2018). Runs of homozygosity: Windows into population history and trait architecture. *Nature Reviews Genetics*, *19*(4), 220–234. https://doi.org/10.1038/nrg.2017.109

Chen, Z.-H., Xu, Y.-X., Xie, X.-L., Wang, D.-F., Aguilar-Gómez, D., Liu, G.-J., Li, X., Esmailizadeh, A., Rezaei, V., Kantanen, J., Ammosov, I., Nosrati, M., Periasamy, K., Coltman, D. W., Lenstra, J. A., Nielsen, R., & Li, M.-H. (2021). Whole-genome sequence analysis unveils different origins of European and Asiatic mouflon and domestication-related genes in sheep. *Communications Biology*, *4*(1), 1–15. https://doi.org/10.1038/s42003-021-02817-4

Cleyet-Merle, J.-J., & Madelaine, S. (1991). La pendeloque magdalénienne gravée d'un «Equus hydruntinus» de la grotte du Putois II, commune de Montmaurin (Haute-Garonne). *Paléo, Revue d'Archéologie Préhistorique*, *3*(1), 119–129. https://doi.org/10.3406/pal.1991.1042

Crees, J. J., & Turvey, S. T. (2014). Holocene extinction dynamics of *Equus hydruntinus*, a late-surviving European megafaunal mammal. *Quaternary Science Reviews*, *91*, 16–29. https://doi.org/10.1016/j.quascirev.2014.03.003

Dabney, J., Knapp, M., Glocke, I., Gansauge, M.-T., Weihmann, A., Nickel, B., Valdiosera, C., García, N., Pääbo, S., Arsuaga, J.-L., & 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. https://doi.org/10.1073/pnas.1314445110

Davis, S. J. (1980). Late Pleistocene and Holocene equid remains from Israel. *Zoological Journal of the Linnean Society*, *70*(3), 289–312. https://doi.org/10.1111/j.1096-3642.1980.tb00854.x

Der Sarkissian, C., Ermini, L., Schubert, M., Yang, M. A., Librado, P., Fumagalli, M., Jónsson, H., Bar-Gal, G. K., Albrechtsen, A., Vieira, F. G., Petersen, B., Ginolhac, A., Seguin-Orlando, A., Magnussen, K., Fages, A., Gamba, C., Lorente-Galdos, B., Polani, S., Steiner, C., … Orlando, L. (2015). Evolutionary genomics and conservation of the endangered Przewalski's horse. *Current Biology*, *25*(19), 2577–2583. https://doi.org/10.1016/j.cub.2015.08.032

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. https://doi.org/10.1093/molbev/mss075

Eisenmann, V., & Baylac, M. (2000). Extant and fossil *Equus* (Mammalia, Perissodactyla) skulls: A morphometric definition of the subgenus *Equus*. *Zoologica Scripta*, *29*, 89–100.

Eisenmann, V., & Mashkour, M. (1999). The small equids (Perissodactyla, Mammalia) of the Pleistocen of Binagady (Azerbaijan) and Qazvin (Iran): *E.hemionus binagadensis* nov. subsp. and *E. hydruntinus*. *Geobios*, *32*(1), 105–122.

Eisenmann, V., & Mashkour, M. (2000). Data Base for teeth and limb bones of modern hemiones. In J. Desse & N. Desse-Berset (Eds.), *Fiche d'OsteÂologie Animale pour l'ArcheÂologie SeÂ rie B: Mammifères*. Centre de Recherches ArcheÂologiques du CNRS, Editions APDCA.

Fages, A., Hanghøj, K., Khan, N., Gaunitz, C., Seguin-Orlando, A., Leonardi, M., McCrory Constantz, C., Gamba, C., Al-Rasheid, K. A. S., Albizuri, S., Alfarhan, A. H., Allentoft, M., Alquraishi, S., Anthony, D., Baimukhanov, N., Barrett, J. H., Bayarsaikhan, J., Benecke, N., Bernáldez-Sánchez, E., … Orlando, L. (2019). Tracking five millennia of horse management with extensive ancient genome time series. *Cell*, *177*(6), 1419–1435.e31. https://doi.org/10.1016/j.cell.2019.03.049

Forsten, A., & Ziegler, R. (1995). *The horses (Mammalia, Equidae) from the early Wuermian of villa Seckendorff, Stuttgart-bad Cannstatt*. Staatliches Museum für Naturkunde.

Geigl, E.-M., & Grange, T. (2012). Eurasian wild asses in time and space: Morphological versus genetic diversity. *Annals of Anatomy – Anatomischer Anzeiger*, *194*(1), 88–102. https://doi.org/10.1016/j.aanat.2011.06.002

Green, R. E., Krause, J., Briggs, A. W., Maricic, T., Stenzel, U., Kircher, M., Patterson, N., Li, H., Zhai, W., Fritz, M. H.-Y., Hansen, N. F., Durand, E. Y., Malaspinas, A.-S., Jensen, J. D., Marques-Bonet, T., Alkan, C., Prüfer, K., Meyer, M., Burbano, H. A., … Pääbo, S. (2010). A draft sequence of the neandertal genome. *Science*, *328*(5979), 710–722. https://doi.org/10.1126/science.1188021

Guimaraes, S., Arbuckle, B. S., Peters, J., Adcock, S. E., Buitenhuis, H., Chazin, H., Manaseryan, N., Uerpmann, H.-P., Grange, T., & Geigl, E.-M. (2020). Ancient DNA shows domestic horses were introduced in the southern Caucasus and Anatolia during the Bronze Age. *Science Advances*, *6*(38), eabb0030. https://doi.org/10.1126/sciadv.abb0030

Guindon, S., & Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology*, *52*(5), 696–704. https://doi.org/10.1080/10635150390235520

Günther, T., & Nettelblad, C. (2019). The presence and impact of reference bias on population genomic studies of prehistoric human populations. *PLoS Genetics*, *15*(7), e1008302. https://doi.org/10.1371/journal.pgen.1008302

Günther, T., Valdiosera, C., Malmström, H., Ureña, I., Rodriguez-Varela, R., Sverrisdóttir, Ó. O., Daskalaki, E. A., Skoglund, P., Naidoo, T., Svensson, E. M., Castro, J. M. B. d., Carbonell, E., Dunn, M., Storå, J., Iriarte, E., Arsuaga, J. L., Carretero, J.-M., Götherström, A., & Jakobsson, M. (2015). Ancient genomes link early farmers from Atapuerca in Spain to modern-day Basques. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(38), 11917–11922. https://doi.org/10.1073/pnas.1509851112

Hordecki, J. (2020). Why did they still use the tell? An analysis of changes in post-chalcolithic settlement at Çatalhöyük, Turkey. In A. Otto, M. Herles, K. Kaniuth, L. Korn, & A. Heidenreich (Eds.), *Proceedings of the 11th international congress on the archaeology of the ancient near east* (1st ed., pp. 141–150). Harrassowitz Verlag. https://doi.org/10.2307/j.ctv10tq3zv.15

Huang, J., Zhao, Y., Bai, D., Shiraigol, W., Li, B., Yang, L., Wu, J., Bao, W., Ren, X., Jin, B., Zhao, Q., Li, A., Bao, S., Bao, W., Xing, Z., An, A., Gao, Y., Wei, R., Bao, Y., … Dugarjaviin, M. (2015). Donkey genome and insight into the imprinting of fast karyotype evolution. *Scientific Reports*, *5*(1), 14106. https://doi.org/10.1038/srep14106

Jónsson, H., Schubert, M., Seguin-Orlando, A., Ginolhac, A., Petersen, L., Fumagalli, M., Albrechtsen, A., Petersen, B., Korneliussen, T. S., Vilstrup, J. T., Lear, T., Myka, J. L., Lundquist, J., Miller, D. C., Alfarhan, A. H., Alquraishi, S. A., Al-Rasheid, K. A. S., Stagegaard, J., Strauss, G., … 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. https://doi.org/10.1073/pnas.1412627111

Jun, G., Wing, M. K., Abecasis, G. R., & Kang, H. M. (2015). An efficient and scalable analysis framework for variant extraction and refinement from population-scale DNA sequence data. *Genome Research*, *25*(6), 918–925. https://doi.org/10.1101/gr.176552.114

Kaczensky, P., Lkhagvasuren, B., Pereladova, O., Hemami, M., & Bouskila, A. (2015). *Equus hemionus*. The IUCN Red List of Threatened Species 2015: e.T7951A45171204.

Kılınç, G. M., Omrak, A., Özer, F., Günther, T., Büyükkarakaya, A. M., Bıçakçı, E., Baird, D., Dönertaş, H. M., Ghalichi, A., Yaka, R., Koptekin, D., Açan, S. C., Parvizi, P., Krzewińska, M., Daskalaki, E. A., Yüncü, E., Dağtaş, N. D., Fairbairn, A., Pearson, J., … Götherström, A. (2016). The demographic development of the first farmers in Anatolia. *Current Biology*, *26*(19), 2659–2666. https://doi.org/10.1016/j.cub.2016.07.057

Koptekin, D., Yapar, E., Vural, K. B., Sağlıcan, E., Altınışık, N. E., Malaspinas, A.-S., Alkan, C., & Somel, M. (2023). Pre-processing of paleogenomes: Mitigating reference bias and postmortem damage in ancient genome data (p. 2023.11.11.566695). *bioRxiv* https://doi.org/10.1101/2023.11.11.566695

Korneliussen, T. S., Albrechtsen, A., & Nielsen, R. (2014). ANGSD: Analysis of next generation sequencing data. *BMC Bioinformatics*, *15*(1), 356. https://doi.org/10.1186/s12859-014-0356-4

Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution*, *35*(6), 1547–1549. https://doi.org/10.1093/molbev/msy096

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*, *25*(14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & 1000 Genome Project Data Processing Subgroup. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Martin, L., & Russell, N. (2012). Cooking meat and bones at Neolithic Çatalhöyük, Turkey. In S. R. Graff & E. Rodríguez-Alegría (Eds.), *The menial art of cooking: Archaeological studies of cooking and food preparation*. University Press of Colorado.

Mashkour, M., Fontugne, M., & Hatte, C. (1999). Investigations on the evolution of subsistence economy in the Qazvin Plain (Iran) from the Neolithic to the Iron Age. *Antiquity*, *73*(279), 65–76. https://doi.org/10.1017/S0003598X00087846

Meyer, M., & Kircher, M. (2010). Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harbor Protocols*, *2010*(6), pdb.prot5448. https://doi.org/10.1101/pdb.prot5448

Moehlman, P., & Kebede, F. (2014). IUCN red list of threatened species: *Equus africanus*. *IUCN Red List of Threatened Species*. https://www.iucnredlist.org/en

Mualim, K., Theunert, C., & Slatkin, M. (2021). Estimation of coalescence probabilities and population divergence times from SNP data. *Heredity*, *127*(1), 1–9. https://doi.org/10.1038/s41437-021-00435-8

Newman, D., & Pilson, D. (1997). Increased probability of extinction due to decreased genetic effective population size: Experimental populations of *Clarkia pulchella*. *Evolution*, *51*(2), 354–362. https://doi.org/10.1111/j.1558-5646.1997.tb02422.x

Nores, C., Muñiz, A. M., Rodríguez, L. L., Bennett, E. A., & Geigl, E.-M. (2015). The Iberian zebro: What kind of a beast was it? *Anthropozoologica*, *50*(1), 21–32. https://doi.org/10.5252/az2015n1a2

Orlando, L., Ginolhac, A., Zhang, G., Froese, D., Albrechtsen, A., Stiller, M., Schubert, M., Cappellini, E., Petersen, B., Moltke, I., Johnson, P. L. F., Fumagalli, M., Vilstrup, J. T., Raghavan, M., Korneliussen, T., Malaspinas, A.-S., Vogt, J., Szklarczyk, D., Kelstrup, C. D., … Willerslev, E. (2013). Recalibrating *Equus* evolution using the genome sequence of an early Middle Pleistocene horse. *Nature*, *499*(7456), 74–78. https://doi.org/10.1038/nature12323

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. https://doi.org/10.1111/j.1365-294X.2006.02922.x

Orlando, L., Metcalf, J. L., Alberdi, M. T., Telles-Antunes, M., Bonjean, D., Otte, M., Martin, F., Eisenmann, V., Mashkour, M., Morello, F., Prado, J. L., Salas-Gismondi, R., Shockey, B. J., Wrinn, P. J., Vasil'ev, S. K., Ovodov, N. D., Cherry, M. I., Hopwood, B., Male, D., … 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. https://doi.org/10.1073/pnas.0903672106

Paradis, E., & Schliep, K. (2019). Ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, *35*, 526–528. https://doi.org/10.1093/bioinformatics/bty633

Patterson, N., Moorjani, P., Luo, Y., Mallick, S., Rohland, N., Zhan, Y., Genschoreck, T., Webster, T., & Reich, D. (2012). Ancient admixture in human history. *Genetics*, *192*(3), 1065–1093. https://doi.org/10.1534/genetics.112.145037

Patterson, N., Price, A. L., & Reich, D. (2006). Population structure and Eigen analysis. *PLoS Genetics*, *2*(12), e190. https://doi.org/10.1371/journal.pgen.0020190

Pawłowska, K. (2020). Towards the end of the Çatalhöyük east settlement: A faunal approach. *Near Eastern Archaeology*, *83*(3), 146–154. https://doi.org/10.1086/709999

Pawłowska, K., & Barański, M. Z. (2020). Conceptualization of the Neolithic world in incised equid phalanges: Anthropomorphic figurine from Çatalhöyük (GDN area). *Archaeological and Anthropological Sciences*, *12*(1), 18. https://doi.org/10.1007/s12520-019-01006-z

Pedersen, J. S., Valen, E., Velazquez, A. M. V., Parker, B. J., Rasmussen, M., Lindgreen, S., Lilje, B., Tobin, D. J., Kelly, T. K., Vang, S., Andersson,

R., Jones, P. A., Hoover, C. A., Tikhonov, A., Prokhortchouk, E., Rubin, E. M., Sandelin, A., Gilbert, M. T. P., Krogh, A., … Orlando, L. (2014). Genome-wide nucleosome map and cytosine methylation levels of an ancient human genome. *Genome Research*, *24*(3), 454–466. https://doi.org/10.1101/gr.163592.113

Posada, D. (2008). jModelTest: Phylogenetic model averaging. *Molecular Biology and Evolution*, *25*(7), 1253–1256. https://doi.org/10.1093/molbev/msn083

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., Bakker, P. I. W.d., Daly, M. J., & Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics*, *81*(3), 559–575. https://doi.org/10.1086/519795

Quinlan, A. R., & Hall, I. M. (2010). BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics*, *26*(6), 841–842.

Rambaut, A. (2014). FigTree. http://tree.bio.ed.ac.uk/software/figtree/

Regalia, E. (1907). Sull'Equus (Asinus) hydruntinus Regalia della grotta di Romanelli. *Archivio per l'Antropologia e l'Etnologia*, *37*, 375–390.

Reimer, P. J., Austin, W. E. N., Bard, E., Bayliss, A., Blackwell, P. G., Ramsey, C. B., Butzin, M., Cheng, H., Edwards, R. L., Friedrich, M., Grootes, P. M., Guilderson, T. P., Hajdas, I., Heaton, T. J., Hogg, A. G., Hughen, K. A., Kromer, B., Manning, S. W., Muscheler, R., … Talamo, S. (2020). The IntCal20 northern hemisphere radiocarbon age calibration curve (0–55 cal kBP). *Radiocarbon*, *62*(4), 725–757. https://doi.org/10.1017/RDC.2020.41

Renaud, G., Petersen, B., Seguin-Orlando, A., Bertelsen, M. F., Waller, A., Newton, R., Paillot, R., Bryant, N., Vaudin, M., Librado, P., & Orlando, L. (2018). Improved de novo genomic assembly for the domestic donkey. *Science Advances*, *4*(4), eaaq0392. https://doi.org/10.1126/sciadv.aaq0392

Revell, L. (2012). Phytools: An R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution*, *3*, 217–223. https://doi.org/10.1111/j.2041-210X.2011.00169.x

Robinson, J. A., Räikkönen, J., Vucetich, L. M., Vucetich, J. A., Peterson, R. O., Lohmueller, K. E., & Wayne, R. K. (2019). Genomic signatures of extensive inbreeding in Isle Royale wolves, a population on the threshold of extinction. *Science Advances*, *5*, eaau0757. https://doi.org/10.1126/sciadv.aau0757

Rogers, R. L., & Slatkin, M. (2017). Excess of genomic defects in a woolly mammoth on Wrangel Island. *PLoS Genetics*, *13*(3), e1006601. https://doi.org/10.1371/journal.pgen.1006601

Ross, J. C., Steadman, S. R., McMahon, G., Adcock, S. E., & Cannon, J. W. (2019). When the Giant falls: Endurance and adaptation at Çadır Höyük in the context of the Hittite empire and its collapse. *Journal of Field Archaeology*, *44*(1), 19–39. https://doi.org/10.1080/00934690.2018.1558906

Schubert, M., Lindgreen, S., & Orlando, L. (2016). AdapterRemoval v2: Rapid adapter trimming, identification, and read merging. *BMC Research Notes*, *9*(1), 88. https://doi.org/10.1186/s13104-016-1900-2

Schubert, M., Mashkour, M., Gaunitz, C., Fages, A., Seguin-Orlando, A., Sheikhi, S., Alfarhan, A. H., Alquraishi, S. A., Al-Rasheid, K. A. S., Chuang, R., Ermini, L., Gamba, C., Weinstock, J., Vedat, O., & Orlando, L. (2017). Zonkey: A simple, accurate and sensitive pipeline to genetically identify equine F1-hybrids in archaeological assemblages. *Journal of Archaeological Science*, *78*, 147–157. https://doi.org/10.1016/j.jas.2016.12.005

Skoglund, P., Mallick, S., Bortolini, M. C., Chennagiri, N., Hünemeier, T., Petzl-Erler, M. L., Salzano, F. M., Patterson, N., & Reich, D. (2015). Genetic evidence for two founding populations of the Americas. *Nature*, *525*(7567), 104–108. https://doi.org/10.1038/nature14895

Skoglund, P., Northoff, B. H., Shunkov, M. V., Derevianko, A. P., Pääbo, S., Krause, J., & Jakobsson, M. (2014). Separating endogenous ancient DNA from modern day contamination in a Siberian neandertal. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(6), 2229–2234. https://doi.org/10.1073/pnas.1318934111

Spassov, N., & Iliev, N. (2002). The animal bones from the prehistoric necropolis near Durankulak (NE Bulgaria) and the latest record of *Equus hydruntinus* Regalia. In H. Todorova (Ed.), *Durankulak II*. Deutsches Archäologisches Institut.

Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*(9), 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Steadman, S. R., Hackley, L. D., Selover, S., Yıldırım, B., von Baeyer, M., Arbuckle, B., Robinson, R., & Smith, A. (2019). Early lives: The late chalcolithic and early bronze age at ÇadırHöyük. *Journal of Eastern Mediterranean Archaeology and Heritage Studies*, *7*(3), 271–298.

Steadman, S. R., McMahon, G., & Ross, J. C. (2019). Chalcolithic, Iron Age, and byzantine investigations at ÇadırHöyük: The 2017 and 2018 seasons. In S. R. Steadman & G. McMahon (Eds.), *The archaeology of Anatolia volume III: Recent discoveries (2017–2018)* (pp. 32–52). Cambridge Scholars Publishing.

Stehlin, H. G., & Graziosi, P. (1935). *Ricerche sugli Asinidi fosili d'Europa*. Birkhäuser.

Stoletzki, N., & Eyre-Walker, A. (2011). Estimation of the neutrality index. *Molecular Biology and Evolution*, *28*(1), 63–70. https://doi.org/10.1093/molbev/msq249

Stuiver, M., & Reimer, P. J. (1993). Extended 14C data base and revised CALIB 3.0 14C age calibration program. *Radiocarbon*, *35*(1), 215–230. https://doi.org/10.1017/S0033822200013904

Toews, D. P. L., & Brelsford, A. (2012). The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology*, *21*(16), 3907–3930. https://doi.org/10.1111/j.1365-294X.2012.05664.x

Twiss, K. C., Wolfhagen, J., Madgwick, R., Foster, H., Demirergi, G. A., Russell, N., Everhart, J. L., Pearson, J., & Mulville, J. (2017). Horses, hemiones, hydruntines? Assessing the reliability of dental criteria for assigning species to southwest Asian equid remains. *International Journal of Osteoarchaeology*, *27*(2), 298–304. https://doi.org/10.1002/oa.2524

van Asperen, E. N. (2012). Late Middle Pleistocene horse fossils from northwestern Europe as biostratigraphic indicators. *Journal of Archaeological Science*, *39*(7), 1974–1983. https://doi.org/10.1016/j.jas.2012.02.025

Vilstrup, J. T., Seguin-Orlando, A., Stiller, M., Ginolhac, A., Raghavan, M., Nielsen, S. C. A., Weinstock, J., Froese, D., Vasiliev, S. K., Ovodov, N. D., Clary, J., Helgen, K. M., Fleischer, R. C., Cooper, A., Shapiro, B., & Orlando, L. (2013). Mitochondrial phylogenomics of modern and ancient equids. *PLoS One*, *8*(2), e55950. https://doi.org/10.1371/journal.pone.0055950

Wade, C. M., Giulotto, E., Sigurdsson, S., Zoli, M., Gnerre, S., Imsland, F., Lear, T. L., Adelson, D. L., Bailey, E., Bellone, R. R., Blöcker, H., Distl, O., Edgar, R. C., Garber, M., Leeb, T., Mauceli, E., MacLeod, J. N., Penedo, M. C. T., Raison, J. M., … Lindblad-Toh, K. (2009). Genome sequence, comparative analysis, and population genetics of the domestic horse. *Science*, *326*(5954), 865–867. https://doi.org/10.1126/science.1178158

Wang, C., Li, H., Guo, Y., Huang, J., Sun, Y., Min, J., Wang, J., Fang, X., Zhao, Z., Wang, S., Zhang, Y., Liu, Q., Jiang, Q., Wang, X., Guo, Y., Yang, C., Wang, Y., Tian, F., Zhuang, G., … Zhong, J. (2020). Donkey genomes provide new insights into domestication and selection for coat color. *Nature Communications*, *11*(1), 6014. https://doi.org/10.1038/s41467-020-19813-7

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag. ISBN 978-3-319-24277-4. https://ggplot2.tidyverse.org

Yang, Z. (2007). PAML 4: Phylogenetic Analysis by Maximum Likelihood. *Molecular Biology and Evolution*, *24*(8), 1586–1591. https://doi.org/10.1093/molbev/msm088

Yi, X., Liang, Y., Huerta-Sanchez, E., Jin, X., Cuo, Z. X. P., Pool, J. E., Xu, X., Jiang, H., Vinckenbosch, N., Korneliussen, T. S., Zheng, H., Liu, T., He, W., Li, K., Luo, R., Nie, X., Wu, H., Zhao, M., Cao, H., … Wang, J. (2010). Sequencing of 50 human exomes reveals adaptation to high altitude. *Science (New York, N.Y.)*, *329*(5987), 75–78. https://doi.org/10.1126/science.1190371

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Özkan, M., Gürün, K., Yüncü, E., Vural, K. B., Atağ, G., Akbaba, A., Fidan, F. R., Sağlıcan, E., Altınışık, E. N., Koptekin, D., Pawłowska, K., Hodder, I., Adcock, S. E., Arbuckle, B. S., Steadman, S. R., McMahon, G., Erdal, Y. S., Bilgin, C. C., Togan, İ., … Somel, M. (2024). The first complete genome of the extinct European wild ass (*Equus hemionus hydruntinus*). *Molecular Ecology*, *33*, e17440. https://doi.org/10.1111/mec.17440