ARCHAEOGENOMICS OF ANATOLIAN WILD ASS

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

MUSTAFA ÖZKAN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY
IN
BIOLOGY

JULY 2023

Approval of the thesis:

ARCHAEOGENOMICS OF ANATOLIAN WILD ASS

submitted by **MUSTAFA ÖZKAN** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy in Biology Department, Middle East Technical University** by,

Prof. Dr. Halil Kalıpçılar
Dean, Graduate School of **Natural and Applied Sciences** _____

Prof. Dr. Mesut Muyan
Head of Department, **Biology** _____

Prof. Dr. Mehmet Somel
Supervisor, **Biology, METU** _____

Examining Committee Members:

Prof. Dr. Cemal Can Bilgin
Biology, METU _____

Prof. Dr. Mehmet Somel
Biology, METU _____

Assoc. Prof. Dr. Derya Silibolatlaz Baykara
Migration Institute, Gaziantep University _____

Assist. Prof. Dr. Gülşah Merve Kılınç
Bioinformatics, Hacettepe University _____

Assist. Prof. Dr. İsmail Kudret Sağlam
Molecular Biology, Koç University _____

Date:13.07.2023

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Surname: Mustafa Özkan

Signature :

ABSTRACT

ARCHAEOGENOMICS OF ANATOLIAN WILD ASS

Özkan, Mustafa
Ph.D., Department of Biology
Supervisor: Prof. Dr. Mehmet Somel

July 2023, 107 pages

European wild ass (*E. h. hydruntinus* or hydruntines) puzzled scientists since their first discovery nearly a century ago. Taxonomy, population history, and extinction mechanisms of this species are not clear since they show a mixture of characteristics of other asses and ancient DNA research focusing on hydruntines utilized mitochondrial sequences so far. In this thesis, I present three nuclear genomes of hydruntines excavated from Iron Age Central Anatolia.

The material of this work involves three equid genomes from Çadır Höyük and Çatalhöyük, of unknown species identity. I generated genomes of coverage 6.38x-0.57x. Using this data I defined their status as wild asses. The collective evidence showed that these three individuals were hydruntines. Because no nuclear genome data has yet been published for this species, this allowed me to study the phylogenetic and demographic history of this species for the first time. These analyses confirmed that hydruntines and hemiones/hemippes from Gobi/Tibet and the Middle East region were sister taxa. I next showed that hydruntines (as represented by the three Anatolian individuals) had a distinct gene pool from hemiones, onagers, and hemippus; with the latter, Asian asses sharing more genetic drift with each other than with hydruntines.

Whole genome analyses also suggested that hydruntines from Anatolia and Middle Eastern subspecies of hemiones experienced a gene flow event.

Next, I inspected homozygous tracks and heterozygosity levels of the Anatolian hydruntine genome to elucidate extinction dynamics. These analyses showed a decrease in heterozygosity and an increase in homozygous regions relative to extant Asian wild asses, in accordance with a declining population scenario.

In this thesis, I thus shed light on the population genomics of hydruntines. I showed that these last known hydruntines experienced a population decline before going extinct, as zooarchaeologists predicted.

Keywords: Anciet DNA, Equus hydruntinus, demography, taxonomy, popultion genetics

ÖZ

ANADOLU YABAN EŞEĞİNİN ARKEOGENOMİK ANALİZİ

Özkan, Mustafa
Doktora, Biyoloji Bölümü
Tez Yöneticisi: Prof. Dr. Mehmet Somel

Temmuz 2023 , 107 sayfa

Avrupa yaban eşekleri (*E. h. hydruntinus* ya da hidruntinler) yaklaşık bir yüzyıl önce keşfedilmelerinden bugüne araştırmacıların kafasını karıştırmıştır. Bu türün karışık morfolojik özellikleri ve tür üzerine yapılan antik DNA çalışmalarının yalnızca mitokondrial dizilere dayanması, türün taksonomisi, popülasyon tarihi ve soy tükenme mekaniklerinin anlaşılamamasına sebep olmuştur. Bu tez çalışması kapsamında, Demir Çağ Orta Anadolu'sundan çıkarılmış örneklerden üç tanesinin nükleer genomunu sunuyorum.

Bu tezin materyalini tür tayini yapılmamış üç atgil örneği teşkil etmektedir. 6.38x ile 0.57x arasında değişen kapsamda genomlar ürettim. Bu veriyi kullanarak yaban eşekleri olduğunu tespit ettim. Birçok kanıt gösterdi ki bu üç birey hidruntindi. Bu türün tüm genom verisi üretilmemiş olması bu türün filogenetik ve demografik tarihini ilk kez çalışmamı sağladı. Analizler hidruntinler ile Gobi/Tibet hemiyonları ve Orta Doğu hemiplerinin kardeş türler olduğunu doğruladı. Asya ve Orta Doğu eşeklerinin birbirleri ile daha fazla genetik sürüklenme paylaşmasından ötürü bu tez kapsamında üç birey ile temsil edilen hidruntinlerin farklı bir gen havuzunu temsil ettiğini göster-

dim. Tüm genom analizleri ayrıca Anadolu hidruntinleri ile Orta Doğu'dan hemiyon alt türleri arasında gen akışı olduğunu gösterdi.

Daha sonra, soy tükenme dinamiklerini açıklamak için Anadolu hidruntin genomlarının homozigot bölgeleri ve heterozigotluklarını inceledim. Bu analizler küçülmekte olan bir popülasyon senaryosuna uygun olarak, yaşayan Asya eşeklerine oranla azalan heterozigotluğu ve artan homozigot bölgeleri gösterdi.

Bu tezde hidruntinlerin popülasyon genomiğine ışık tuttum. Zooarkeologların tahmin ettiği gibi bu son hidruntinlerin soyları tükenmeden önce popülasyon daralması geçirdiğini gösterdim.

Anahtar kelimeler: Antik DNA, Equus hydruntinus, demografi, taksonomi, popülasyon genetiği

Anahtar Kelimeler: Antik DNA, Equus hydruntinus, demografi, taksonomi, popülasyon genetiği

To all the animals with whom we share this Pale Blue Dot

ACKNOWLEDGMENTS

This thesis work is not a struggle of a single person but rather a collective effort of many people, from a friendly touch to endless hours of staring at computer screen contests.

I want to express my sincere gratitude and appreciation to my thesis supervisor, Mehmet Somel. He has always had unwavering faith in me, and I am genuinely thankful for his support. With his advice, persistence in exploring my limits, and guidance in terra incognita of a field I was alien to, I am the scientist I am now.

During my journey, Füsun Özer was an exceptional role model, and I greatly benefited from her guidance. Her calm demeanor, focus on the solution rather than the problem, and excellence in finding a way to communicate in delicate matters have made me a good scientist and man. The collaboration we had for the project that serves as the foundation of my thesis amidst the challenging times our world faced is a significant aspect of my life.

I would like to thank Derya Silibolatlaz Baykara and Cemal Can Bilgin for their indispensable contributions to my thesis as members of the thesis monitoring committee. Their expertise and continuous support improved this thesis over the years.

I want to extend my deepest gratitude to Eva-Maria Geigl, Thierry Grange, Anders Götherström, and İnci Togan for their valuable expertise and support of this thesis. I have never felt lost during this work. As senior scientists, they have inspired and influenced me to follow in their footsteps willingly.

I will be forever indebted to current members and alums of Comparative and Evolutionary Biology (CompEvo) at METU Biology. I would like to especially thank Ali Akbaba for his indispensable help in wet lab steps; Kıvılcım Başak Vural for her assistance in the dry lab process and tireless efforts in dataset construction; Eren Yüncü and Kanat Gürün for their assistance in mitochondrial DNA analyses and Gözde Atağ

for her contribution to nuclear divergence time estimations. I want to thank Ezgi Altınışık, Dilek Koptekin, Fatma Rabia Fidan, and Ekin Sağlıcan not only for their scientific contributions but also for their unyielding support and friendship despite my never-ending questions. Also, I thank Sevgi Yorulmaz and Duygu Deniz Kazancı for their tireless efforts in the wet lab organization. I thank all members of CompEvo for their valuable contributions during lab meetings and retreats.

The archaeological teams of Çadır Höyük and Çatalhöyük also should be acknowledged for their faith in this thesis and project. I would like to thank Benjamin Arbuckle, Sarah Adcock, Gregory McMahon, and Sharon Steadman from the Çadır Höyük team and Kamilla Pawlowska, Mikolaj Lisowski, and Ian Hodder from the Çatalhöyük team for providing equid samples and their willingness to discuss zooarchaeological questions.

I am grateful to Ecology and Evolutionary Biology Society (EkoEvo) for providing a smooth transition from one discipline to another through their efforts in sharing knowledge and increasing the quality of ecology and evolutionary biology research in Turkey.

I acknowledge support from the Scientific and Technical Research Council in Turkey (TÜBİTAK, no. 117Z991 to Füsun Özer), H2020 ERC Consolidator Grant (no. 772390 to Mehmet Somel), and H2020 TWINNING Grant (no. 952317 to Mehmet Somel).

I cannot express my gratitude to my friends for their continuous support during the tumultuous times experienced by both the world and this student. The friendship of the whole Düşünkara people, Burak Derkuş, Cemil Can Eylem, Utku Karakaya, Onur Yıldırım, Bengisu Biray, Ebru Konyar, Onur Altıntaş, Yeşim Kuzhan, Merve Nur İlga, and Gözde Vatansever kept me from falling into despair, gave me strength and peace of mind. My acquaintance with Ali Akbaba, Ezgi Altınışık, Kıvılcım Başak Vural, and Dilek Koptekin started as labmates but turned into a lifelong friendship during this rocky journey.

Lastly, words cannot express how lucky I am to have my family. In the direst conditions, I always felt that I am important to someone. Their unyielding faith in me

always made me continue, even in the darkest hour.

TABLE OF CONTENTS

ABSTRACT v

ÖZ vii

ACKNOWLEDGMENTS x

TABLE OF CONTENTS xiii

LIST OF TABLES xvii

LIST OF FIGURES xviii

LIST OF ABBREVIATIONS xxii

CHAPTERS

1 INTRODUCTION 1

1.1 Equid Evolution 2

1.1.1 Progenitors of horses, donkeys and asses 2

1.1.2 European wild ass 4

1.2 Archaeogenomics 8

1.2.1 Equid evolution 8

1.2.1.1 Horse domestication 11

1.2.1.2 Donkey domestication 12

1.2.2 European wild ass 13

1.2.3 Other extinct species 14

1.3 Thesis Outline 16

2 MATERIAL AND METHODS 19

2.1 Archaeological samples 19

2.1.1 Çatalhöyük 19

2.1.2 Çadır Höyük 20

2.2 Radiocarbon dating 20

2.3 Ancient DNA isolation and sequencing 21

2.4 Data preprocessing 23

2.5 Authentication of genetic data and trimming 25

2.6 Complete mitogenome sequence 26

2.7 Taxonomic identification and sex determination 26

2.8 Phylogenetic analyses of mtDNA data 26

2.9 SNP panel construction 28

2.10 Genotyping and dataset construction 29

2.11 Outgroup f3-statistics 30

2.12 Multidimensional scaling 30

2.13 D-statistics 31

2.14 Divergence time estimation 32

2.15 Runs of homozygosity 33

2.16 Genome-wide heterozygosity estimation 34

3 RESULTS 37

3.1 Genetic Analyses Assign the Iron Age Anatolian Wild Ass Individuals to *E.h.hydruntinus*/*E.hydruntinus* 42

3.2 Whole Genome Analyses Support Hemiones and Hydruntines as Sister Clades . . . 44

3.3 Hemiones and Hydruntines Are Reciprocally Monophyletic . . . 46

3.4 Evidence for Gene Flow Between Hydruntines and Middle Eastern Wild Asses . . . 48

3.5 The Timing of the Hydruntine and Asian Wild Ass Split . . . 51

3.6 Severely Depleted Genetic Diversity in the Anatolian Wild Ass Signals Population Decline . . . 53

4 DISCUSSION . . . 55

4.1 Ancient DNA Conservation and Sequencing Characteristics . . . 55

4.2 Sex Distribution of the Samples . . . 56

4.3 Biases in population genomic analyses . . . 57

4.4 European Wild Asses of Anatolia . . . 58

4.5 Population History of European Wild Asses . . . 59

4.5.1 Divergence between African and Eurasiatic wild ass branches . 59

4.5.2 Degrees of variation within the Eurasiatic wild ass gene pool . 60

4.5.3 Gene flow between Eurasiatic wild ass species . . . 61

4.6 Extinction Dynamics of European Wild Ass . . . 61

4.7 Conclusions . . . 62

REFERENCES . . . 65

APPENDICES

A . . . 89

B . . . 95

C . . . 97

D 99

E 101

F 103

LIST OF TABLES

TABLES

Table 3.1 Genome information and radiocarbon dates of the three newly sequenced ancient Anatolian equids. 39

Table A.1 Bronze and Iron Age archaeological sites where Equids have been reported among faunal remains in modern-day Turkey. EBA: Early Bronze Age; MBA: Middle Bronze Age; LBA: Late Bronze Age; BA: Bronze Age; IA: Iron Age. 89

Table B.1 Pre-screening statistics of 15 equid samples. Samples which were sequenced further (see Supplementary Table 2) are indicated in bold. 5' damage and 3' damage indicate the proportion of C->T and G->A mismatches at first and last positions across all reads (Undet: Undetermined) . 96

Table C.1 Deep sequencing statistics of selected individuals. Merged library (screening + deep sequencing) statistics are given in bold. 98

LIST OF FIGURES

FIGURES

Figure 1.1 Phylogenic tree of Equidae (adapted after Macfadden et al.[1]) . 3

Figure 1.2 *E. hydruntinus* skull from Emine-Bair-Khosar excavation site, from Van Asperen et al.[2]. A) Lateral and B) Occlusal view. 5

Figure 1.3 Plot showing proctone length versus occlusal surface of first two molars. A-C) Upper teeth of *E. h. hydruntinus* from Ramandils cave. D-E) from Crouzade Cave. F) Right P4 of *Equus graziosi*. Adapted after Eisenmann et al. [3] 7

Figure 1.4 *E. hydruntinus* distribution in Holocene. (Circles: *E. hydruntinus*; diamonds: questionable records of *E. hydruntinus*; squares: *E. hemionus*. From Crees and Turvey[4]) 8

Figure 1.5 Wild ass phylogeny based on mitochondrial control region. Estimated mean node heights given in red and representing kya. Line thickness is proportional with posterior clade probability. After[5] . . . 14

Figure 2.1 Summarised workflow of ancient DNA wet lab process. 22

Figure 2.2 Schematic representation of A) outgroup-f3 and B) D-statistics with which I carried demographic analyses. In this schema, X1, X2, and X3 shows test populations while X4 represents outgroup. Allele frequency differences on branches calculated and multiplied. Gene flow between X3 and X2 yields negative D while between X3 and X1 yields positive D. Adapted after[6]. 31

Figure 3.1 A) The map shows the geographical locations of Çadır Höyük and Çatalhöyük (the excavation sites where the wild ass remains analyzed in this study were recovered), as well as the approximate dispersal areas of extant and extinct ass taxa. B) Shows the locations of three ancient hemione individuals and one hemippe individual the partial genomes of which were previously published[7, 8]. C) Timeline of dated *E. hydruntinus* remains published in the literature including samples reported in this study (adapted after Crees and Turvey[4]). . . . 38

Figure 3.2 Post-mortem damage signals of cdh008, cdh010 and chh003. . . 40

Figure 3.3 Mitochondrial DNA median joining network of the Eurasiatic wild asses from Central/East Asia, Iran, India, Syria, Caucasus, Anatolia, Balkan and Western Europe[9, 10, 5, 11], and sequences reported in this article generated using 249 bp diagnostic region of D-loop (Appendix D, Methods). *H1*, which colored with light blue here, represents the main cluster of hydruntine haplotypes. (Capital letters denote haplogroups and names with italic font denote sample names, see Appendix D.) 41

Figure 3.4 A Bayesian phylogenetic tree constructed using a 13,042 bp long consensus sequence of the whole mitochondrial DNA among various equid lineages. Numbers on the internal nodes show posterior probability support. X axis denotes time in millenial scale. 43

Figure 3.5 Multidimensional scaling plot of A) donkey and wild asses and B) only the Eurasiatic wild asses generated using a distance matrix based on (1-outgroup f3) 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*). 44

Figure 3.6 Heatmap generated using outgroup-f3 data on autosomal loci total of 2,146,416 positions on individual level. 45

Figure 3.7 D-statistics calculated using autosomal data between regional groups. Test groups are shown on the top. All D-statistic point estimates belongs to a comparison and boxplot represents the interquartile ranges. The color gradient from pink to black represents the fraction of individual comparisons that are nominally significant (|Z|>3). Regional groups include: Anatolia (cdh008, cdh010, chh003), Gobi/Tibet (Aw1, Aw2, Aw3, Kia1, Kia2), Middle East (GT64, Hm_1864, Onager, Hasanlu3459, Sagzabad) and Africa (Asinus, Somalicus). 46

Figure 3.8 D statistics results in the form of D(Horse,Anatolia; MiddleEast, Gobi/Tibet) based on the main autosomal dataset consisting of 2,146,461 variable sites. These are the same results as presented in summary form in Figure 3.7. Pink color represents significant results at |Z|>3 and gray color |Z|<3. Confidence intervals represents 2 SE. 48

Figure 3.9 D statistics results in the form of D(Horse,Anatolia; MiddleEast, Gobi/Tibet) based on the secondary autosomal variable loci (2,408,064 positions). Blue color represents significant results at |Z|>3 and gray color |Z|<3. Confidence intervals represents 2 SE. 49

Figure 3.10 D statistics results in the form of D(Horse,Anatolia; MiddleEastX, MiddleEastY) based on the main autosomal dataset consisting of 2,146,461 variable sites. These are the same results as presented in summary form in Figure 3.7. Pink color represents significant results at |Z|>3 and gray color |Z|<3. Confidence intervals represents 2 SE. 50

Figure 3.11 D statistics results in the form of D(Horse,Anatolia; MiddleEastX, MiddleEastY) based on the secondary autosomal variating loci(2,408,064 positions). Blue color represents significant results at |Z|>3 and gray color |Z|<3. Confidence intervals represents 2 SE. 51

Figure 3.12 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 (Methods). Horizontal lines represents F(A|B) values for cdh008, *E. a. asinus* and *E. h. onager*. Observed and expected F(A|B) values, calculated using a) $e^{-T(2N)3}$ and b) population genetic simulations, are given for divergence times ranging from 100kya to 2 mya, and for population sizes (Ne) ranging from 5k to 80k by the color key. 52

Figure 3.13 A) Total number versus the total length (sum) of runs of homozygosity (ROH) tracks over 1.5 Mb. B) Heterozygosity calculated across 2,146,416 transversion sites. (ANAT: Anatolian samples reported in this article; ASI: *E.asinus*; HEM: *E.hemionus*; KIA: *E.h.kiang*; ONA: *E.h.onager*; SOM: *E.somalicus*). 53

LIST OF ABBREVIATIONS

aDNA	Ancient DNA
AD	Anno Domini
BCE	Before Common Era
bp	base pairs
cM	Centimorgan
DNA	Deoxyribonucleic acid
kya	Thousand years ago
LGM	Last Glacial Maximum (26 - 20 kya)
MIS	Marine isotope stage
Mya	Million years ago
MDS	Multidimensional scaling analysis
NGS	Next Generation Sequencing
PCR	Polymerase Chain Reaction
PMD	Post Mortem Damage
ROH	Runs of Homozygosity

CHAPTER 1

INTRODUCTION

Horses and donkeys have been companion animals of humankind since the early periods of civilization; thanks to their strength and endurance, they not only shared the burden of developing society but also contributed to reaching far distances in relatively short times and facilitated trade and knowledge networks. Collectively known as Equids, these animals are entrenched into ancient and modern culture, from idioms such as "dark horse," which define a person or thing that has not been popular before but gains significance according to circumstances, to quotes, for example, "A wilful plow-ox should be put back in the track, a balking ass should be made to take the straight path." from Anzu, a lesser divinity in ancient Mesopotamian religions. Domestication of horses and donkeys has been considered one of humankind's most significant achievements. They once played an essential role in the daily life of some cultures until the advent of combustion engines. Even today, they are referenced in art, such as in the movie "The Lord of the Rings: Two Towers." In the film, the king of the fictional land Rohan, Thèoden, laments the absence of horses and riders before confronting a much larger enemy force, uttering, "Where is the horse... and the rider?" On the other hand, donkeys have become a mascot for numerous organizations, as they are known for their stubbornness.

In this thesis work, I studied the archaeogenomics of extinct wild asses of Anatolia. Throughout history, domesticated equids have received more attention in archaeogenomics, leaving the place of their wild counterparts in the equid family open to question. While previous research based on mitochondrial DNA sequences offered limited genomic knowledge, most of the conclusions about the history and fate of this species were derived from morphological comparisons and presence-absence data on

excavated remains and their radiocarbon dates. Furthermore, it is crucial to comprehend the specific factors that led to the extinction of this animal and how this event transpired. My thesis work aims to identify the position of hydruntines in the equid family tree and uncover the complex relationship network between equids. Additionally, it examines the various challenges that different species faced throughout time, eventually leading to their extinction.

1.1 Equid Evolution

Equidae is one of the three members of odd-toed ungulates (Perissodactyla, Mammalia) along with rhinos (Rhinocerotidae) and tapirs (Tapiridae). A single genus, Equus, today represents the equid family. Within Equus, three major subgenera consist of several species: 1) Feral (*Equus przewalski*) and domesticated horses (*Equus caballus*) constitute caballine horses; 2) Donkey (*Equus asinus*), Somali wild ass (*Equus asinus somaliensis*) and Asian wild asses (*Equus hemionus*, *Equus kiang*, *Equus hemionus onager*) form one part of non-caballine or stenonid horses with zebras (*Equus hartmanne*, *Equus quagga*, *Equus greyvi*) compose the other. Apart from modern-day populations, over 42 genera identified in the Equid family are now extinct and have only been defined through the fossil record[12].

1.1.1 Progenitors of horses, donkeys and asses

Equids received scientific attention due to their lasting impact on human history. Research on various topics, from their origins to behavioral patterns, has been conducted over two centuries. The history of the Equid family gained popularity in discussions of evolution during the late 19th century. Thomas H. Huxley, a prominent defender of evolution, analyzed Othniel C. Marsh's collection of equid remains, which included small mammal remains that exhibited subtle resemblances to modern horses. Although his views were on only one of many lineages through equid evolution (Figure 1.1), it grew into a textbook example of the transition from small-sized, multi-digital, and browsing organisms to large, single-toed grazing equids in modern times[13].

The origins of equids can be traced back to small-sized, browsing species of hyra-

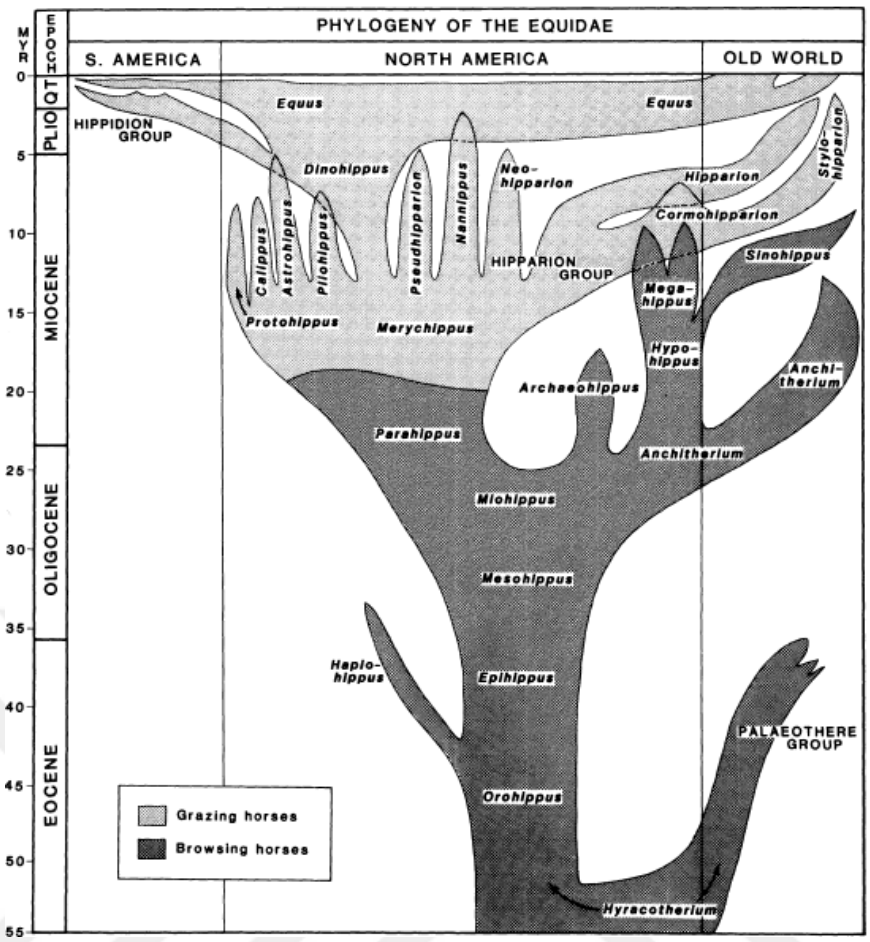


Figure 1.1: Phylogenic tree of Equidae (adapted after Macfadden et al.[1])

cotheres, which lived approximately 55 million years ago (Mya) in North America[14, 15]. Between Eocene (55 – 34 Mya) and mid-Miocene Climatic Optimum (mMCO, 18 – 15 Mya), there was little morphological change in the equid fossil record[16]. An increase in temperature in mMCO brought ecological changes in biomes, namely a shift from woodlands to mixed woodland-savannah biomes. It opened a significant number of vacant niches and triggered explosive radiation of grazing horses with at least 19 different species, 16 of which lived contemporaneously[16].

A thorough understanding of the exact timing of equid dispersal into Eurasia is still needed. However, it is believed that it occurred in at least two separate events. The earliest dispersal took place approximately 2.6 Mya[17, 18], bringing a diverse group of herbivores in Eurasia, and one of these were stenonine-like equids, which were ancestors of donkeys, wild asses, and zebras. The second distinct event was char-

acterized as the dispersal of caballine horses, ancestors of modern-day feral and domesticated horses, into Eurasia toward the end of the Villafranchian (1 – 0.9 Mya) period[19]. In this period, caballine horses replaced stenonine horses almost everywhere in Eurasia, leaving Africa as the sole region where stenonine horses took refuge[20].

Caballine horses reached central-east Europe by about 1 Mya and started to differentiate to fill various niches by 0.9 Mya. In contrast, the fate of stenonine horses for this period is unclear. There are several hypotheses on the dispersal and differentiation of ancestors of donkeys, wild asses, and zebras. According to one theory, the process of stenonine differentiation initially occurred in Asia and subsequently these spread to Africa around 0.5 million years ago[21]. This hypothesis was weakened by an ass-like fossil report, dated to 1.8 Mya, excavated in Algeria[22]. The other view supports the idea of the differentiation of stenonines in Africa and the recolonization of Asia afterward. Support for this hypothesis depends on the megafaunal change following the late Villafranchian dispersal of caballines and the eradication of stenonines from Eurasia. Because divergence between asses and zebras is shown to have occurred 2 Mya[23] using archaeogenomic analyses and the oldest stenonine from the Africa dated to 2.3 Mya[18], the African origin hypothesis of modern stenonines is logical.

1.1.2 European wild ass

One of the most puzzling species of extinct equids is *Equus hydruntinus*, also known as European wild ass. The first remains described as *Equus hydruntinus* were reported by Regalia[24], from an Upper Paleolithic context of Romanelli Cave, attributed to the Epigravettian. Following this report, extensive research on this species' anatomy and geographical distribution was also published[25]. It is known that very little information exists about the origins of this enigmatic species. Some research placed *E. hydruntinus* as descendants of *Equus altidens*[26, 27, 28], based on their size and chronology of fossils. Also, *Equus tabeti* or *Equus cf. tabeti* were thought to be had affinities with *E. hydruntinus*[29]. However, the current idea about hydruntine origins claims an arrival from Asia rather than a local evolution in Europe[30]. This origin model suggests that hydruntines spread to Europe and filled the niche of *E. altidens*,

after the decline of the latter in the Early-middle Pleistocene. At the height of their dispersal, hydruntines were to be found from Western Europe to the Middle East, including Iran in the south and the Caucasus in the north[9].



Figure 1.2: *E. hydruntinus* skull from Emine-Bair-Khosar excavation site, from Van Asperen et al.[2]. A) Lateral and B) Occlusal view.

The sporadic presence of *Equus hydruntinus* in the fossil record is likely a result of its rarity[31]. The oldest known remains attributed to European wild ass are reported from Vallparadis, Spain, and dated to slightly younger than 0.6 Mya, by analyses based on other faunal remains[32]. The sequence in Spain continues with findings in the Sierra de Atapuerca complex, with remains considered younger than the ones in Vallparadis[33]. The oldest remains found in Italy, where the species was first identified, were dated slightly younger than sightings in Spain and accepted as one of the significant markers of the beginning of the Middle Aurelian period[34]. Lunel-Viel cave in France was long thought to be containing the oldest fossils before findings in Spain[35]; however, older French sites also presented hydruntine presence, such as Achenheim[36]. Other researchers found evidence of hydruntine dispersal in Germany[37] and British Isles[38]. Based on reports from regions outside of Eu-

rope, most significantly those of from Transcaucasia[39], Syria[40] and Jordan[41]), it seems that hydruntines migrated to Europe during the Marine Isotope Stage 15 (approximately 621,000 years ago), followed by a second wave of migration into Europe between Marine Isotope Stages 11-9 (427-300,000 years ago) [31].

The elusive nature of these enigmatic creatures makes it challenging to pinpoint their position in the equid family, given their mixed characteristics. The presence of various subspecies and ecomorphs in literature further complicates the understanding of their taxonomy (Figure 1.3). For example, *Equus hydruntinus minor* is a term used to describe smaller European wild asses population found in Lunel-Viel[35]. On the other hand, a larger variant of hydruntines is also reported from Romania, named *Equus hydruntinus danubiensis*[42]. Furthermore, *Equus petralonensis*, dated to 0.4 Mya, is also considered to be within the variation of *E. h. hydruntinus* as *Equus hydruntinus petralonensis*[43]. A species known as *Equus hydruntinus davidii* was initially described based on a single tooth, but it was later discovered to be prevalent in Southwestern France and identified as a variation of the hydruntine species. This information was documented in a study by Alimen in 1946[44].

The morphological features of hydruntines have long been discussed with respect to their positioning in the equid family tree. *E. hydruntinus* is described as having slender limb bones adapted for running[45], small teeth proportions, short proctones on the upper cheek teeth, and primitive enamel pattern on the lower cheek teeth[46]. Limb bone morphology of *E. hydruntinus* resembles those of Asian wild asses (*E. hemionus*), whereas upper teeth characteristics are similar with donkeys, whose ancestry has been shown to lie in Africa[46]. On the other hand, lower teeth characteristics have an affinity to extinct *E. stenonis*[46]. In summary, *E. hydruntinus* shows a mosaic of characteristics described in other equids that lived in or survived through the Pleistocene, except caballine horses[47, 45]. These mixed features had till now complicated efforts to shed light on their taxonomic status. Some researchers classified them with zebras[48], although this was later shown to be false[49, 10]. Others suggested asses[25] as precursors; sometimes hydruntines were thought to be an older offshoot of stenonids[47, 50].

The questions of how and when *E. hydruntinus* went extinct is also unclear. While

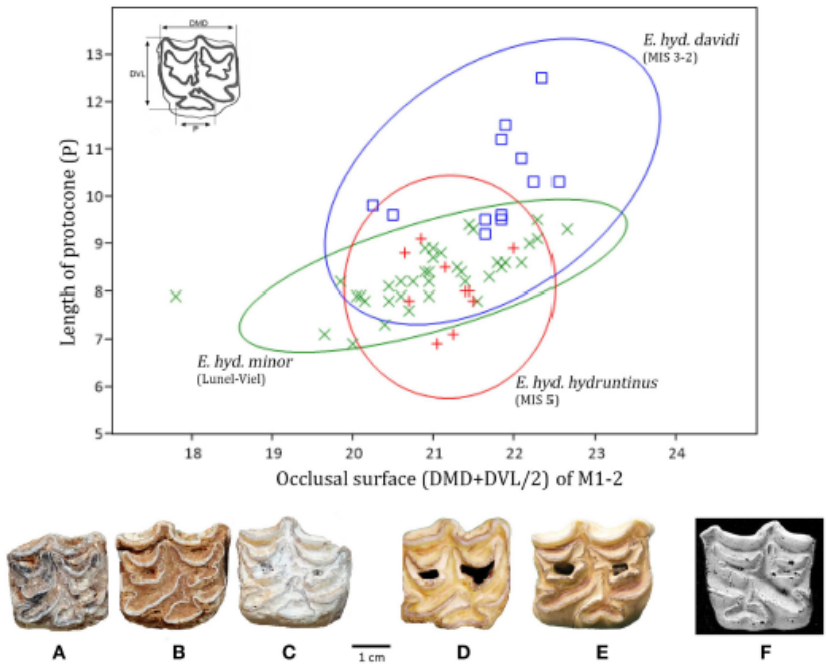


Figure 1.3: Plot showing proctone length versus occlusal surface of first two molars. A-C) Upper teeth of *E. h. hydruntinus* from Ramandils cave. D-E) from Crouzade Cave. F) Right P4 of *Equus graziosi*. Adapted after Eisenmann et al. [3]

this species' distribution covered a large area of Eurasia in the Late Pleistocene and they persisted through the Pleistocene-Holocene transition, it shrank considerably during the Holocene. At some point between the Neolithic and Iron Age, they were estimated to go extinct[4]. Although medieval equid individuals called zebro from Portugal were suggested to be European wild asses [51], studies over the last decades showed that the genetic signature of zebro remains was closer to donkeys than hydruntines[10]. Crees and Turvey examined the extinction event using radiocarbon dating of identified individuals and point pattern analysis based on occurrence data, compared with paleo-ecological changes in possible biome areas [4] (Figure 1.4). The extinction event was predicted to be caused by the combination of highly fragmented habitats that created distinct subpopulations and human exploitation. The authors reported that hydruntines went extinct in the Danube region in the third millennium BC and the global extinction date was estimated as the sixth century BC[4]. It should also be noted that Anatolia's most recent findings assigned as *E. hydruntinus* were excavated from Bronze Age culture layers (see Appendix A).

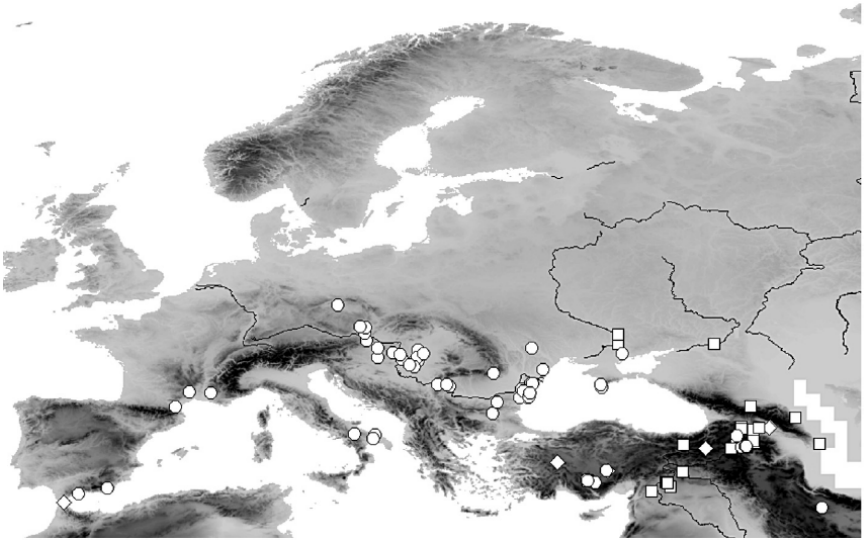


Figure 1.4: *E. hydruntinus* distribution in Holocene. (Circles: *E. hydruntinus*; diamonds: questionable records of *E. hydruntinus*; squares: *E. hemionus*. From Crees and Turvey[4])

1.2 Archaeogenomics

1.2.1 Equid evolution

While equid family history is studied chiefly with paleontological methods, with the emergence of specific protocols and the development of new sequencing methods, ancient DNA became increasingly involved in equid research. Since the beginning of ancient DNA research, equids have been familiar subjects. In fact, the first ancient DNA ever extracted was from a museum specimen of *Equus quagga*, commonly known as the Quagga zebra[52]. The authors of this study compared two fragments, one from the cytochrome oxidase I region and the other belonging to an unidentified reading frame, with zebra, cow, and human sequences to validate authenticity. The analysis clustered ancient DNA fragments with zebra as expected. Retrieving mitochondrial DNA fragment belonging to an extinct species showed that genetic material persists in tissues long after death. Following this demonstration, followed by studies focuses on other organisms, it was understood that two caveats of retrieving and analyzing ancient DNA were 1) fragmentation of DNA through time and 2) accumulation of deamination at the ends of fragments[53]. To overcome the first caveat,

researchers developed specially designed DNA isolation methods, mainly utilizing the adhesion of genetic material to an appropriate medium followed by elution into a new solution[54, 55, 56, 57]. Alternatively, methods to fish highly fragmented DNA molecules out of cell-matrix solution have also been developed[58]. Accumulation of deamination of cytosine nucleotide to uracil and its appearance as thymine in sequenced reads were also addressed in three ways. One way is treating sequencing libraries with an enzyme mainly responsible for DNA repair, such as Uracil DNA Glycosylase (UDG), to eliminate the postmortem damage patterns[59]. The other way utilizes in silico elimination of genetic variants, which would be confused with postmortem damage signal (PMD signal), namely transitions from a purine-to-purine or a pyrimidine-to-pyrimidine. Another way to overcome this obstacle is trimming sequencing reads from 5' and 3' ends for about 5 or 10 bases, on which the PMD patterns most likely occur. However, since the deamination reaction would also occur beyond the tenth base in some samples, this method might still cause false transition-type alleles. The postmortem damage pattern phenomenon is also a criterion of the authenticity of DNA extracted from the archaeological sample. A low level of PMD signal means failure to extract endogenous ancient DNA i.e., modern contamination[60].

After nearly a decade following the first ancient DNA extraction equids again appeared in ancient DNA research. In one ancient mitochondrial DNA study, researchers sampled three South American samples, two hippidons, and one equid. The Hippidon was a species that once inhabited South America. It diverged from equids and lived alongside them on the continent but is now extinct. Their analyses, which focused on 241-394 bp long control region and 172-296 bp long cytochrome b sequences, revealed that the individuals identified as hippidons were, in fact, equids, casting doubt on morphological distinctions between these two genera[61]. Meanwhile, another ancient mitochondrial DNA study placed hippidons in proximity with equids, specifically to caballine (true horses) than stenonine horses. They also revealed that stilt-legged horses were not related to hemiones of the Old World, but rather, an endemic species of North America[62]. A thorough analysis of partial mitochondrial sequences from 22 equid specimens, both ancient and from the Old and New World, revealed that hippidons are closely related to genuine horses. This study also sug-

gested that several species, including the Cape zebra, quagga, and plains zebras, as well as *E. hydruntinus* and *E. hemionus*, were over-divided and should be considered synonymous[10].

The study of ancient DNA saw significant progress thanks to the emergence of advanced sequencing technologies such as Illumina, along with reduced sequencing expenses. This progress allowed researchers to incorporate more samples into their research and utilize whole genome data in addition to mitochondrial sequences. A whole ancient genome DNA research on horse samples excavated from North American permafrost yielded the oldest ancient DNA sequences until then, from horse remains dated to 560 - 780 thousand years before present (kyr BP)[63]. This 1.12x coverage genome, together with a Pleistocene horse dated to 43 kyr BP, five modern horses, and a modern donkey, researchers estimated that all modern equid species originated from a lineage originated 4.0 to 4.5 million years ago (Mya BP), nearly twice what paleontological research presented before. The same study also pointed out that domesticated horses and Przewalski horses diverged 38 - 72 kyr BP without recent gene flow. The analysis supported the idea that Przewalski horses are the only remaining truly wild horse population. Another study on historical Przewalski horses calibrated this divergence time to 45 kya BP while underlining that the effect of captivity over a hundred years could be characterized by reduced heterozygosity[64]. Sequences of ancient horses predating domestication found genetic structure within Eurasia in the late Pleistocene while pointing out a "cost of domestication" in modern breeds due to accumulation of deleterious mutations[65].

The first ancient equid genomes from Anatolia were published in an article focusing on *Leopard Complex* (LP) phenotype of caballine horses[66]. This phenotype, associated with a single nucleotide polymorphism in the *TRPM1* gene, affects coat color variation by bringing contrasting spots. Evidence from various forms of media, including cave art, reliefs, and tapestries, indicates that this physical characteristic existed during the Pleistocene period. The partial nuclear genomes from domesticated horses of Kırklareli-Kanlıgeçit, dated to the 3rd millennium BC, showed a positive selection pressure on LP phenotype. However, since homozygous individuals with this phenotype are also night-blind, authors reported selection against this phenotype in the late Bronze Age.

1.2.1.1 Horse domestication

The analysis of ancient genomes has contributed to determining the time and location of horse domestication, which has been a topic of debate for a long time. Genetic research on modern mitochondrial sequences revealed that the Iberian Peninsula harbored high levels of genetic diversity[67] and might be considered a potential domestication center[68]. On the other hand, pastoralists from the Pontic-Caspian steppe were nominated as the first horse riders based on linguistic research[69]. Other researchers pointed out the Levant as a horse domestication center[70]. Aridification of the Levant is known to have driven existing horse populations to extinction during early Holocene, and the late Holocene prevalence of horses was therefore interpreted as implying a human-driven domestication event. Finally, being a domestication center for several farm animals[71], Anatolia was considered to be where horse domestication happened[72]. However, the candidacy of Anatolia ended after a study sampling ancient mitochondrial sequences from Anatolia and Caucasus from the Neolithic to Iron Age[73]. This study showed that Anatolia's diversity of mitochondrial DNA rapidly increased in the 3rd millennium BC. This explosive increase in mitochondrial DNA haplogroups can be explained by the introduction of non-local lineages into Anatolia from the Caucasus. The local haplotypes prevalent before the dispersal of domesticates such as P persisted in Bronze Age individuals, suggesting the incorporation of wild mares into domesticated stock.

Another candidate for horse domestication was Central Asian steppes. In 2009, a study discovered horse remains dated to 5500 BC with wear marks and archaeometric measurements similar to those of domesticated horses in the Bronze Age[74]. During the excavation of this settlement linked to Botai culture in Kazakhstan, pottery shards were discovered and subjected to lipid residue analysis. The results indicated that horses were milked in the area. The authors presented these findings as circumstantial evidence of horse domestication in Botai. However, an ancient DNA study on this horse remains[75] revealed that these Botai individuals did not significantly contribute to the gene pool of modern domesticated horses. All individuals sampled in the last 4000 years showed only 2.7% Botai ancestry. On the other hand, these Botai horses clustered with Przewalski horses in phylogenetic analyses. Other lin-

eages of wild horses were also the focal point in research on domestication event. A massive study that reported 149 whole ancient genomes[7] (129 of them >1x nuclear coverage) pointed out that two contemporaneous lineages with predicted to be the source of domestication, one from the Iberian Peninsula and the other from Siberia, did not contribute to modern domesticated horses. They also reported a decrease in genetic diversity that arose with modern breeding practices, more intense than human management practices applied in previous millennia. The debate on horse domestication finally concluded as the researchers published 273 ancient genomes and reported that domestication occurred in Western Eurasia steppes, namely lower Don-Volga region[76]. Horses domesticated in this region nearly replaced all other local populations after 2000 BC, which was evident that Przewalski horses were not true wild horses[75].

1.2.1.2 Donkey domestication

Ancient DNA research also contributed to the understanding of the domestication of another equid species, the donkey. In 2004, genetic analysis based on 479 bp long mitochondrial sequences of the control region of modern donkeys revealed that donkeys clustered with African wild asses and determined their origin as Africa[77]. Further research to elucidate donkey domestication dynamics focused on Africa to find the local population which gave rise to the domesticates. One ancient DNA study examined the partial mitochondrial sequences of Nubian and Somali wild asses and found a significant contribution from Nubian wild asses to modern donkey mitogenome. In contrast, Somali wild ass signature was non existent[78]. Another study that generated 207 modern and 31 ancient donkey nuclear genomes provided evidence for a single domestication event in Africa, in around 5000 BC[79], supporting the findings from modern microsatellite sequences[80]. Donkeys spread into Eurasia after domestication in northeastern Africa and ultimately returned to the continent in historical times.

1.2.2 European wild ass

Perplexing features of *E. hydruntinus* such as mixed morphological characteristics, the uncertainty of the taxonomic position, and extinction mechanisms attracted ancient DNA scientists from the early 21st century. The beginning of ancient DNA research on hydruntines happened during an era when mitochondrial DNA was the principal genetic material of ancient DNA research since next-generation sequencing was not yet widely available to sequence whole ancient genomes. In the first such study, mitochondrial ancient DNA sequences corresponding to 134-288 bp long section of the control region and 143 bp long cytochrome b fragments were isolated from three individuals assigned as *E. hydruntinus* from Crimea and Iran. The results showed that hydruntines were genetically closer to Asiatic wild asses, also known as hemiones, than any other equid, including African wild ass, donkeys, and other stenonid lineages such as zebra[9] and confirming morphological findings towards this affinity[46]. The authors also pointed out that dental markers which suggest the proximity of hydruntines to asses and zebras produce poor resolution in equid taxonomic analyses.

Further research based on a 229 bp-long hypervariable region on mitochondrial DNA[5] classified non-African wild asses into eleven clades. In this study, hydruntine and hemione clades clustered together, confirming previous observations (Figure 1.5). The authors noted the lack of reciprocal monophyly in the hydruntine and hemione branches in their data. This result led to the deduction that *E. hydruntinus* was a subspecies of *E. hemionus*, rather than a separate species of itself. The authors labeled the most prevalent mitochondrial haplogroup among hydruntines as H1; interestingly, S, CI, and I clades of hemiones from Syria, Caucasus, and Iran, respectively, were not the closest relatives of hydruntines. Instead, TI and Kh clades of Turkmen and Indian individuals were the most related samples to hydruntines in this dataset. This observation suggested different population histories of hydruntines than of Syria hemippes and Iran onagers. Another surprising finding was the occurrence of a second haplogroup, labeled as H2, in one of the ancient hydruntines from France and another individual sampled from Iran. This finding hints at a complex genetic structure within the hydruntine gene pool. However, such a structure would be hard to unravel by ex-

amining mitogenomes, requiring higher quality nuclear genomes from a wide range of ancient settlements.

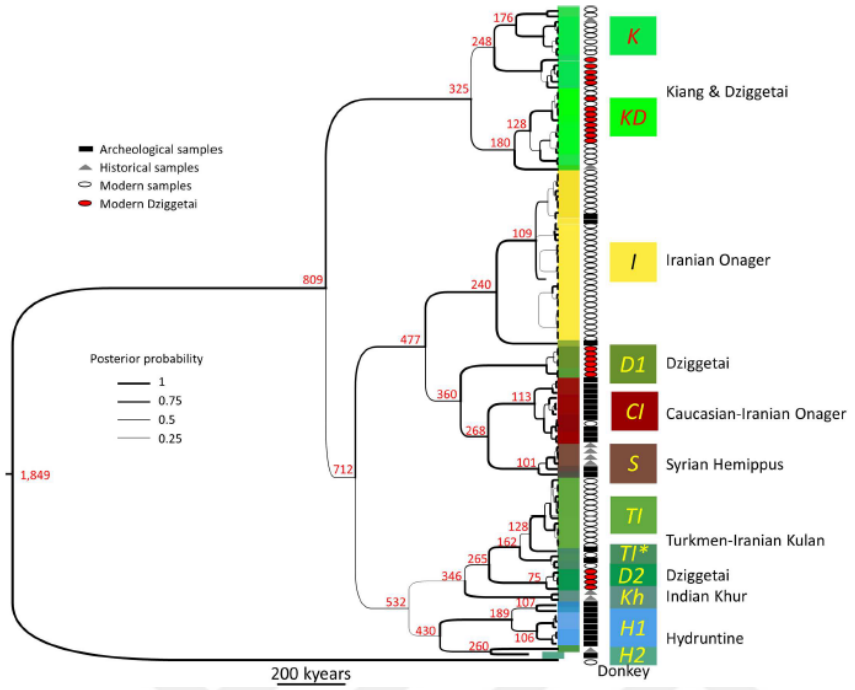


Figure 1.5: Wild ass phylogeny based on mitochondrial control region. Estimated mean node heights given in red and representing kya. Line thickness is proportional with posterior clade probability. After[5]

A near complete ancient mitochondrial genome extracted from a sample dated 21-23 kya was also published and reached the same result that hydruntines are closer to Asiatic wild asses[11]. Genetic data from this individual, excavated from Sicily, supported the idea that South Europe was a refuge in the Last Glacial Maximum (LGM). Meanwhile, none of these studies analysed nuclear genomic data from *E. hydruntinus*.

1.2.3 Other extinct species

The study of archaeogenomics is a valuable tool for understanding how extinctions occur. It allows researchers to analyze genomes from different historical or archaeological periods and identify patterns of genetic diversity that have been lost over time. Several studies have been focused on already extinct species and have provided ge-

nomic snapshots of mammalian populations on the verge of extinction, a selection of which I will present here.

One of these studies focused on blue antelope *Hippotragus leucophaeus*, which was the only large mammal species that went extinct in Africa in historical times, and their phylogenetic classification was not fully understood due to possible gene flow with other antelope species. In a recent study, researchers recovered two genomes from South Africa from samples aged 9500 and 200 years [81]. Armed with one archaeological and one historical genome, the authors confirmed previous studies reporting that blue and sable antelopes were sister taxa. Moreover, they presented evidence of gene flow from roan *Hippotragus equinus* into the blue antelope gene pool. The results also show that the genetic diversity of blue antelopes was much lower than sable and roan antelopes, indicating a lower population size. This estimation of decreased population size concorded with archaeological observations since blue antelope becomes scarcer after the Pleistocene-Holocene transition in the zooarchaeological record. Despite the low population size, they argued that the persistence of blue antelope in the Holocene was a sign that anthropogenic factors such as colonialism in the region were the main reason for blue antelope extinction.

The Sardinian dhole, *Cynotherium sardous*, was an endemic canid species of Sardinia and Corsica islands and went extinct in Late Pleistocene. Since its morphological features resemble a score of other canid species, researchers could not define its phylogenomic position on the canid tree and closest relative species by examining zooarchaeological features. An ancient DNA study based on c.21000-year-old remains revealed that Sardinian dhole was a distinct taxon, separated from other living canids from Eurasia, Africa, and North America[82]. The results indicated a divergence time between the Sardinian dhole and Asian dhole of about 885 kya. Nonetheless, these two populations experienced gene flow, which ended at 500-300 kya. This date concurred with the severance of Sardinia from mainland Italy by sea. Again, this extinct species showed lower genome-wide diversity compared to extant canid species, and this observation suggested the long-term effect of isolation and might contribute to their extinction.

The Sussemionus, *Equus ovodovi*, was one of the other extinct equids whose phy-

logenic position was has remained controversial. While this species was thought to go extinct in the Middle Pleistocene, emerging ancient DNA studies presented evidence that they persisted until the Late Pleistocene[10, 83]. Once ranging from Africa to Siberia, the extinction mechanisms of this species were also poorly understood. Researchers showed evidence of Sussemiones in the Holocene contexts of Northern China; radiocarbon dating evidenced an extinction date near 3.5 kya. They also confirmed that Sussemiones diverged from non-caballine horses, associated with the first equid dispersal into Eurasia from America, 2.3-2.7 mya, and possibly affected by the second wave of spread which brought true horses to the Old World. Results showed decreased levels of heterozygosity but no elevated levels of runs of homozygosity. The authors also noted that their work presented evidence of the persistence of species in cryptic population pockets[84].

Recently, researchers achieved a significant milestone in ancient DNA research- the retrieval of genetic material from individuals who lived over one million years ago. In this work, researchers isolated DNA from three mammoths (*Mammuthus* sp.) excavated from Eastern Siberia permafrost. According to the findings, there were two separate lineages of mammoths. One of these lineages eventually evolved into woolly mammoths, while the other served as the ancestors of mammoth species that inhabited North America. Furthermore, the Columbian mammoth of North America was a hybrid between these two lineages, with roughly the same amount of admixture proportions. The researchers also showed that most of the mutations associated with cold adaptation were already present one million years ago[85] Another study on mammoth decline revealed that accumulation of deleterious variants such as greater numbers of deletions and premature stop codons can be associated with genomic meltdown due to low effective population size[86].

1.3 Thesis Outline

Despite the developments in zooarchaeological and archaeogenomic methods, the population history of *E. hydruntinus* is still shrouded in mystery. This is mainly because of two factors; 1) since hydruntines were wild animals, entire skeletal remains are nearly non-existent, and 2) lack of complete genomes from which demographic

and ancestry could be established. Since hydruntines had mixed morphological characteristics and only mitochondrial fragments extracted from their archaeological remains, which can harbor a different population history than the nuclear genome[87], extinct *E. h. hydruntinus* deserves thorough archaeogenomic research.

This thesis aims to better understand the taxonomic status of *E. hydruntinus* by leveraging whole nuclear ancient genomes generated from Central Anatolian samples. To achieve this goal, I carried out ancient DNA isolation from 15 samples, initially thought of as horses, from two Central Anatolian excavation sites, the important Neolithic site of Çatalhöyük and the Iron Age settlement of Çadır Höyük. After the library preparation steps required for the next-generation sequencing, raw sequencing reads were aligned to the horse reference genome, and three samples stood out as they have enough genetic material to run downstream computational analyses. Phylogenetic analyses run on ancient mitochondrial sequences placed the equid remains of Central Anatolia in a cluster with *E. hydruntinus*, indicating their species identity as European wild ass, both on hypervariable region and whole mitochondrial sequence. Nuclear genome analyses presented four distinct groups in a multidimensional scaling analysis: 1) Donkeys and African wild asses from Somalia; 2) Asiatic wild asses from Central Asia, namely hemiones and kiangs; 3) subgroups of hemiones distributed in the Middle East, which are from Iran and Syria; and finally, 4) European wild ass (hydruntines) samples I produced from Central Anatolia. The lack of monophyly of hydruntines in mitochondrial sequences reported previously was not the case in nuclear genomes. The clustering of Anatolian hydruntines in shared drift analysis, suggesting a separate and distinct gene pool, weakened the idea that hydruntines were a subgroup of hemiones.

Since these three Anatolian samples were the most recent and genetically confirmed individuals of European wild ass, I also took a snapshot of several genetic markers which present evidence of a declining population. Runs of homozygosity test showed a genetic load similar to Somali wild ass, which was reported as a highly inbred population, and donkeys, known to have gone through a domestication bottleneck. Heterozygosity analyses confirmed this observation since Anatolian hydruntines had the lowest amount of heterozygosity, even lower than Somali wild ass and donkeys. Combined, I present evidence for a declining population, supporting the idea of dis-

creet subpopulations of hydruntines followed by habitat fragmentation.

With this thesis, my results provide a clearer picture of taxonomy and extinction mechanism of enigmatic *E. hydruntinus* species by reporting the whole genomes for the first time.

CHAPTER 2

MATERIAL AND METHODS

2.1 Archaeological samples

In this thesis work, I studied eleven Çatalhöyük and four Çadır Höyük animal skeletal samples identified as equid.

2.1.1 Çatalhöyük

Çatalhöyük is a tell site near Konya in central Turkey. First excavated by James Mellaart in the 1960s, a new project began in 1993 under the direction of Ian Hodder. At first the focus was on surface study and survey, followed by excavation from 1995 to 2017. The site is of international significance because it reached a large size (13.5 ha) at an early date, had a dense population (1,500 to 5,000 people), has rich symbolism and subfloor burials, and was occupied for 1,500 years; the Neolithic East mound dates from 7100 BCE to 5900 BCE with the Chalcolithic West Mound overlapping in time in the last quarter of the seventh millennium BCE and continuing on until 5600 BCE. Its upper layers have also yielded remains dating to Bronze and Iron Ages, and later periods[88]. The site was inscribed on the UNESCO World Heritage list in 2012.

This archaeological background information about Çatalhöyük was provided by Ian Hodder and Kamilla Pawlowska.

2.1.2 Çadır Höyük

Çadır Höyük is located in the Yozgat Province in north central Turkey, near the modern city of Sorgun. The mound rises 32 m above the floodplain and covers an area of approximately 5 hectares, with an additional occupation area north of the mound. Investigations have been ongoing at Çadır since 1994. A 2 x 2 m sounding on the mound's lower southern slope documented that occupation at Çadır extends back to at least the Middle Chalcolithic (Beta 146707, 5220–4940 BC [Cal BP 7170–6890]). The mound has been continuously occupied since that time, with final abandonment in the 14th century CE[89, 90, 91].

The Iron Age occupation extends across much of the mound, but is currently largely covered by the Byzantine settlement[92, 91]. The USS 4 trench, from which the four samples were excavated, is located on the upper southern slope. Rather than residential, this area was likely outside the main occupation area and was devoted to craft production including textiles, followed by what appears to be felt and leather production[93, 94]. The CDH08 (USS4 F128, FCN11128) and CDH010 (USS4 L198, FCN11130) samples derive from a stone wall (F128) set into a subfloor (L198) that have been phased to the Middle Iron Age. The CDH013 sample comes from a matrix phased to the transition between the Early and Middle Iron Age. The CDH022 (USS 4 L281, FCN 14393) sample comes from the fill in a pit (F212) that rests on the cusp between the very end of the Late Bronze Age and the beginning of the Early Iron Age.

This archaeological background information about Çadır Höyük was provided by Benjamin S. Arbuckle and Sharon Steadman.

2.2 Radiocarbon dating

Accurately determining the age of archaeological samples is crucial for answering archaeological inquiries. Radiocarbon dating, discovered in the 1940s, provides a wide range of timeframes that are useful for research, particularly for those studying the Holocene Period. Developed by Willard Libby from Chicago University, this

technique utilizes the radioactive carbon element that is continuously produced in the Earth's atmosphere. The heavier radioactive carbon, known as ^{14}C, combines with oxygen to form radioactive carbon dioxide, which is then absorbed by plants. Herbivores consume these plants, while carnivores acquire the radioactive carbon by consuming herbivores. However, once the supply of ^{14}C stops, such as after death, the amount of radioactive carbon decreases due to radioactive decay. The proportion of radioactive and non-radioactive carbon increases over time after death, providing a means of determining the age of the sample. Since the half-life of ^{14}C is 5,730 years, the upper limit of this technique is around 50,000 years[95].

It is important to note that while there is a proportional relationship between the age of a sample and its radioactive carbon content, local deviations can occur due to atmospheric conditions caused by climate change. To ensure accuracy and eliminate these deviations, calibration curves are created. The most recent calibration curve was published in 2020[96].

I prepared all three equid samples that showed wild ass genetic signatures for radiocarbon dating. For each sample, approximately 3 grams of bone or tooth material was cut using non-carbon-based discs attached to a Dremel tool and sent to the TÜBİTAK-MAM (Gebze, Turkey) AMS facility for carbon-14 dating. The dates were calibrated using the R package "Bchron" using the IntCal20[96] curve.

2.3 Ancient DNA isolation and sequencing

One of the characteristics of ancient DNA is having short fragment lengths. Apart from modern DNA extraction techniques, ancient DNA isolation requires specific procedures to capture as many short fragments as possible. In the past, ancient DNA studies primarily used the phenol extraction method[52]. However, since the latter half of the 21st century, it has been proven that techniques utilizing the binding of genetic material to a silica column are much more effective for isolation[56, 57, 54]. In this thesis work, I applied one of the most common ancient DNA isolation methods with minor modifications[55] (Figure 2.1). All pre-PCR experiments were conducted in the METU Ancient DNA Clean Room, in a dedicated aDNA research laboratory

located in a different building from the post-PCR laboratory.

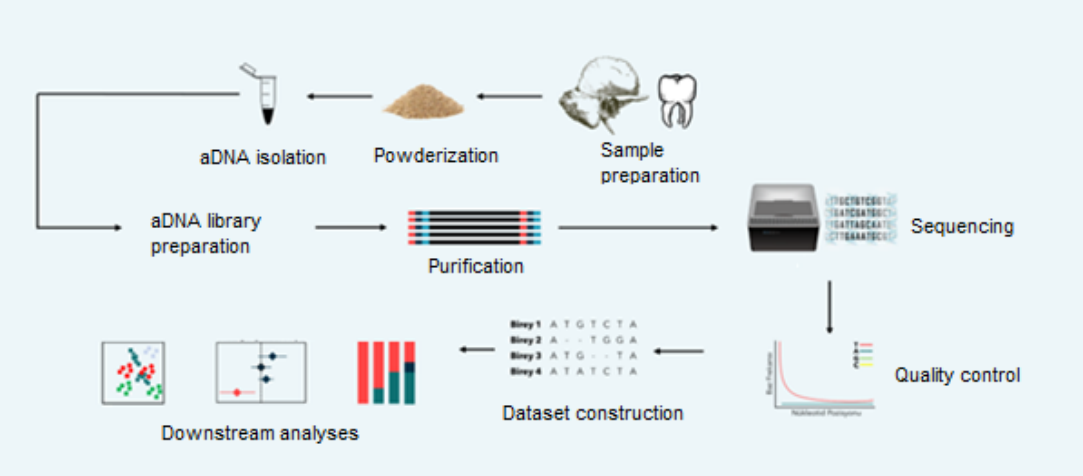


Figure 2.1: Summarised workflow of ancient DNA wet lab process.

In brief, I cleaned the surfaces of archaeological samples with a damp paper cloth, cut 100-200 mg of bone or tooth sample using discs attached to a Dremel, pulverized them with mortar and pestle, and transferred them into 2 mL screw-top tubes. Each sample powder was treated with 1 mL extraction buffer (0.45 M EDTA and 0.25 mg/mL Proteinase K) in a 37º C rotating incubator for 18 hours. This step breaks the bone matrix and releases DNA and proteins into the solution. 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, 90 mM Sodium Acetate) to bind DNA fragments to Qiagen (Germany) spin columns. I washed DNA fragments twice with a washing buffer and collected fragments after two consecutive elution steps into 50 µL Qiagen (Germany) EB buffer.

I procured ancient genetic material and subsequently generated double-stranded sequencing libraries according to the Meyer and Kircher protocol[97]. To this end, I turned the sticky ends of ancient DNA fragments into blunt ones using DNA polymerase (ThermoFisher, USA) using 20 µL isolated DNA and cycling 15 minutes in 25ºC and 5 minutes 12ºC. Then, adapter regions that would be recognized by the sequencing platform and indices used for demultiplexing step were added using DNA ligase (ThermoFisher, USA). The reaction took place in 22ºC incubation in 30 minutes followed by adapter filling step. Finally, libraries were washed using Qiagen

(Germany) wash buffer included in MinElute PCR kit three times to ensure their purity. This crucial step is essential because the unprocessed isolate fragments lack the universal adapter regions required to interact with the Illumina sequencing platform lane. Additionally, I added demultiplexing indices, which are 7 bp-long sequences, to the fragments during this procedure. Blunt-end ancient DNA libraries were prepared using Meyer and Kircher protocol[97] with a single indexing approach, i.e, only one index sequence per individual used. All experiments on Çadır Höyük and Çatalhöyük samples were conducted on different days with fresh reagents prepared for each archaeological site.

Libraries were sequenced on the Illumina HiSeq 4000 platform in SciLife facility in Sweden, in paired-end fashion (Appendix B). Using 47 thousand - 79 million reads (median 20 million) obtained in initial low-coverage sequencing, I 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. I further sequenced cdh008 three times, cdh010, and chh003 one more time (Appendix C). Vandela K. Lagerholm assisted with preparation and sending the samples to the SciLife facility for sequencing.

2.4 Data preprocessing

In this section I first provide background about the data preprocessing steps. I then describe the technical details such as the software and parameters used.

The Illumina sequencing output was converted to *.fastq* text files at the SciLife facility in Sweden. These files contain information about total sequencing reads. This file was then demultiplexed, using 7 bp long indices, at the same facility to obtain individual-based forward and reverse reads. This approach enables researchers to sequence multiple samples in one lane between 20 and 100.

Demultiplexing is in silico separation of the raw reads according to indices which are unique for all samples, and yields a separate pair of *.fastq* files for each library. Following this, adapter and index sequences are removed using programs such as *AdapterRemoval*[98] to obtain information only from sequenced individuals. These

are then subject to quality checks using softwares such as FastQC (www.bioinformatics.babraham.ac.uk /projects/fastqc/). The quality check procedure consists of several basic statistics such as base and read quality scores, fragment length distribution, GC content distribution, and amount of missing bases. Files passing the quality check are subjected to an algorithm (for example *BWA*) that ascertains read positions on the reference genome, and this step is called "mapping." More stringent parameters yield lower levels of mapped reads, while relaxed parameters can map as many reads as possible. Since ancient DNA fragments are short and include post-mortem damage (PMD) patterns, a balanced set of parameters has been established in the ancient DNA community.

Since a PCR reaction occurs before sequencing and this step introduces technical duplicates, the duplicated reads with the same starting and ending positions are removed to eliminate technical positive hits. Moreover, reads shorter than 35 bp length are also removed due to uncertainty of their exact position on the genome, with multiple hits that might correspond to repeat, telomeric and centromeric sequences. The length threshold depends on the sample; while 35 is a good practice, the threshold for older samples which might have a great importance might be reduced to 30 or 25. The last step in preprocessing consists of eliminating reads with a high mismatch-to-length ratio. This step is usually applied to ensure authentication of the reads in human ancient DNA studies. Still, it might present a problem of removing genuine reads when the organism of interest is mapped to a reference genome of a different species, such as in this thesis work. I inspected the proportion of the reads eliminated in this step for deep sequencing libraries, and the results showed that only 2 of the six libraries have a higher than 1% elimination rate, and none of them belonged to the highest quality sample, cdh008 (median=0.86%).

Raw sequencing reads demultiplexed in the SciLife facility were transferred to Ankara and processed following previous work[99, 100] by Kıvılcım Başak Vural. First, adapter sequences were removed, and paired-end sequencing reads were collapsed using *AdapterRemoval* (v. 2.3.1), requiring a minimum of 11 bp overlap between pairs[98]. Collapsed .fastq files were aligned to the horse reference genome (EquCab2.0; including 31 autosomal chromosomes, the X chromosome, and the mitochondrial genome) [101] using the *aln* module of *BWA* software (v. 0.7.15)[102]

with the parameters "-n 0.01 -o 2" and disabled the seed with "-l 16500". While BWA software has an alternative module called *mem*, which is reported to be faster than *aln*[103], another study showed that *mem* parameterisation effected precision negatively[104]. We thus opted to use the *aln* module. All alignment files of the same individual were merged using *SAMtools* (v. 1.9) *merge*[105], PCR duplicated reads were removed using *FilterUniqueSAMCons.py*[97] and reads with length <35 bp were discarded. Reads with a mismatch to fragment length ratio >10% were also removed (Appendix B and C).

2.5 Authentication of genetic data and trimming

One of the two main characteristics of ancient DNA material is post-mortem damage patterns accumulated at both ends of the reads[53]. After death, the DNA chain undergoes a deamination reaction at several points, most commonly the deamination of cytosine to uracil. This process destabilizes the chain and results in the fragmentation of the DNA. While sequencing, uracil bases are read as thymine, and digital files contain thymine bases in positions expected cytosine. Although it looks like a setback, this feature is one of the safety checks utilized in assessing reads as endogenous, i.e., belongs to the organism of interest.

I verified the authenticity of the genetic data using *PMDtools* (v. 0.60)[106], which quantifies post-mortem damage (PMD) signals at both ends of mapped reads, characteristic of ancient DNA[53]. PMD profiles for each sample were generated using the PMDtools "–deamination" parameter. Deamination signals of cytosine to thymine transitions at 5'-most positions were 0.31, 0.29, and 0.36 for cdh008, cdh010, and chh003, respectively, which were above the minimum expected PMD signal value of 0.20 (Figure 3.2). Average read lengths were 74 bp for cdh008, 75 bp for cdh010, and 67 bp for chh003, again supporting authenticity[107]. The other authentication methods, such as contamination estimates based on mitochondrial sequences, were inapplicable in this thesis since they require a reference mtDNA diversity panel.

Reads obtained from libraries were trimmed ten bp from both ends using *trimBAM* command of the *bamUtil* software[108] to remove residual deamination.

2.6 Complete mitogenome sequence

I generated a complete mitogenome sequence of our highest coverage sample cdh008 (coverage: 140.04X) by consensus calling of the most common bases using *Geneious Prime* (v. 2022.2) with default parameters. Most of the unidentified bases (Ns: 2% of the whole sequence) were observed near the end of the sequence. This occurs due to the circular nature of mitochondrial DNA and can be overcome by adding the first 100 bases to the ending of sequence and the last 100 bases to the beginning, in *FASTA file*. The sequence was deposited to the NCBI Genbank with accession no: OP448588.

2.7 Taxonomic identification and sex determination

I investigated the taxonomic identity of these three samples using the *Zonkey* pipeline implemented in *PaleoMix* (v. 1.2.14)[109]. Zonkey generates a PCA[110], TreeMix[111] and mitochondrial phylogeny graphs based on a reference panel consisting of horse, zebra, donkey, and Asian wild ass (hemione) genomes. SNP counts used in Zonkey analysis were 163,921 for cdh008, 158,448 for cdh010, and 162,739 for chh003. All three samples were clustered together with the *Equus hemionus onager* and *Equus kiang* (Asian wild ass) group, confirming their Eurasian wild ass status. Sex estimation was also performed using Zonkey, based on autosomal versus X chromosome coverages. For chh003, the X chromosome coverage was approximately half the autosomal coverage, suggesting this sample was male. At the same time, the data for cdh008 and cdh010 showed similar coverage in autosomes and the X chromosome, meaning they were female.

2.8 Phylogenetic analyses of mtDNA data

Phylogenetic analysis was carried out using mtDNA consensus sequences generated from BAM files using *angsd* (v. 0.934)[112] with “-doFasta 2”, “-minQ 30”, “-minMapQ 30” and “-setMinDepth 3” parameters, following software’s manual (www.popgen.dk/angsd/index.p While "-minQ 30" and "-minMapQ 30" values set quality thresholds the same as the whole genome mapping procedure, "-doFasta 2" prints out the most common allele

in a given position. The "-setMinDepth 3" parameter was incorporated to avoid sequencing errors in positions with less than three reads.

MEGA X[113] was used to construct phylogenetic trees using maximum likelihood (ML) and neighbor-joining (NJ). Neighbor-joining is a clustering method that is based on genetic distances[114]. However, different substitution rates among the sites lower the accuracy of NJ[115]. On the other hand, maximum likelihood trees are based on models which are inferred from sequencing data[114, 116]. While the ML method requires higher computational power, it cand yield more reliable results, assuming there is sufficient information available.

To this end, a 361 bp fragment of the D-loop in the control region, which was used by previous studies, of the mitochondrial DNA and a 13,042 bp long reference sequence of the whole donkey mitochondria were used (Genbank). For the maximum likelihood analysis using 138 D-loop sequences (Appendix D), the substitution model was chosen as Hasegawa-Kishino-Yano (HKY) with a Gamma-distributed rate (no invariant sites) using *jModeltest* (v.0.1.1)[117, 118]. HKY model[119], incorporating variable base frequency rates, is different from the other models such as general time reversible (GTR)[120], which incorporates a symmetrical substitution matrix. A maximum composite likelihood model was used for the neighbor-joining tree construction. The nodal support was evaluated by 1000 bootstrap replicates for both methods. For the median-joining (MJ) network, 121 partial D-loop samples were trimmed to 249 bp based on the shortest fragment in the dataset. The MJ network was constructed with *NETWORK* (v.5) maximum parsimony post-processing (http://www.fluxus-engineering.com).

The mitochondrial DNA divergence times of chosen equid species were estimated with the full mitochondrial DNA sequences of 24 individuals using Bayesian inference with *BEAST* (v. 1.8.4)[121]. For this analysis, the best substitution model was found to be GTR with a Gamma distributed rate and invariant sites using *jModeltest* (v. 0.1.1). An uncorrelated relaxed clock with a lognormal distribution and Yule speciation was assumed. The convergences were visualized using *Tracer* (v. 1.6)[122]. The best tree was found and drawn using *TreeAnnotator* (v. 1.8.4)[121] and *FigTree* (v. 1.4.3)[122]. *Equus caballus* was used as an outgroup, and the molecular clock

was calibrated using a normal prior of 4.25±0.125 Mya for the divergence of *E. caballus* from other equid species[83]. The analysis was run for 600,000,000 steps, logging every 60,000 states and discarding the first 10% states as burn-in, resulting in effective sample sizes (ESS) greater than 1000 for all traces. The mitochondrial divergence between African and non-African asses was estimated as 3.57 [1.18-4.23] mya, which is higher than but includes previous estimates (e.g., 2.62 mya)[83].

Kanat Gürün and Eren Yüncü provided support for the phylogenetic analyses.

2.9 SNP panel construction

Demographic analyses incorporating autosomal sequences require a variation panel, preferably generated using an adequate number of individuals and populations. As there is currently no comprehensive panel on donkey and wild ass species, I have created a series of variation panels to assess how effectively they represent the wild ass gene pool.

In my efforts to create such a panel, I first used 12 modern-day and high coverage (>5x) genomes[123, 23, 124, 125], of which six were domesticated donkeys (*Equus asinus asinus*) and six were Asian wild asses (three *Equus hemionus hemionus*, one *Equus hemionus onager*, two *Equus hemionus kiang*) (Appendix D). The *E.h.onager* sample was traced back by Eva-Maria Geigl in their lineage to confirm its status as an onager. It revealed that this individual's (studbook n.818) father (studbook n.682) was an onager with known origin while the mother was a non-specified onager. Their father's (studbook n.682) father was a non-specified onager, and their mother was a zoo onager (studbook n.618). The lineage of n.618 individual was captured from Iran in the wild between 1955-1973, 6-7 generations before the birth of n.682. Two horse samples were included as an outgroup[64, 63].

First, I performed de novo SNP calling using *angsd* (v. 0.934)[112] from the donkey and wild asses, separately this time. I limited de novo SNP calls to positions with a minimum depth of 4 and a maximum depth of threefold of average coverage. I further required that the SNPs were non-missing across all the individuals. A total of 31,592,013 and 38,029,882 SNPs were thus called from donkeys and wild asses,

respectively. Limiting these to autosomal positions, applying a minor allele frequency filter of MAF > 0.1 separately which requires at least two copies of an allele in this sample, a Hardy-Weinberg filter of "1e-30," and keeping only biallelic transversion positions yielded 2,146,416 SNPs, which I refer to as the main SNP panel.

To confirm results from this main SNP panel using a published alternative, I used a second SNP panel of equid variation as generated by Bennett et al[8]. This panel was derived from the panel used in the *Zonkey* algorithm that was established from a single genome of each of 9 Equids species or subspecies, including horses, African and Asiatic wild asses, and zebras[109]. This set of SNPs had been previously filtered to eliminate all SNPs that fell within repeats in the reference genome based on the assumption that repeats are more prone to artifactually mapping with ancient DNA fragments, in particular when mapping to the genome of a distantly related species[8]. I further filtered this dataset to keep positions polymorphic in donkeys and wild asses and used only transversions, leaving 2,408,064 SNPs. I refer to this as the secondary SNP panel.

2.10 Genotyping and dataset construction

I genotyped both modern-day and ancient genomes (Appendix D) using both SNP panels. All modern-day individuals were genotyped using *SAMtools*[105] (v. 1.9) *mpileup* module in a diploid fashion. Ancient individuals were genotyped using *sequenceTools* (v. 1.4.0.5) (https://github.com/stschiff/sequenceTools) using the option "-randomHaploid." The *pileupCaller* module of this program was run on ancient samples, genotyping only the positions in the SNP panel. Ancient genomes were pseudohaploidized by randomly choosing a single read and using its allele as the individual's genotype. This was carried out due to avoid genotyping biases caused by differences in terms of coverage among samples and sequencing errors. Genotypes of modern-day and ancient individuals were merged using *PLINK* (v. 1.90)[126] and *EIGENSOFT* (v. 7.2.1)[110] to form the final versions of the dataset. I thus prepared two datasets: one I refer to as the main dataset and the other as the secondary dataset. These were used in all population genomic analyses.

2.11 Outgroup f3-statistics

A way to measure genetic similarities between individuals and between groups is through demographic analysis using outgroup-f3 statistics[127]. In this analysis, trios are generated in the form of f3(Outgroup; PopA,PopB) (Figure 2.3). The genetic affinity is then calculated between the outgroup and PopA and PopB based on differences in shared allele frequencies. Since I used an outgroup, the f3 value is always positive. Normalization based on heterozygosity is often applied[128]; however, this approach yields enormously high f3 values when the outgroup is inbred, such as a domesticated horse. To overcome this obstacle, I incorporated "–f3 vanilla" parameter which calculates f3 value without normalization (f3=(p3-p1)(p3-p2), where p1, p2 and p3 represent allele frequencies of population A, population B, and the outgroup, respectively [129]. The higher values point to elevated affinity while lower values practically mean PopA and PopB are distant.

I performed outgroup-f3 statistics on the autosomal chromosomes dataset using *popstats*[129] with "–not23" and "–f3 vanilla" parameters and using the *PLINK* (v. 1.90) .tped/.tfam file format. Analyses were run in the forms of f3(Horse; Anatolia, Modern), f3(Horse; Modern, Modern), f3(Horse; Anatolia, Anatolia), and f3(Horse; Anatolia, AncientHemione) in a pairwise manner, where Horse represents two horse genomes used as outgroup, Anatolia represents one of three Anatolian ancient wild ass individuals reported in this article, Modern represents a modern-day donkey, Somali wild ass, or Asian wild ass genome, and AncientHemione represents the three published ancient hemione and three ancient hemippe genomes[7, 8]. A heatmap graph was generated from the f3 results using the *heatmap.2* command of the R package *gplots* (https://cran.r-project.org/web/packages/gplots/index.html).

2.12 Multidimensional scaling

To derive a distance measurement from f3 values, one can simply subtract the value from 1. This will yield distance values that can be effectively utilized for creating a distance matrix. This distance matrix then can be used in generating a multidimensional scaling (MDS) plot. This analysis is one of the dimension reduction techniques,

such as principal component analysis, that allow researchers to utilize multidimensional data into two or three-dimensional representations.

A genetic distance matrix was generated by subtracting pairwise outgroup-f3 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 visualized in R.

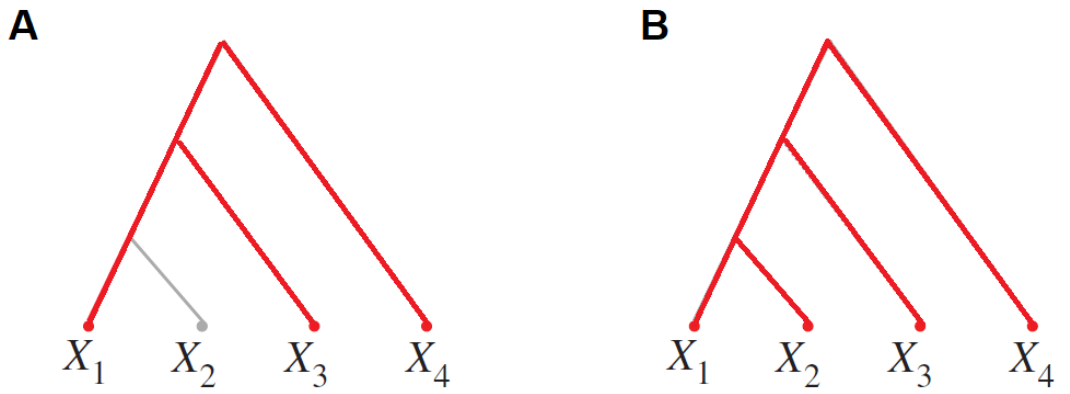


Figure 2.2: Schematic representation of A) outgroup-f3 and B) D-statistics with which I carried demographic analyses. In this schema, X1, X2, and X3 shows test populations while X4 represents outgroup. Allele frequency differences on branches calculated and multiplied. Gene flow between X3 and X2 yields negative D while between X3 and X1 yields positive D. Adapted after[6].

2.13 D-statistics

Based on the same logic with outgroup-f3 analysis, D-statistics are calculated between four individuals or populations in the form of D(Outgroup, PopA; PopB, PopC)[128]. Since the first population is an outgroup, which should be equally distant to PopA, PopB, and PopC, this analysis quantifies the affinity between PopA and two other populations. This analysis traces ancestral (A) and derived (B) allele frequencies between three populations if the fourth population is an outgroup that is equally distant to the other three. Two scenarios, 1) ABBA, where X2 and X3 share derived allele, and 2) BABA, where X1 and X3 share derived allele, are considered. Then D statis-

tic is calculated as D = (nABBA-nBABA)/(nABBA+nBABA). D-statistic results are between -1 and 1 and are calculated from overlapping SNPs in all four populations (including the outgroup). Negative results indicate an affinity between PopA and PopB relative to PopC, while positive results suggest an affinity between PopA and PopC relative to PopB. The presence of this similarity suggests that a gene flow may have occurred between different populations. Furthermore, distant groups' common ancestry or a historical connection can produce significant outcomes. The standard error (SE) of D is calculated using a block-jackknife approach. In this, the genome is divided into several blocks with the same length, preferably larger than the extent of linkage disequilibrium. Variance in D scores is then calculated in each block, multiplied by the number of blocks and square rooted. The result's significance is then inspected by calculating a Z-score, the ratio of D-statistics to SE. As a rule of thumb, any Z-score above 3 means a significant result. A Z-score below three can be interpreted as PopA being equally distant to PopB and PopC.

Gene flow and genetic affinity between ancient and modern groups were inspected with D-statistics using main and secondary dataset in *EIGENSTRAT* file format (.geno / .ind / .snp), running the *qpDstat* module of *AdmixTools* (v. 7.0)[128] with default parameters. The analysis was run on all possible trio combinations, using the horse as an outgroup. I limited the analyses to pairs with with at least 2000 overlapping SNPs (for example, excluding the historical hemippe individual, Hm_1892). The significance of each comparison was inspected using the Z-score (the ratio of the D-statistic to the standard error, the latter estimated using a jackknife procedure using 5 Mb blocks across the genome). Results with absolute Z-scores > 3 (corresponding to roughly p<0.001) were defined as nominally significant. Results were visualized using the R package, *ggplot2*[130].

2.14 Divergence time estimation

In order to estimate the divergence time between the hydruntines and Asian wild assess, I used the F(A|B) statistic[131], which calculates the probability of observing a derived allele in individual A at the sites where individual B is heterozygous. This approach was taken because an alternative method called TT[132] depends on

diploid genomes. Since the ancient genome I generated was not suitable for diploid genotypinh, the TT method was not used. For this case, individual A corresponds to best quality hydruntine cdh008, and individual B corresponds to the kiang sample Kia2[125], representing the Asian lineage. In total, n=2,626,614 heterozygous positions were determined in the Kia2 genome. The alleles were sorted as ancestral or derived according to the horse genome (EquCab 2.0). Genotype calling in cdh008 was carried out using *samtools* (v. 1.9) *mpileup*[105], with a minimum read and base quality of 30. The probability was calculated by choosing a random allele at each site and recording whether it was derived or ancestral. This probability was then computed for different divergence times and population sizes using the formula $e^{T/(2N)}/3$ [133] and simulations using *msprime* (v. 1.0)[134], with a generation time of 8 years, a mutation rate of 7.242×10^{9} and a recombination rate of 1cM/1Mbp[63]. I simulated 100 chunks of 1Mb and calculated the mean F(A|B). Using the formula H=4Neμ, I estimated N_e for Kia2 as 62,319.3. Expected F(A|B) values were calculated for divergence times ranging from 100kya and 2.1mya and population sizes ranging from 5k to 80k. I determined the expected F(A|B) values under various split times that correspond to the observed F(A|B) value to find the estimated population split time. For comparison, I also estimated split times of *E. h. onager* and *E. a. asinus* each from *E. kiang*, again using F(A|B). I also repeated the F(A|B) calculation using only transversions (n=811,814) instead of all SNPs, which resulted in a highly similar estimate (0.145 using all SNPs vs. 0.149 using transversions only). I note that using transversions limits the chance of homoplasy and misidentification of the ancestral state.

Divergence time estimates based on nuclear sequences were done with assistance of Gözde Atağ.

2.15 Runs of homozygosity

A reliable method of identifying inbreeding due to small population size and/or mating among closely related individuals is to examine homozygous tracks on the genome, known as runs of homozygosity[135, 136]. In this analysis, a set of SNPs are inspected to determine whether they form a string of homozygous states, and due to the

lower quality of ancient samples and sequencing errors, some level of heterozygosity is allowed. It is a good practice to limit the positions depth-wise since lower levels of depth are more prone to sequencing or genotyping errors. In runs of homozygosity analysis, elevated amounts of long tracks of homozygous SNPs tend to signify recent inbreeding, i.e., the mating of closely related individuals. High levels of short tracks also indicate a small or shrinking population[135]. Analyzing runs of homozygosity in this thesis elucidated the population trajectory and presented evidence of a decrease in variation in the gene pool.

Runs of homozygosity (ROH) of representative samples for each group (Appendix D) of *E.a.asinus*, *E.h.kiang*, *E.h.onager*, *E.h.hemionus*, *E.a.somalicus* and "cdh008" for the Anatolian wild ass, (or E.hydruntinus) were calculated from equal coverage BAM files. For this, I downsampled all BAM files to 6.38x to match the lowest coverage case, cdh008. To calculate ROH, I used *PLINK* (v. 1.90)[126] with the following 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". In order to overcome issues arising from the low quality of ancient individual (cdh008), 3 heterozygous positions were allowed within a homozygous track. The resulting data were merged into a single file and visualized using R. A recent work has shown that ROH can be called efficiently using this approach from >5x genomes[135].

2.16 Genome-wide heterozygosity estimation

Decreases in genetic variation and shrinking population sizes tend to manifest their effects on heterozygosity. A reduction in heterozygosity is often associated with a declining population[137] which can also be sensitive to loss of diversity through genetic drift. Coupled with the results from runs of homozygosity analysis, I conducted heterozygosity calculation based on polymorphic sites with minimum depth of 5. I also limited this calculation to transversion sites in order to eliminate PMD-based false heterozygosity calls since the PMD pattern of cdh008 extended beyond the first ten bases.

In order to estimate genome-wide heterozygosity for modern-day and ancient equid genomes, I first downsampled BAM files from modern-day equids, namely three *E.a.asinus*, one *E.a.somalicus*, one *E.h.onager*, two *E.h.hemionus* and one *E.h.kiang* to 6.38x coverage, i.e., the coverage of the Anatolian cdh008. Using these downsampled BAM files and that of cdh008, I called de novo variants in each genome at all positions with 5-13x depth using *samtools* (v. 1.9)[105]. I then determined heterozygous loci for each genome and further filtered these sites only to keep transversion positions to avoid PMD-related false heterozygosity calls. Finally, the total number of heterozygous sites was divided by the number of all positions with 5-13x depth in the downsampled genomes. Results were visualized using the R package, *ggplot2*[130].

CHAPTER 3

RESULTS

I performed ancient DNA extraction on 15 equid tooth and bone samples excavated from two Central Anatolian sites, eleven from Çatalhöyük and four from Çadır Höyük (Figure 3.1A). I generated genome-wide genetic data from these samples using shotgun sequencing (Appendix B) and obtained 0.02%-11.35% (median=0.07%) of endogenous DNA by mapping to the horse reference genome (version EquCab2.0)[101]. I chose three samples (cdh008, cdh010 and chh003) with >5% endogenous DNA all of which exhibited typical postmortem damage (PMD) signals for further sequencing (Table 3.1, Appendix C, Figure 3.2). I thus produced three genomes with nuclear coverages 6.38x, 0.72x and 0.57x, respectively (Table 1, Appendix C). Radiocarbon dating of the skeletal material placed all three individuals within the early/mid 1st millennium (Iron Age) in Anatolia (Table 1, Figure 3.1C). I also generated a near-complete mitogenome sequence of the highest coverage (140.04x) sample cdh008 (Material and Methods).

In table 3.1 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. Radiocarbon (C14) date intervals (95%), calibrated using IntCal20[96] curve, are given as calibrated years Before Present (BP).

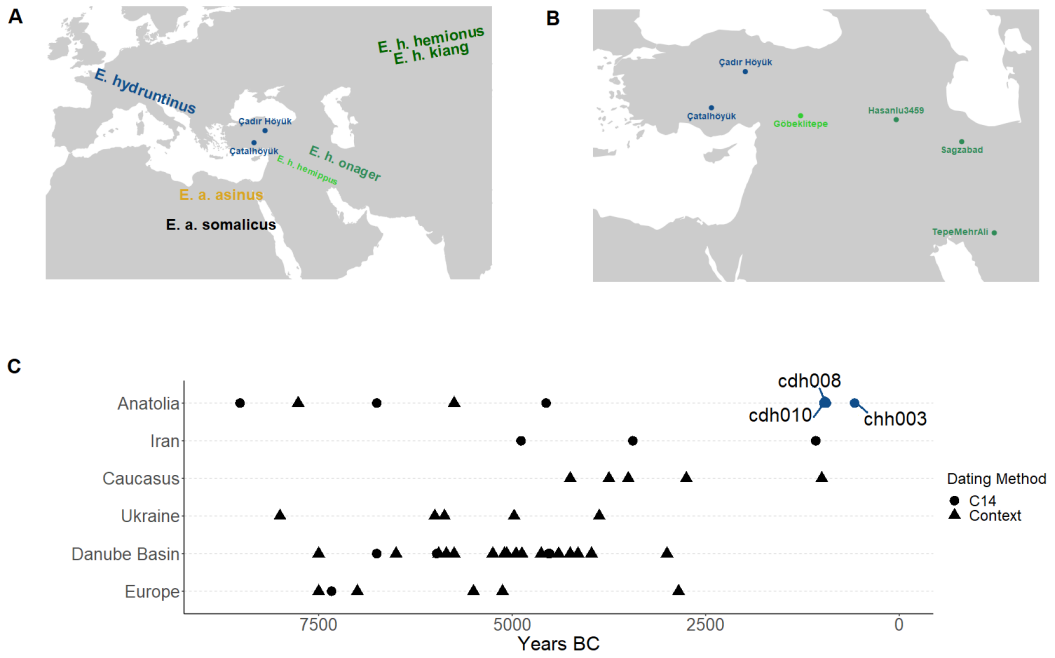


Figure 3.1: A) The map shows the geographical locations of Çadır Höyük and Çatalhöyük (the excavation sites where the wild ass remains analyzed in this study were recovered), as well as the approximate dispersal areas of extant and extinct ass taxa. B) Shows the locations of three ancient hemione individuals and one hemippe individual the partial genomes of which were previously published[7, 8]. C) Timeline of dated *E. hydruntinus* remains published in the literature including samples reported in this study (adapted after Crees and Turvey[4]).

Table 3.1: Genome information and radiocarbon dates of the three newly sequenced ancient Anatolian equids.

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

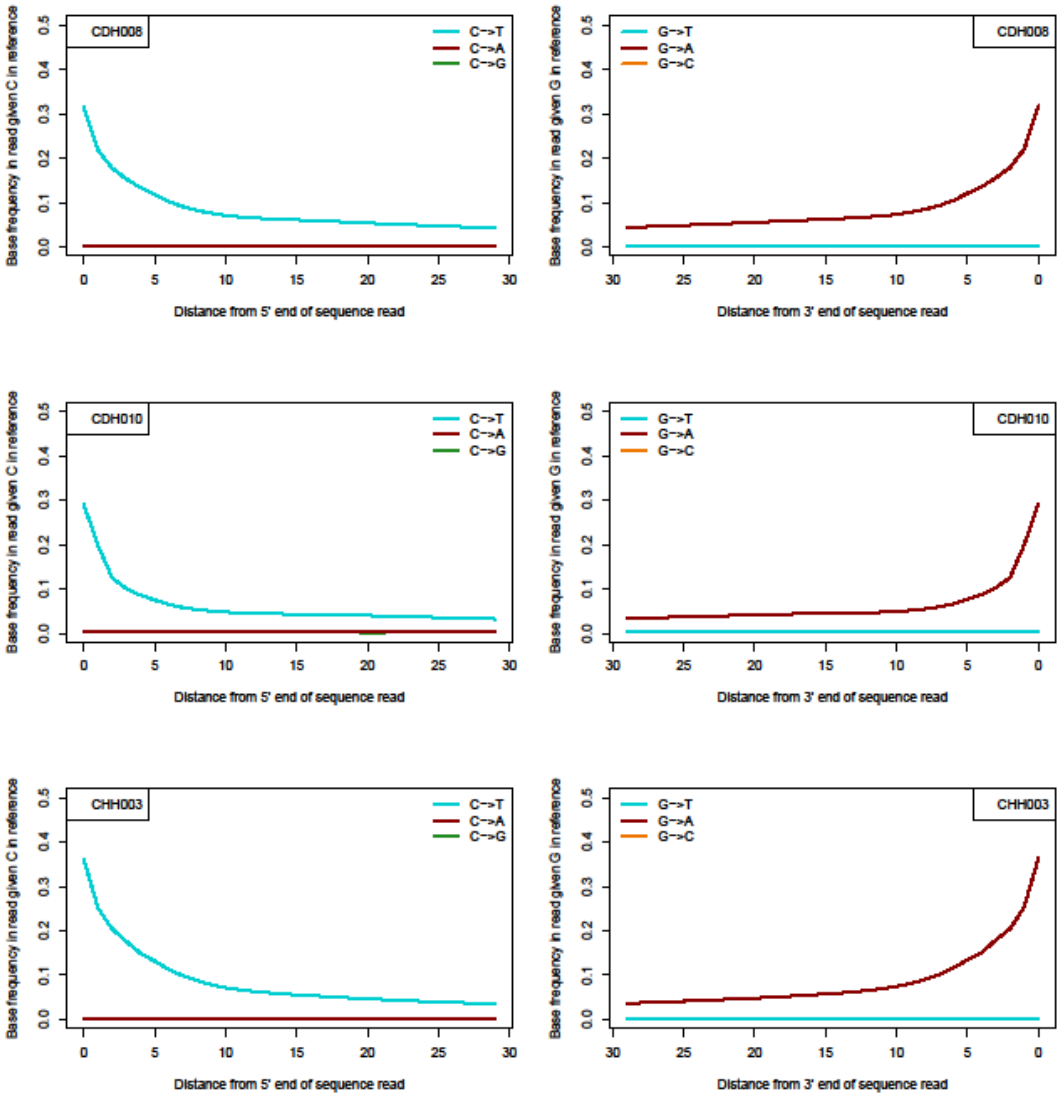


Figure 3.2: Post-mortem damage signals of cdh008, cdh010 and chh003.

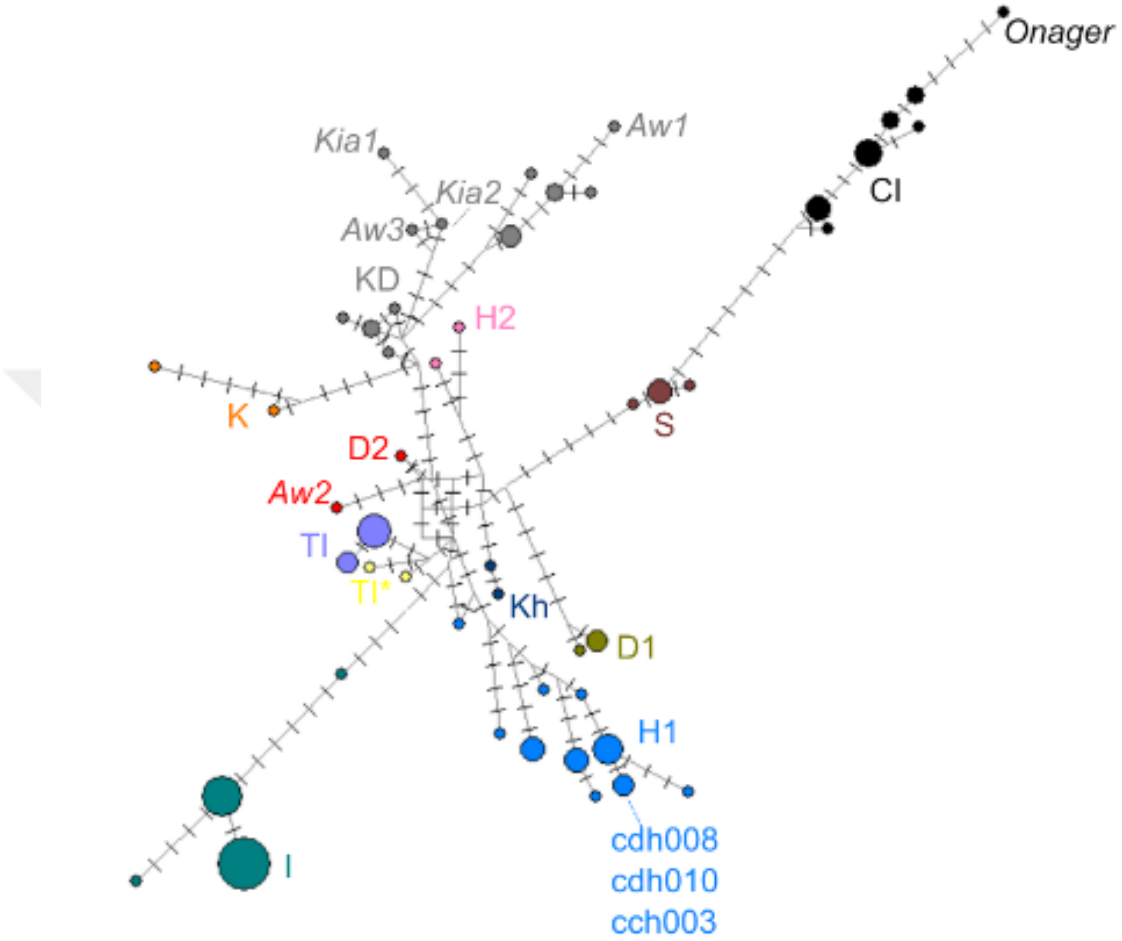


Figure 3.3: Mitochondrial DNA median joining network of the Eurasiatic wild asses from Central/East Asia, Iran, India, Syria, Caucasus, Anatolia, Balkan and Western Europe[9, 10, 5, 11], and sequences reported in this article generated using 249 bp diagnostic region of D-loop (Appendix D, Methods). *H1*, which colored with light blue here, represents the main cluster of hydruntine haplotypes. (Capital letters denote haplogroups and names with italic font denote sample names, see Appendix D.)

3.1 Genetic Analyses Assign the Iron Age Anatolian Wild Ass Individuals to *E.h.hydruntinus/E.hydruntinus*

During the Bronze and Iron Ages, horses and the domestic donkey were common in Anatolia, whilst hemiones were present in southern-eastern Anatolia, and hydruntine populations might also have survived in Anatolia[72] (Appendix A). Due to a lack of sufficient diagnostic characteristics, the three specimens were each identified to the level of the genus *Equus* but no attempt was made to assign species based on morphological criteria (Appendix B). As a first step towards characterizing Anatolian individuals, I used the Zonkey pipeline, which performs taxonomic classification of equid genomic data by comparing them with a reference genome panel[109]. Zonkey classified all 15 individuals examined in this thesis as "Asian wild ass-related" (Appendix B). This algorithm also identified two samples as males (including chh003, one of the three samples deep sequenced) since their X chromosome coverage was half of the autosomal chromosome coverage. The other ten samples were identified as females, including cdh008 and cdh010, which were further sequenced; they had similar coverages on both X and autosomal chromosomes (Appendix B). I observed a genetic sex bias in favor of females (p=0.03). The genetic sex of the three individuals was undetermined.

I then compared mtDNA data from these three Anatolian equids with published partial mtDNA sequences. In total, sequences from 82 hemiones[5, 123, 10, 125] and 21 morphologically-identified hydruntine individuals[5, 11, 10, 98] as well as 7 other equids[5, 23, 125] were collected (Appendix D). All three new sequences were a sub-branch of the *H1* haplotype clade, with no difference in diagnostic region [as defined by [5]] (Figure 3.3). *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 of Neolithic and Bronze Ages, c.7,850-2,500 BCE[5]. Together, these results indicate that the mitochondrial lineage of the three Anatolian equids belonged to wild asses, and was more closely related to hydruntines than to hemiones.

My colleagues and I investigated this further using complete mitochondrial DNA sequences, taking advantage of a recently published almost complete mtDNA sequence

of a morphologically-identified *E.h.hydruntinus*/*E.hydruntinus* specimen[11]. I compiled full mtDNA sequences from a total of 9 ancient and modern-day equids and outgroup lineages (Material and Methods, Appendix D). Again, Bayesian trees constructed using BEAST (Figure 3.4) clustered the Anatolian wild asses with the published hydruntine with full support.

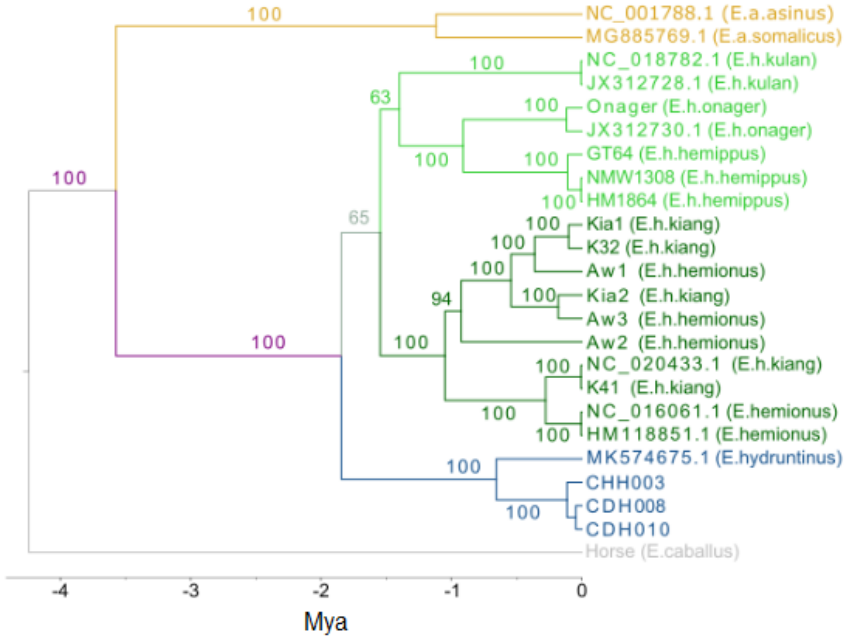


Figure 3.4: A Bayesian phylogenetic tree constructed using a 13,042 bp long consensus sequence of the whole mitochondrial DNA among various equid lineages. Numbers on the internal nodes show posterior probability support. X axis denotes time in millenial scale.

Mitochondrial DNA phylogenies may be incongruent with the overall phylogenetic history of a species due to drift or selection[87]. I hence explored the whole genome data for evidence that the three Anatolian wild asses might represent a hydruntine. For this, I compiled a list of c.2 million autosomal SNPs positions using published asinus and hemionus genomes (Material and Methods, SNP panel construction). I genotyped the three Anatolian wild ass genomes, as well as 13 published modern-day ass and donkey genomes at these positions (Table 1, Appendix D). For comparison, I further included three published ancient onager genomes from Iran (Tepe-MehrAli, Hasanlu3459 and Sagzabad)[7], and three ancient hemippe genomes, one from Göbeklitepe, Turkey (GT64), and two from museum specimens (Hm_1864 and

Hm_1892)[8](Figure 1C, Appendix D). I then constructed a multidimensional scaling (MDS) summary plot of genomic diversity based on the outgroup f3-statistic (Material Methods), including the modern-day African and Asian asses and the ancient individuals (Figure 3.5, Appendix E). This revealed a unique position for the three Anatolian wild asses, distinct from African asses as well as the Asian hemiones, including ancient onagers from Iran and the hemippes. Accordingly, the three Anatolian wild asses appear phylogenetically distinct from either the hemione or the hemippe clades. The hydruntine is the only other wild ass lineage identified in the fossil record to has lived in Holocene Anatolia. Together with the mtDNA evidence, this strongly suggests the three wild asses belonged to the extinct hydruntine lineage.

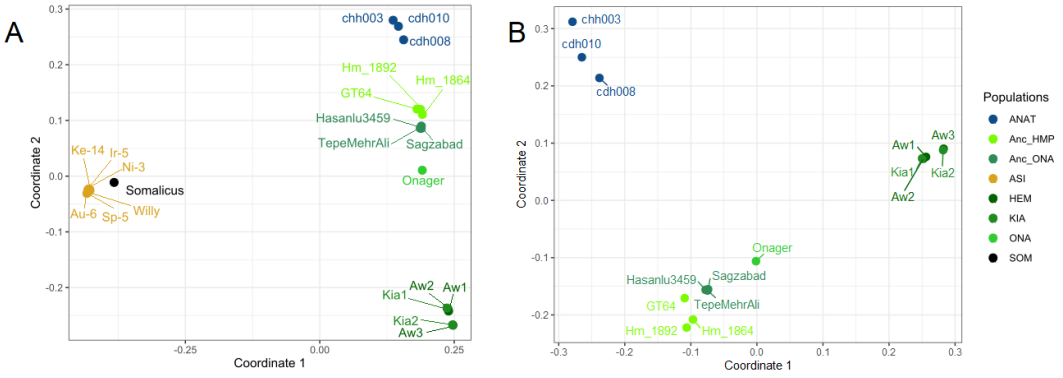


Figure 3.5: Multidimensional scaling plot of A) donkey and wild asses and B) only the Eurasiatic wild asses generated using a distance matrix based on (1-outgroup f3) statistics calculated from the autosomal variation dataset consisting of 2,146,461 sites (ANAT: Anatolian samples reported in this article; 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*).

3.2 Whole Genome Analyses Support Hemiones and Hydruntines as Sister Clades

Because the hydruntine has yet only been genetically described through mtDNA, I 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[5, 11, 9, 10]. Our results above are fully consistent with this hypothesis, with Anatolian and modern-day Asian wild ass individuals clustering to the exclusion of African asses in mtDNA analyses (Figures 3.3 and 3.4). Anatolian wild asses are also closer to hemiones in the MDS- and heatmap-based summaries of nuclear genomic variation using outgroup-f3 scores (Figure 3.5 and 3.6). I further tested this pattern with D-statistics, employing the c.2 million autosomal transversion SNP panel (Material and Methods) and using the horse as an outgroup (Appendix F). D-tests of the form D(Horse, Anatolia; Africa, Hemione) and D(Horse, Hemione; Africa, Anatolia) were all significantly positive (|Z|>3; Methods), in line with the notion that hemiones and hydruntines represent sister taxa (Figure 3.7, Appendix F). I replicated the same results using c.2 million transversions in a secondary variation panel[109] (Material and Methods) (Appendix F).

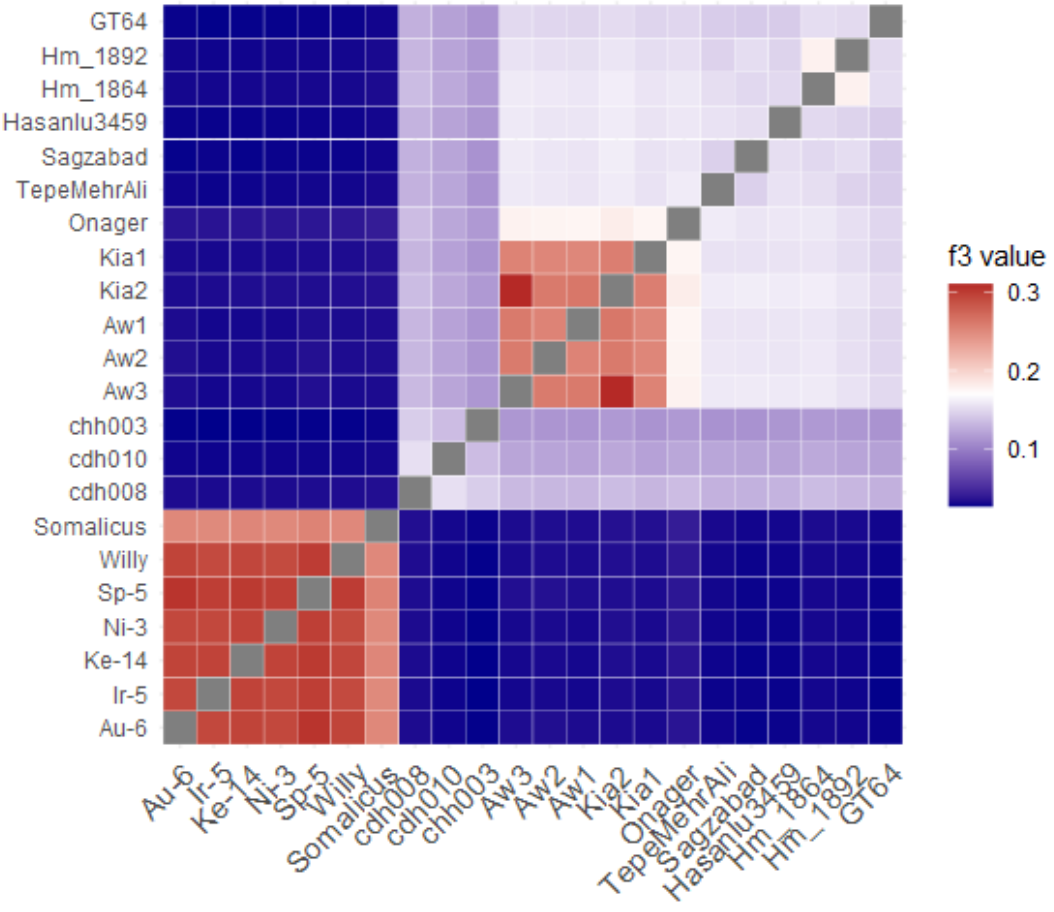


Figure 3.6: Heatmap generated using outgroup-f3 data on autosomal loci total of 2,146,416 positions on individual level.

3.3 Hemiones and Hydruntines Are Reciprocally Monophyletic

Earlier analyses of partial mtDNA sequences had suggested that the hemione-hydruntine divide could be artificial, with some hemiones appearing closer to hydruntines than to other hemiones[5, 10]. Meanwhile, the external positioning of hydruntines to hemiones in my MDS analysis (Figure 3.5) implied the contrary. I tested this question further through a number of approaches. First, I studied mtDNA trees produced using full mtDNA sequences. A Bayesian tree based on full sequences included a hydruntine branch external to hemiones with full support (Figure 3.4).

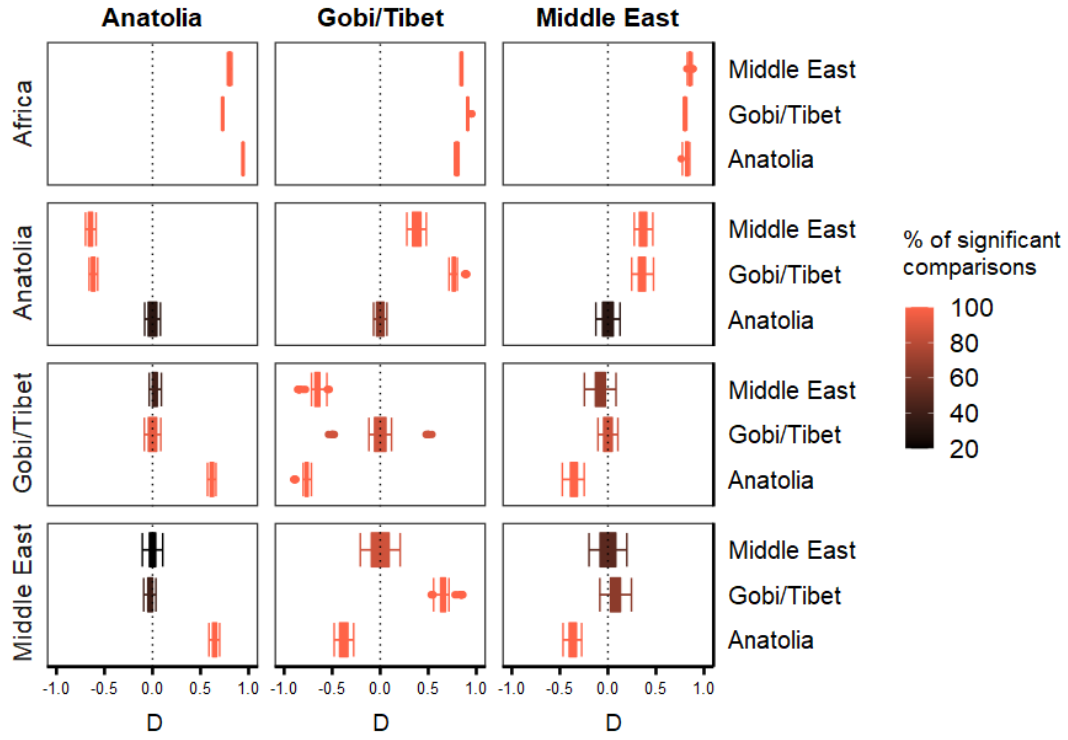


Figure 3.7: D-statistics calculated using autosomal data between regional groups. Test groups are shown on the top. All D-statistic point estimates belongs to a comparison and boxplot represents the interquartile ranges. The color gradient from pink to black represents the fraction of individual comparisons that are nominally significant (|Z|>3). Regional groups include: Anatolia (cdh008, cdh010, chh003), Gobi/Tibet (Aw1, Aw2, Aw3, Kia1, Kia2), Middle East (GT64, Hm_1864, Onager, Hasanlu3459, Sagzabad) and Africa (Asinus, Somalicus).

I further used D-tests to further clarify hemione-hydruntine relationships. I started by testing monophyly of the Anatolian wild ass genomes with D(Horse, AnatoliaX; AnatoliaY, Non-Anatolia), where *AnatoliaX* and *AnatoliaY* refer to different Anatolian wild ass genomes, and *Non-Anatolia* refers to any individual hemione, kiang, onager or hemippe genome (Material and Methods, Appendix F). 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; Methods) (Figure 3.7, Appendix F). I 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 3.7, Appendix F). The evidence thus supports reciprocal monophyly between hemiones/hemippes and hydruntines, represented by Anatolian wild asses.

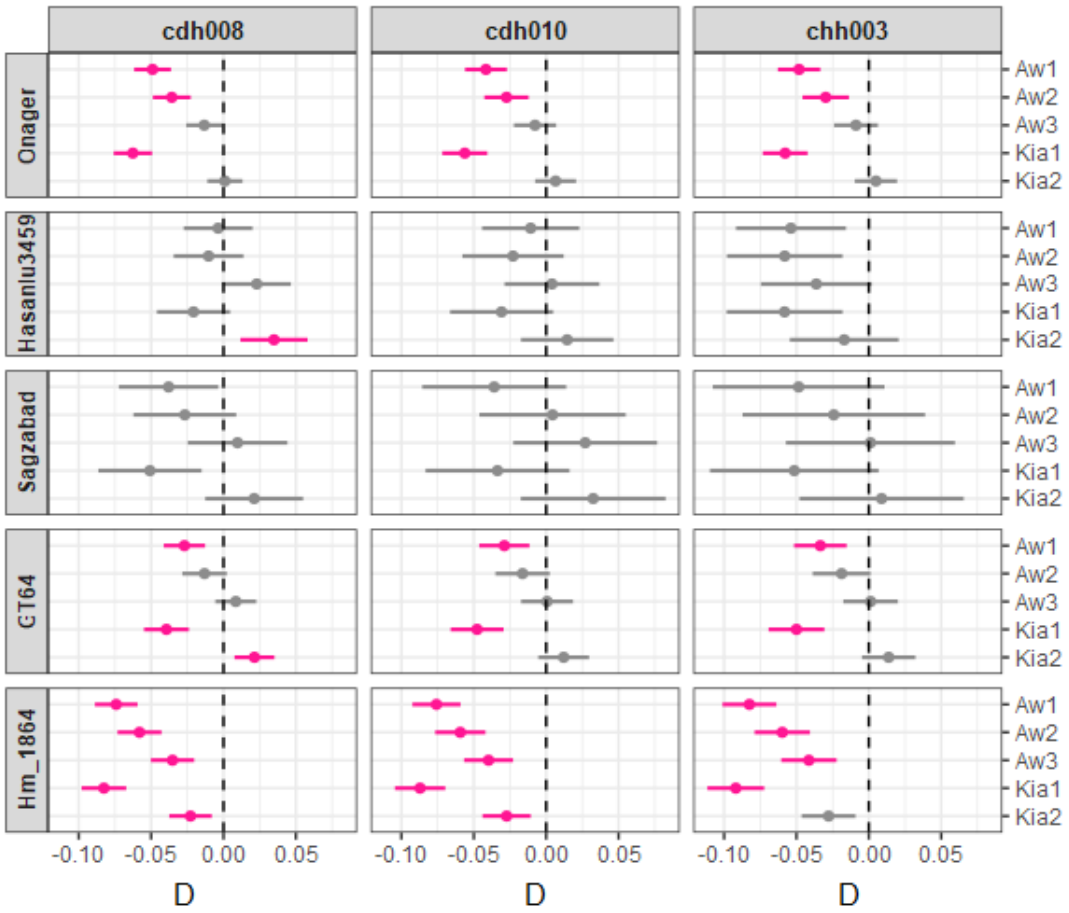


Figure 3.8: D statistics results in the form of D(Horse,Anatolia; MiddleEast, Gobi/Tibet) based on the main autosomal dataset consisting of 2,146,461 variable sites. These are the same results as presented in summary form in Figure 3.7. Pink color represents significant results at |Z|>3 and gray color |Z|<3. Confidence intervals represents 2 SE.

3.4 Evidence for Gene Flow Between Hydruntines and Middle Eastern Wild Asses

I next asked whether the Anatolian wild asses might be symmetrically related to all Asian wild ass lineages. Following our earlier observations (Figures 3.3, 3.4 and 3.5), here I divided Asiatic wild asses into two groups: hemione and kiangs forming a Gobi/Tibet cluster while onager and hemippe forming the Middle East cluster. I then performed D-tests of the form D(Horse, Anatolia; MiddleEast, Gobi/Tibet),

where MiddleEast represents onager and hemippe individuals while Gobi/Tibet consists of hemione and kiangs. I found that all three Anatolian wild asses show higher affinity 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 significant at |Z|>3) (Figures 3.7 and 3.8, Appendix F). This observation is even more pronounced (100% of 75 comparisons at |Z|>3) in comparisons using the secondary variation panel (Figure 3.9, Appendix F). This would be compatible with gene flow between onager/hemippe and hydruntine populations possibly in Southwest Asia. It also indicates the impact of SNP panels and hence ascertainment on D-test results (also see below).

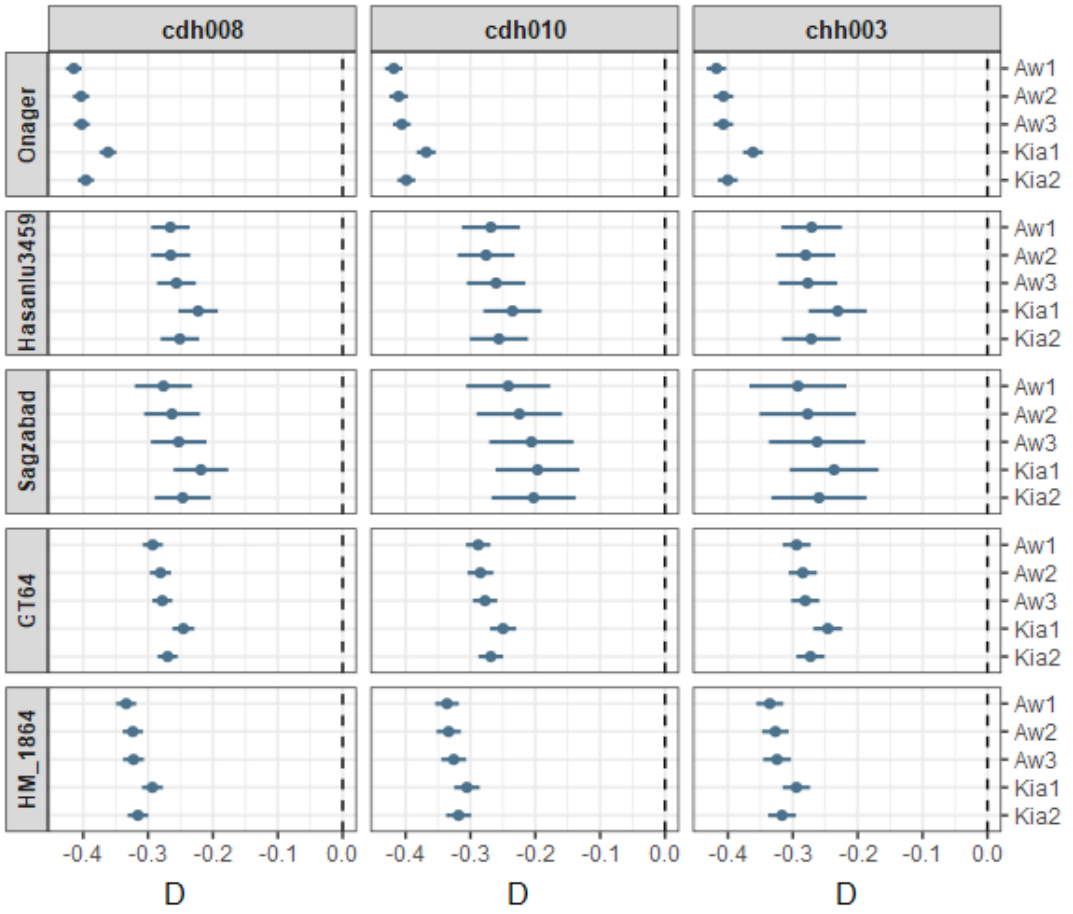


Figure 3.9: D statistics results in the form of D(Horse,Anatolia; MiddleEast, Gobi/Tibet) based on the secondary autosomal variable loci (2,408,064 positions). Blue color represents significant results at |Z|>3 and gray color |Z|<3. Confidence intervals represents 2 SE.

I further investigated this by testing D(Horse, Anatolia; MiddleEastX, Middle-EastY) where I compared genetic affinity of Anatolian hydruntines between the onager and/or hemippe individuals. Using the main variation panel I observed slight affinity toward the 19th century hemippe over the other hemippe and onagers (12/12 comparisons with |Z|>0 and 5 with |Z|>3, Figure 3.10).

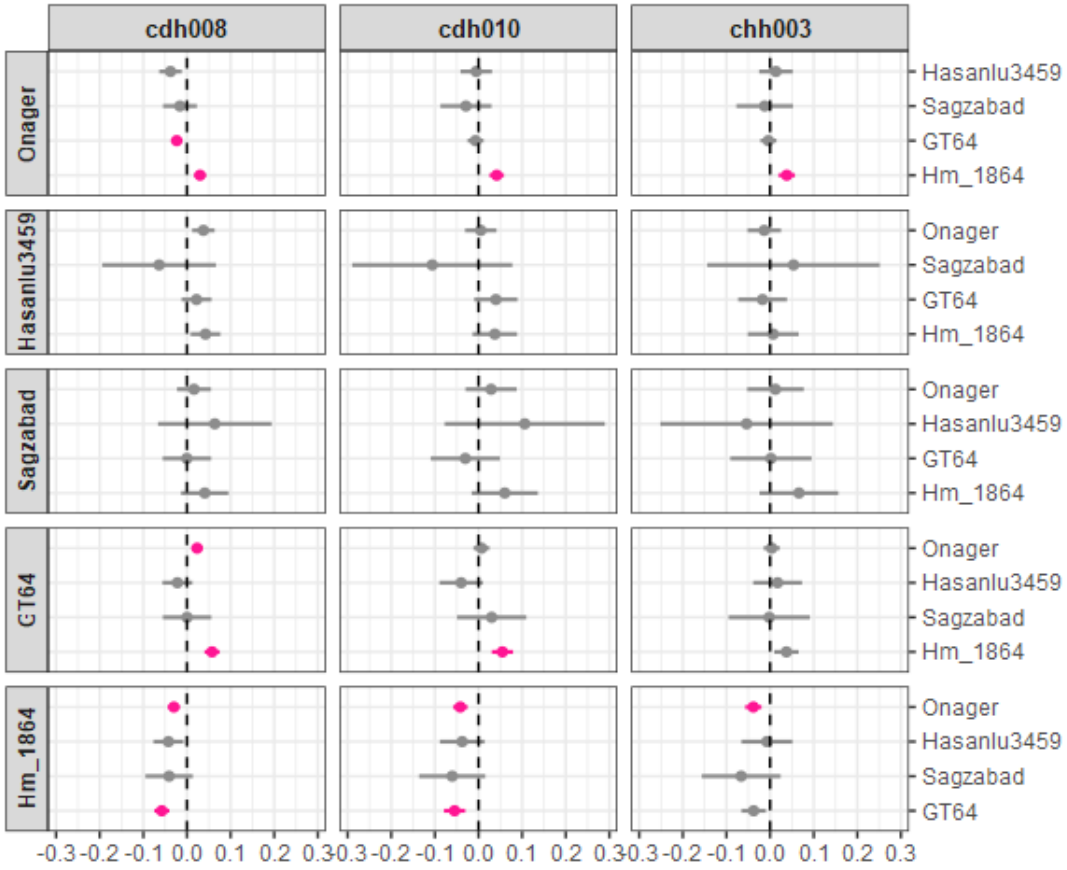


Figure 3.10: D statistics results in the form of D(Horse,Anatolia; MiddleEastX, MiddleEastY) based on the main autosomal dataset consisting of 2,146,461 variable sites. These are the same results as presented in summary form in Figure 3.7. Pink color represents significant results at |Z|>3 and gray color |Z|<3. Confidence intervals represents 2 SE.

In contrast, using the second variation panel I found higher affinity to the modern-day onager over all other Middle East wild ass genomes (12 of 12 comparisons with |Z|>3; Figure 3.11). Thus, although I cannot yet fully determine the exact nature of the possible admixture, I hypothesize that gene flow from hydruntine to the onager

and hemippe lineages could explain the observed patterns (see Discussion).

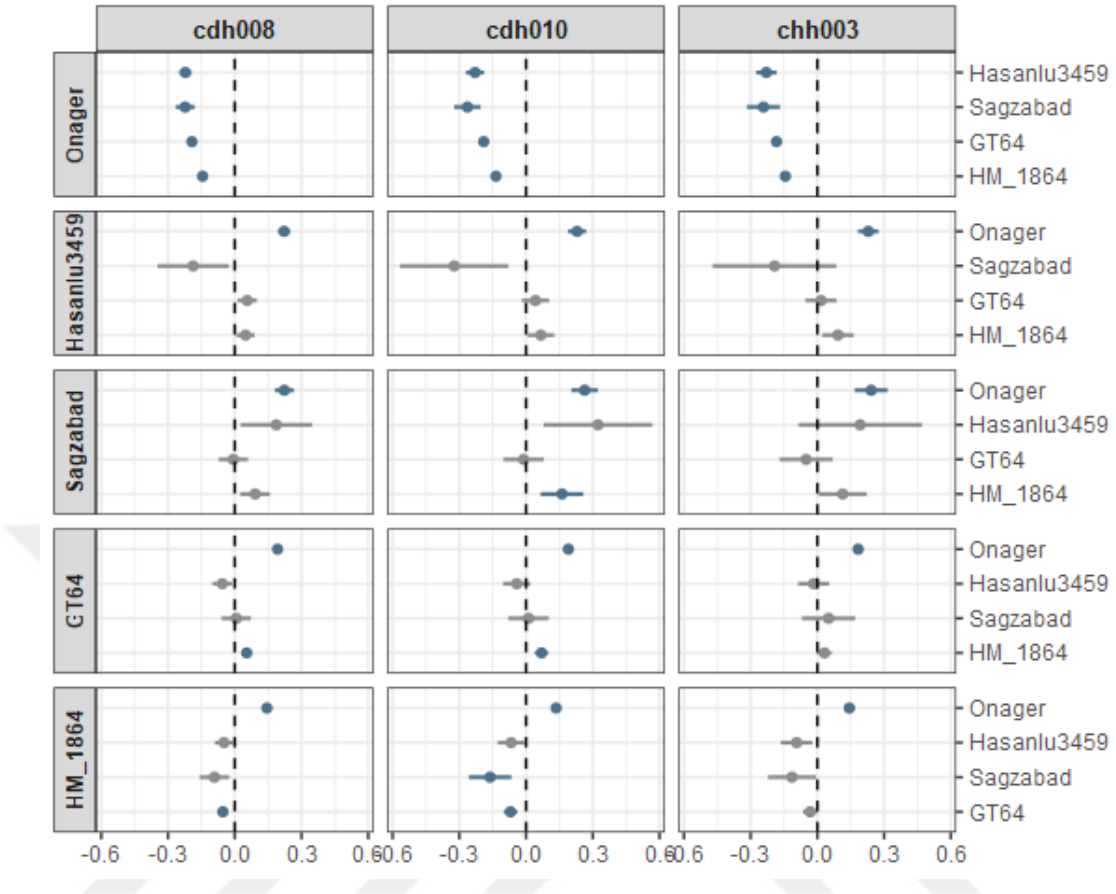


Figure 3.11: D statistics results in the form of D(Horse,Anatolia; MiddleEastX, MiddleEastY) based on the secondary autosomal variating loci(2,408,064 positions). Blue color represents significant results at |Z|>3 and gray color |Z|<3. Confidence intervals represents 2 SE.

3.5 The Timing of the Hydruntine and Asian Wild Ass Split

The above analyses suggest that hydruntines split from other Eurasian wild asses 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 I can estimate the hydruntine-Asian wild ass split time using mtDNA and also autosomal comparisons between hydruntines and Gobi/Tibet wild assess. I first used BEAST for mtDNA split time estimation using the caballine and non-caballine divergence times

of 4.25 mya as an anchor[83] for calibrating splits (Figure 3.4). The mitochondrial divergence estimate between European-Anatolian clade and the Asiatic hemiones was 1.84 [0.38-3.23] mya, while that between the hydruntine specimen from Sicily and Anatolian wild asses was 734 [54-2,056] kya. I note that the split times estimated here tend to be higher than earlier estimates[83, 5, 23] but also have wide confidence intervals that include those earlier estimates.

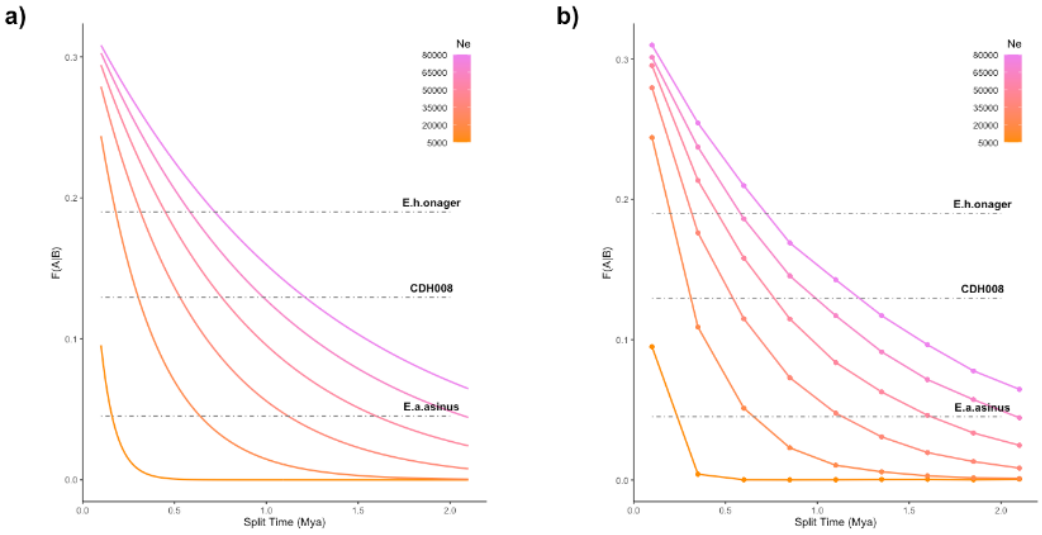


Figure 3.12: Population split time estimation of Anatolian wild ass (cdh008), *E. a. asinus* and *E. h. onager* from Asian wild ass based on the F(A|B) statistic. F(A|B) is the probability of observing a derived allele in the Anatolian wild ass, *E. a. asinus* or *E. h. onager* at the positions where the Asian wild ass genome is heterozygous (Methods). Horizontal lines represents F(A|B) values for cdh008, *E. a. asinus* and *E. h. onager*. Observed and expected F(A|B) values, calculated using a) $e^{-T(2N)3}$ and b) population genetic simulations, are given for divergence times ranging from 100kya to 2 mya, and for population sizes (Ne) ranging from 5k to 80k by the color key.

I then used the F(A|B) method for estimating the split time between hydruntines and kiang from the Gobi/Tibet region using autosomal DNA. This method is designed for split time estimation when a high-quality and a low-quality genome are available for the investigated lineages[131, 133]. I calculated F(A|B), the ratio of derived alleles in the Anatolian wild ass at sites heterozygous in the kiang genome, as 0.145 (and 0.149 using transversions only). Then, I computed the expected F(A|B) values for different

divergence times, using both the theoretical estimate[133] and also the estimate calculated via simulations[134], and using a range of effective population size estimates for the kiang (Material and Methods). The intersection of the observed and expected statistics suggests a most likely population split between 800 kya and 1 mya (Figure 3.12). This is younger than previous mtDNA split time point estimate but within its confidence interval. This autosomal split time would also be compatible with the first fossil records of *E. h. hydruntinus* in Europe c.600 kya[31].

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 yet identified, if not the latest recorded hydruntines to date[73, 4]. Zooarchaeological studies have 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[31, 4]. I thus asked whether I might detect signatures of a stark population decline in our genetic data.

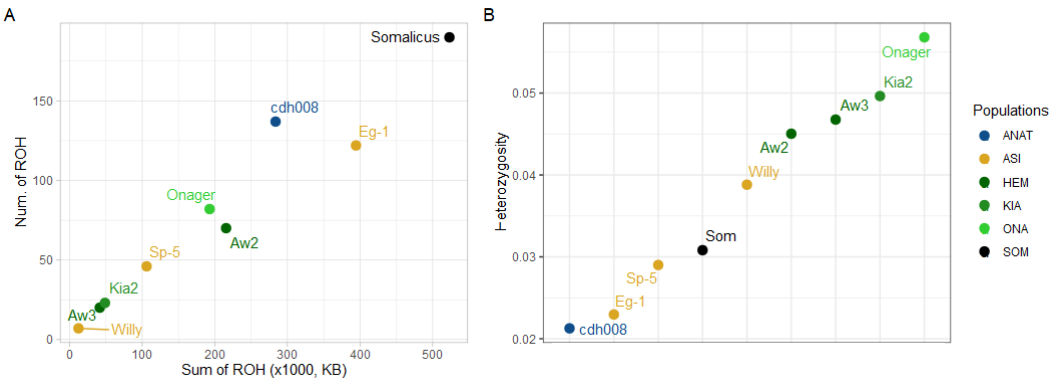


Figure 3.13: A) Total number versus the total length (sum) of runs of homozygosity (ROH) tracks over 1.5 Mb. B) Heterozygosity calculated across 2,146,416 transversion sites. (ANAT: Anatolian samples reported in this article; ASI: *E.asinus*; HEM: *E.hemionus*; KIA: *E.h.kiang*; ONA: *E.h.onager*; SOM: *E.somalicus*).

I first estimated runs of homozygosity (ROH), thus measuring past inbreeding that can be caused by shrinking population size[136]. I called ROH in the highest coverage Anatolian wild ass genome (cdh008) and compared these ROH estimates with those from genomes of other asses, after down-sampling all to the same coverage (Methods). I observed higher numbers of ROH in the African asses (*E. a. asinus* and *E. a. somalicus*) (Figure 3.13A), as expected, since these populations are known to have experienced recent bottlenecks[23, 124]. On the other hand, kiangs, kulans and onagers showed 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.

I investigated this further by estimating heterozygosity from de novo-called transversion variants across the same genomes. Hemiones had the highest heterozygosity, followed by donkeys and the Somali wild ass. The cdh008 genome harbored the lowest heterozygosity, even below the African group known to have undergone several bottlenecks (Figure 3.13B). Combined with ROH analysis, these observations suggest that the Anatolian population was in strong decline by the 1st millennium BCE, which is also supported by its rare occurrence in the archaeological record.

CHAPTER 4

DISCUSSION

In the course of this thesis work, I screened 15 ancient equid individuals from the Central Anatolian settlements of Çadır Höyük and Çatalhöyük using well-established archaeogenomic workflows. I produced one relatively high (cdh008) and two lesser (cdh010 and chh003) quality whole genomes.

4.1 Ancient DNA Conservation and Sequencing Characteristics

Although the three samples subjected to further analyses, cdh008, cdh010, and chh003, possessed 11.35%, 9.15%, and 9.12% endogenous equid proportions, respectively, the other 12 samples yielded little or no ancient equid DNA (fourth highest was chh009 with 0.48% equid proportion). The hot climatic conditions of Central Anatolia have a negative effect on ancient DNA preservation. Previous research conducted by our research group produced double-stranded whole genome libraries of 29 ancient sheep samples, only to find that five samples yielded over 1% endogenous ancient sheep content[138]. Moreover, the research conducted on Neolithic Central Anatolian human populations[99] reported similar results; only one-third of the samples contained over 1% endogenous human DNA. Considering these previous reports, recovering endogenous DNA content over 9% from 20% of screened samples falls within the expected thresholds. The poor preservation of samples from Çatalhöyük is evident from discovering only one usable individual (chh003) excavated from a mixed archaeological context of a midden area used as a trash disposal site. This individual was dated to the Anatolian Iron Age, which was somewhat surprising given that Çatalhöyük is a well-known Neolithic site in Central Anatolia, and most findings

date to the Neolithic period, while the Iron Age deposits are very few.

Sequencing characteristics such as average read length, PMD patterns on both ends of reads, and clonality values after deeper sequencing reflected that I successfully carried out ancient DNA isolation and library preparation processes. This was further corroborated by downstream analyses, where I encountered no effect either stemming from wet lab procedures or manifested as a result of cross-contamination. The blanks were also clean.

The libraries had a limitation where PMD signals extended to inner parts of reads, going beyond the trimming threshold of the tenth base pair. Due to this issue, I could not utilize trimmed files without filtering transitions. Despite such obstacles, the results indicated the successful completion of the first stage of this thesis project, which involved the extraction and sequencing of ancient DNA from Iron Age Central Anatolian equid samples.

4.2 Sex Distribution of the Samples

Fifteen samples were classified as wild asses without considering their hemione or hydruntine ancestry. The majority of these samples were females (10 of 12 individuals whose genetic sex could be assigned, p=0.03). This result is surprising since studies on the remains of wild animals reported male-biased fossil records of wild species such as woolly mammoths, bison, and brown bears[139, 140], or equal distribution of males and female horses before domestication[141] and the reindeer hunted in Last Glacial period in Northern Germany[142]. On the other hand, studies on contemporary hunting of greater sage-grouse in North America have found that females are more frequently targeted than males during their adult years and as juveniles[143]. The exceeding number of female wild asses in Central Anatolian assemblages of Çadır Höyük and Çatalhöyük can be explained by 1) the ease of capture of female individuals due to the attachment of their offspring or 2) harem-like social structure of wild equid herds which might include a couple of dozen females for a handful of males increase the chance of capturing or hunting a female individual than a male. Hunting of females might drive the already declining population to extinction. Re-

gardless, my findings suggest a bias towards capturing or hunting female hydruntines over males in Central Anatolia.

4.3 Biases in population genomic analyses

In this thesis, I present a novel variation panel that aims to minimize ascertainment bias. The panel's efficacy is verified by validating the results against an earlier published dataset[109].

My initial efforts of generating such a dataset yielded failed results. In these phases, I developed several datasets that contained transition variants as well as transversions, although the former that could be confused with PMD signals. These datasets also under-represented wild asses over domesticated donkey variation. Although these datasets yielded expected results (and results consistent with the main SNP panel I later produced) when I generated heatmaps and MDS plots based on f3 statistics, D-statistics calculated on ABBA-BABA values suggested unlikely gene flow patterns. These imbalances in datasets appear to have produced false genetic affinities between domesticated donkeys which originated in Africa and Eurasiatic wild ass species. Possible gene flow or incomplete lineage sorting events in deep population history that may not have been detected so far may also have played a role in this issue.

It is important to note that demographic analyses are primarily based on the datasets (main dataset and secondary dataset) that do not include variations in European wild asses. This is due to the fact that there is only one individual with acceptable sequencing coverage, the one I produced in this thesis, but even that was too low to call diploid variants with high confidence. Moreover, although several subspecies or variations have been reported in zooarchaeological discourse, the literature lacks knowledge about the genetic structure of *E. h. hydruntinus*, if there is any. There is a dire need for more high-quality hydruntine and other wild ass genomes to construct a dataset that reflects the whole Eurasiatic wild ass gene pool, which can be incorporated into studies to resolve possible intricate gene flow networks.

4.4 European Wild Asses of Anatolia

The question of the origin of Anatolian wild asses was the first issue to be addressed using the data I generated. Although the equid-specific taxon identifying algorithm, Zonkey[109], assigned these three samples as *wild ass*, this identification is not only based on a reference panel which consists of a limited number of representatives of wild asses, namely a single Iran onager and a single Central Asian kiang, but also a variation dataset which lacked the greater variation that non-African stenonids harbors. Despite these limitations, the Zonkey algorithm produced a striking result that these three individuals were not domesticated donkeys, a species most likely to be encountered in Iron Age Anatolia at a time when people were already familiar with equid management. The wild status of these Central Anatolian asses raises the possibility that the equid relationship in the past combined management of domesticates within settlement perimeters (i.e., the existence of domesticated horses contemporary with wild asses in Çadır Höyük, unpublished data) and constant input of wild animals from outside. This means that even though domesticates were present and readily available for human exploitation, wild animals were also sought after in activities such as hunting. It is worth noting that zooarchaeologists found no cut marks on the samples of this thesis. It is important to consider that the majority of the material analyzed consisted of individual teeth and bones, which could have influenced the lack of findings.

Since historical occurences of both *E. hemionus* (including hemippe) and *E. hydruntinus* is found within the boundaries of present-day Turkey, wild asses from the Iron Age context would most likely belong to one of these species. In order to establish the taxonomic identity of Anatolian wild asses, my colleagues and I used several strategies that employed mitochondrial and nuclear sequences. The first approach leveraged published modern and ancient sequences of the hypervariable region on mitochondrial DNA due to the high number of already published sequences[5, 9, 11, 23]. This analysis showed that the Anatolian ancient samples were closely related to other hydruntines from Anatolia[5], Crimea[9], Iran[9], Sicily[11] and Europe[5] with full support. The closest relatives of these three Anatolian samples were Central Anatolian hydruntines. Lack of monophyly reported before[5, 10] within the hydruntine

gene pool was also confirmed using mitochondrial DNA data; however, this observation was based on a 229 bp long fragment on mitochondrial DNA. Using fewer individuals with complete mitogenome sequences also confirmed the phylogenic status of the samples, clustering Anatolian samples with the single near-complete mitochondrial sequence of hydruntines[11] published so far, excluding all other asses with Asiatic ancestry with full support.

These three Anatolian asses were the first nuclear genome sequences of *E. h. hydruntinus*, as a hydruntine whole genome that could be coined as a reference sample, was absent. D-statistic results in the form of D(Horse, AnatoliaX; AnatoliaY, Non-Anatolian-Ass), where AnatoliaX and AnatoliaY represent different samples of this thesis, and Non-Anatolian-Ass include the other individuals within the dataset; all showed greater affinity among Anatolian samples than others. The divergence time estimates calculated on mtDNA and nuclear sequences also support the idea of hydruntines representing a distinct lineage. These results present evidence that the gene pool of Anatolian wild asses was distinct and most closely related to *E. h. hydruntinus*; hence, Anatolian wild asses sampled in this thesis were most likely hydruntines.

I also assigned the mitochondrial DNA haplogroup of three Anatolian wild asses as *H1*, which was the most prevalent haplogroup in hydruntines[5]. On the other hand, the *H2* haplogroup observed in two individuals from vastly different locations previously (France and Iran)[5, 9] was lacking in Anatolia, presenting an intriguing situation about the hydruntine gene pool and history. Although this thesis adds merely three individuals to existing mitochondrial DNA knowledge based on the previously reported 19 individuals of hydruntines of Anatolia, Anatolia's lack of an *H2* haplogroup suggests a more profound divergence between these two haplogroups.

4.5 Population History of European Wild Asses

4.5.1 Divergence between African and Eurasiatic wild ass branches

The results of this thesis also confirm the divergence between African and Eurasian wild ass populations. BEAST analysis on partial and near complete mitochondrial

sequences clustered hydruntines from Anatolia, onagers from Iran, hemippe from Syria, and hemiones from Central Asia together, while donkeys and Somali wild ass formed another and distinct cluster, basal to Eurasiatic branch. The divergence between African and Eurasian wild asses was estimated to be 3.57 Mya and 1.2-2 Mya from mitochondrial and nuclear sequences, respectively. D-statistics results showed that hydruntines, onagers, hemippe, and hemiones share more genetic drift than the African wild ass and donkey. This provides evidence from whole genome data for the Eurasia vs. Africa divergence that was previously estimated using mitochondrial sequences[83, 5].

4.5.2 Degrees of variation within the Eurasiatic wild ass gene pool

The significance of this thesis lies in its ability to offer an improved genetic resolution of wild ass phylogenetic and demographic relationships, as it is based on nuclear sequences with over 2 million data points, compared to a subset of mitochondrial sequences that are only about 200 bp long. I took advantage of this dataset's resolution to solve intricate relationship patterns within the Eurasiatic wild ass gene pool. As mentioned earlier, a multidimensional scaling analysis (MDS) based on a distance matrix generated from outgroup f3 statistics yielded clearly separated clusters. The first coordinate split African and Eurasian clusters. The second coordinate divided Eurasian wild asses into three main clusters: modern hemione and kiangs were the first group, while all modern and ancient onager and hemippe formed another cluster, and a third cluster which consisted of hydruntines reported in this thesis work. This last cluster was differentiated from hemiones, onagers, and hemippe, and this divergence was also presented in phylogenetic trees based on whole mitochondrial sequences. D-statistics also supported this separation, providing affinities between hemione, kiang, onager, and hemippe groups, excluding hydruntines. Combining these genetic results with the phenotypical differences reported since the first discovery of hydruntine remains, I concluded that hydruntines were not a subbranch of the greater hemione gene pool; instead, the individuals sampled in this thesis represent a distinct gene pool of themselves. Hydruntines may deserve a separate taxonomic status from other Eurasiatic wild asses; however, this conclusion should be considered cautiously since I present three individuals, and only one has relatively high coverage.

It is also important to be cautious when interpreting the results of this study, as the individuals sampled come from a population that is in decline and probably represents the periphery of the original hydruntine distribution (see section 4.6 Extinction Dynamics of European Wild Ass). This may affect their genetic makeup and make their affinities with other asses less reflective about the past hydruntine gene pool across West Eurasia.

4.5.3 Gene flow between Eurasiatic wild ass species

My findings also presented evidence of gene flow between hydruntines sampled from Anatolia and the modern onager and historical hemippe. This is particularly intriguing because a previous study[5] has shown that hydruntines are closer to Central Asian hemiones, namely Turkmen kulan and Indian khur, than Middle Eastern onager and hemippe populations in analyses based on mitochondrial control region sequence, albeit with limited statistical support. Evidence for a signal of a gene flow between modern onagers and hydruntines and a lack of signal between ancient onagers was also interesting. If true, this suggests that a possible gene flow event that occured after the 2nd Millenium BC. The most closely related individual to our sampling was a historical hemippe from Syria (evidenced in 16 of 27 relevant D-statistic comparisons). These results point out to two conclusions: 1) mitochondrial sequences are unable to capture past gene flow events even in the hyper-variable region, and 2) sympatry of hydruntines, hemiones, and hemippes allowed gene flow among lineages and thus produced a complex population history for these species.

4.6 Extinction Dynamics of European Wild Ass

The last significant output of this thesis was a genetic snapshot of a population on the brink of extinction. The apparent decline in the population is mainly evidenced by runs of homozygosity (ROH) analyses where the high-quality hydruntine from Anatolia (cdh008) shows a clear signal of a high ROH load, comparable with donkeys, which have undergone a severe domestication bottleneck[79] and the Somali wild ass, which has been reported as highly inbred[144]. Heterozygosity analyses in which

the Anatolian hydruntine showed the lowest level of heterozygosity, even lower than the donkey and Somali wild ass, further supported this result. Based on the high levels of ROH and low levels of heterozygosity, it can be deduced that the extinction of hydruntines is likely a result of habitat fragmentation, leading to a reduction in discrete populations[4]. This event should have been accelerated by anthropogenic factors such as hunting, which is suggested by the fact that these individuals were sampled chiefly from areas in ancient settlements where people discarded their food-related residues, such as the midden context of chh003 in Çatalhöyük.

4.7 Conclusions

This thesis work presented one full and two partial nuclear genomes of extinct *E. hydruntinus* from Iron Age contexts of Central Anatolian settlements of Çadır Höyük and Çatalhöyük, for the first time in the literature. Their radiocarbon dating placed these individuals as the last confirmed sightings of this species. The results showed that enigmatic species of *E. hydruntinus*, which had puzzled scientists for its phenotypical features, had a Eurasian origin, as suggested before; however, it had its distinct gene pool other than hemiones, onagers, hemippes, and kiangs. Although their separate species' status has been discussed and rejected in previous studies based on partial mitochondrial sequences, nuclear genomes have revealed a different story.

Since *E. hydruntinus* was a wild species, their interaction with humans was most likely limited to hunter-prey dynamics, although unique interactions involving breeding domestic and wild asses have also been reported[8]. However, human contribution to this extinction event is worth noting since the same dynamic is also a current event considering global climate change and the accelerated rate of extinction in the present time. A genomic snapshot of a species immediately before the extinction event can be a starting point for conservation efforts and present a basis for strategies for a sustainable population for threatened species.

Finally, although my results indicate that hydruntines were an early split from the Asiatic wild ass lineage, I remain cautious about suggesting a different species status since this thesis has limitations, such as the low number of individuals and only a tiny

fragment of hydruntine dispersal sampled. Moreover, the observation of likely gene flow between hydruntines and hemiones indicates that these would not be considered biological species sensu Mayr. However, the lack of interspecific hybridization is not a strong criterion of species definition[145], thus, I would not like to eliminate the distinct species status of European wild asses. Future work should be concentrated on broadening the geographical cover of hydruntine distribution while adding more high-quality ancient genomes to the literature. The past genomic diversity of *E. hydruntinus* should be the focal point of future work to unravel population structure and extinction dynamics.

REFERENCES

[1] B. J. MacFadden, "Patterns of phylogeny and rates of evolution in fossil horses; hipparions from the Miocene and Pliocene of North America," *Paleobiology*, vol. 11, pp. 245–257, July 1985.

[2] E. N. van Asperen, "Late Middle Pleistocene horse fossils from northwestern Europe as biostratigraphic indicators," *Journal of Archaeological Science*, vol. 39, pp. 1974–1983, July 2012.

[3] V. Eisenmann, "Pliocene and Pleistocene equids: Palaeontology versus molecular biology," *CFS Courier Forschungsinstitut Senckenberg*, pp. 71–89, Nov. 2006.

[4] J. J. Crees and S. T. Turvey, "Holocene extinction dynamics of Equus hydruntinus, a late-surviving European megafaunal mammal," *Quaternary Science Reviews*, vol. 91, pp. 16–29, May 2014.

[5] E. A. Bennett, S. Champlot, J. Peters, B. S. Arbuckle, S. Guimaraes, M. Pruvost, S. Bar-David, S. J. M. Davis, M. Gautier, P. Kaczensky, R. Kuehn, M. Mashkour, A. Morales-Muñiz, E. Pucher, J.-F. Tournepiche, H.-P. Uerpmann, A. Bălăşescu, M. Germonpré, C. Y. Gündem, M.-R. Hemami, P.-E. Moullé, A. Ötzan, M. Uerpmann, C. Walzer, T. Grange, and E.-M. Geigl, "Taming the late Quaternary phylogeography of the Eurasiatic wild ass through ancient and modern DNA," *PLOS ONE*, vol. 12, p. e0174216, Apr. 2017. Publisher: Public Library of Science.

[6] B. M. Peter, "A geometric relationship of F2, F3 and F4-statistics with principal component analysis," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 377, p. 20200413, Apr. 2022. Publisher: Royal Society.

[7] A. Fages, K. Hanghøj, N. Khan, C. Gaunitz, A. Seguin-Orlando, M. Leonardi, C. McCrory Constantz, C. Gamba, K. A. S. Al-Rasheid, S. Albizuri, A. H.

Alfarhan, M. Allentoft, S. Alquraishi, D. Anthony, N. Baimukhanov, J. H. Barrett, J. Bayarsaikhan, N. Benecke, E. Bernáldez-Sánchez, L. Berrocal-Rangel, F. Biglari, S. Boessenkool, B. Boldgiv, G. Brem, D. Brown, J. Burger, E. Crubézy, L. Daugnora, H. Davoudi, P. de Barros Damgaard, M. de Los Ángeles de Chorro Y de Villa-Ceballos, S. Deschler-Erb, C. Detry, N. Dill, M. do Mar Oom, A. Dohr, S. Ellingvåg, D. Erdenebaatar, H. Fathi, S. Felkel, C. Fernández-Rodríguez, E. García-Viñas, M. Germonpré, J. D. Granado, J. H. Hallsson, H. Hemmer, M. Hofreiter, A. Kasparov, M. Khasanov, R. Khazaeli, P. Kosintsev, K. Kristiansen, T. Kubatbek, L. Kuderna, P. Kuznetsov, H. Laleh, J. A. Leonard, J. Lhuillier, C. Liesau von Lettow-Vorbeck, A. Logvin, L. Lõugas, A. Ludwig, C. Luis, A. M. Arruda, T. Marques-Bonet, R. Matoso Silva, V. Merz, E. Mijiddorj, B. K. Miller, O. Monchalov, F. A. Mohaseb, A. Morales, A. Nieto-Espinet, H. Nistelberger, V. Onar, A. H. Pálsdóttir, V. Pitulko, K. Pitskhelauri, M. Pruvost, P. Rajic Sikanjic, A. Rapan Papeša, N. Roslyakova, A. Sardari, E. Sauer, R. Schafberg, A. Scheu, J. Schibler, A. Schlumbaum, N. Serrand, A. Serres-Armero, B. Shapiro, S. Sheikhi Seno, I. Shevnina, S. Shidrang, J. Southon, B. Star, N. Sykes, K. Taheri, W. Taylor, W.-R. Teegen, T. Trbojević Vukičević, S. Trixl, D. Tumen, S. Undrakhbold, E. Usmanova, A. Vahdati, S. Valenzuela-Lamas, C. Viegas, B. Wallner, J. Weinstock, V. Zaibert, B. Clavel, S. Lepetz, M. Mashkour, A. Helgason, K. Stefánsson, E. Barrey, E. Willerslev, A. K. Outram, P. Librado, and L. Orlando, "Tracking Five Millennia of Horse Management with Extensive Ancient Genome Time Series," *Cell*, vol. 177, pp. 1419–1435.e31, May 2019.

[8] E. A. Bennett, J. Weber, W. Bendhafer, S. Champlot, J. Peters, G. M. Schwartz, T. Grange, and E.-M. Geigl, "The genetic identity of the earliest human-made hybrid animals, the kungas of Syro-Mesopotamia," *Science Advances*, Jan. 2022. Publisher: American Association for the Advancement of Science.

[9] L. Orlando, M. Mashkour, A. Burke, C. J. Douady, V. Eisenmann, and C. Hänni, "Geographic distribution of an extinct equid (Equus hydruntinus: Mammalia, Equidae) revealed by morphological and genetical analyses of fossils," *Molecular Ecology*, vol. 15, no. 8, pp. 2083–2093, 2006. Number: 8.

[10] L. Orlando, J. L. Metcalf, M. T. Alberdi, M. Telles-Antunes, D. Bonjean,

M. Otte, F. Martin, V. Eisenmann, M. Mashkour, F. Morello, J. L. Prado, R. Salas-Gismondi, B. J. Shockey, P. J. Wrinn, S. K. Vasil'ev, N. D. Ovodov, M. I. Cherry, B. Hopwood, D. Male, J. J. Austin, C. Hänni, and A. Cooper, "Revising the recent evolutionary history of equids using ancient DNA," *Proceedings of the National Academy of Sciences*, vol. 106, pp. 21754–21759, Dec. 2009. Publisher: National Academy of Sciences Section: Biological Sciences.

[11] G. Catalano, A. Modi, G. Mangano, L. Sineo, M. Lari, and L. Bonfiglio, "A mitogenome sequence of an Equus hydruntinus specimen from Late Quaternary site of San Teodoro Cave (Sicily, Italy)," *Quaternary Science Reviews*, vol. 236, p. 106280, May 2020.

[12] P. D. Moehlman, *Equids : zebras, asses and horses : status survey and conservation action plan*. IUCN, 2002.

[13] B. J. MacFadden, *Fossil Horses: Systematics, Paleobiology, and Evolution of the Family Equidae*. Cambridge University Press, 1st paperback edition ed., June 1994.

[14] D. J. Froehlich, "Quo vadis eohippus? The systematics and taxonomy of the early Eocene equids (Perissodactyla)," *Zoological Journal of the Linnean Society*, vol. 134, no. 2, pp. 141–256, 2002. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1046/j.1096-3642.2002.00005.x.

[15] B. J. MacFadden, "Fossil Horses–Evidence for Evolution," *Science*, vol. 307, pp. 1728–1730, Mar. 2005. Publisher: American Association for the Advancement of Science.

[16] B. J. MacFadden and R. C. Hulbert, "Explosive speciation at the base of the adaptive radiation of Miocene grazing horses," *Nature*, vol. 336, pp. 466–468, Dec. 1988. Number: 6198 Publisher: Nature Publishing Group.

[17] E. H. Lindsay, N. D. Opdyke, and N. M. Johnson, "Pliocene dispersal of the horse Equus and late Cenozoic mammalian dispersal events," *Nature*, vol. 287, pp. 135–138, Sept. 1980. Number: 5778 Publisher: Nature Publishing Group.

[18] L. Rook, R. L. Bernor, L. S. Avilla, O. Cirilli, L. Flynn, A. Jukar, W. Sanders, E. Scott, and X. Wang, "Mammal Biochronology (Land Mammal Ages) Around the World From Late Miocene to Middle Pleistocene and Major Events in Horse Evolutionary History," *Frontiers in Ecology and Evolution*, vol. 7, 2019.

[19] A. Azzaroli, "Quaternary mammals and the "end-Villafranchian" dispersal event — A turning point in the history of Eurasia," *Palaeogeography, Palaeoclimatology, Palaeoecology*, vol. 44, pp. 117–139, Nov. 1983.

[20] A. Forsten, "Middle Pleistocene replacement of stenonid horses by caballoid horses — ecological implications," *Palaeogeography, Palaeoclimatology, Palaeoecology*, vol. 65, pp. 23–33, May 1988.

[21] D. R. Prothero, R. M. Schoch, and A. P. o. S. a. M. R. M. Schoch, *Horns, Tusks, and Flippers: The Evolution of Hoofed Mammals*. JHU Press, 2002.

[22] Y. Sam, "African origins of modern asses as seen from paleontology and DNA: what about the Atlas wild ass?," *Geobios*, Jan. 2020.

[23] H. Jónsson, M. Schubert, A. Seguin-Orlando, A. Ginolhac, L. Petersen, M. Fumagalli, A. Albrechtsen, B. Petersen, T. S. Korneliussen, J. T. Vilstrup, T. Lear, J. L. Myka, J. Lundquist, D. C. Miller, A. H. Alfarhan, S. A. Alquraishi, K. A. S. Al-Rasheid, J. Stagegaard, G. Strauss, M. F. Bertelsen, T. Sicheritz-Ponten, D. F. Antczak, E. Bailey, R. Nielsen, E. Willerslev, and L. Orlando, "Speciation with gene flow in equids despite extensive chromosomal plasticity," *Proceedings of the National Academy of Sciences*, vol. 111, pp. 18655–18660, Dec. 2014. Publisher: National Academy of Sciences Section: Biological Sciences.

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

[25] H. G. Stehlin and P. Graziosi, *Ricerche sugli Asinidi fosili d'Europa*. Birkhäuser, 1935. Google-Books-ID: bDzXZwEACAAJ.

[26] A. Azzaroli, “Ascent and decline of monodactyl equids: a case for prehistoric overkill,” *Annales Zoologici Fennici*, vol. 28, no. 3/4, pp. 151–163, 1991. Publisher: Finnish Zoological and Botanical Publishing Board.

[27] M. T. Alberdi, E. Ortiz-Jaureguizar, and J. L. Prado, “A quantitative review of European stenonoid horses,” *Journal of Paleontology*, vol. 72, pp. 371–387, Mar. 1998. Publisher: Cambridge University Press.

[28] M. T. Alberdi and M. R. Palombo, “The late Early to early Middle Pleistocene stenonoid horses from Italy,” *Quaternary International*, vol. 288, pp. 25–44, Mar. 2013.

[29] V. Eisenmann, “Equus granatensis of Venta Micena and evidence for primitive non-stenonid horses in the Lower Pleistocene,” in *Los homínidos y su entorno en el Pleistoceno inferior y medio de Eurasia: actas del Congreso Internacional de Paleontología Humana, Orce 1995, 1999, ISBN 84-8416-938-3, págs. 175-189*, pp. 175–189, Museo de Prehistoria y Paleontología "J. Gibert", 1999. Section: Los homínidos y su entorno en el Pleistoceno inferior y medio de Eurasia: actas del Congreso Internacional de Paleontología Humana, Orce 1995.

[30] M. R. Palombo and M. T. Alberdi, “Light and Shadows in the Evolution of South European Stenonid Horses,” *Fossil Imprint*, vol. 73, pp. 115–140, Aug. 2017.

[31] N. Boulbes and E. N. van Asperen, “Biostratigraphy and Palaeoecology of European Equus,” *Frontiers in Ecology and Evolution*, vol. 7, p. 301, 2019.

[32] J. Aurell Garrido, J. Madurell Malapeira, D. Alba, and S. Moyà-Solà, “Los équidos estenonianos y caballinos de la sección pleistocena de Vallparadis (Terrassa, Cataluña, España),” *Cidaris*, vol. 30, pp. 61–66, Jan. 2010.

[33] P. Saladié, A. Rodríguez-Hidalgo, J. Marín, J. Vallverdú i Poch, and E. Carbonell, “The top of the Gran Dolina (Atapuerca, Spain) sequence: A zooarchaeological and occupational perspective,” *Quaternary Science Reviews*, vol. 195, pp. 48–71, Sept. 2018.

[34] F. Marra, S. Nomade, A. Pereira, C. Petronio, L. Salari, G. Sottili, J. J. Bahain, G. Boschian, G. Di Stefano, C. Falguères, F. Florindo, M. Gaeta, B. Giaccio, and M. Masotta, "A review of the geologic sections and the faunal assemblages of Aurelian Mammal Age of Latium (Italy) in the light of a new chronostratigraphic framework," *Quaternary Science Reviews*, vol. 181, pp. 173–199, Feb. 2018.

[35] M. F. Bonifay, ""Equus hydruntinus Regalia minor n.ssp. from the caves of Lunel-Viel (Hérault, France),"," in *Equids in the Ancient World, Vol. II, Beihefte Zum Tübinger Atlas des Vorderen Orients, Reihe A* (R. H. Meadow and H. P. Uerpmann, eds.), pp. 179–216, Wiesbaden: Naturwissenschaften, 1991.

[36] A. Forsten, "The Pleistocene horses from Achenheim, near Strasbourg (Alsace, France), and a comparison with contemporaneous finds," *Neues Jahrbuch für Geologie und Paläontologie - Abhandlungen*, pp. 163–177, Sept. 1996. Publisher: Schweizerbart'sche Verlagsbuchhandlung.

[37] T. Van Kolfschoten, E. Buhrs, and I. Verheijen, "The larger mammal fauna from the Lower Paleolithic Schöningen Spear site and its contribution to hominin subsistence," *Journal of Human Evolution*, vol. 89, pp. 138–153, Dec. 2015.

[38] P. Pettitt and M. White, *The British Palaeolithic: Human Societies at the Edge of the Pleistocene World*. London ; New York: Routledge, 1st edition ed., Jan. 2012.

[39] G. Baryshnikov, "Local biochronology of Middle and Late Pleistocene mammals from the Caucasus," *Russian Journal of Theriology*, vol. 1, pp. 61–67, June 2002.

[40] J. Clutton-Brock, "A re-consideration of the fossil fauna from C-Spring, Azraq,," in *The Hammer on the Rock: Studies on the Early Paleolithic of Azraq* (L. Copeland and F. Hours, eds.), BAR International Series, pp. 391–397, Oxford, 1989.

[41] N. Reynaud Savioz and P. †, "La faune de Nadaouiyeh Aïn Askar (Syrie cen-

trale, Pléistocène moyen) : aperçu et perspectives," *Revue de Paleobiologie*, vol. volume spécial 10, pp. 31–35, Jan. 2005.

[42] P. Samson, "Les équidés fossiles de Roumanie (Pliocène moyen – Pléistocène supérieur)," *Geologica Romana*, vol. 14, pp. 165–352, 1975.

[43] V. Eisenmann, J. Howe, and M. Pichardo, "Old world hemiones and new world slender species (Mammalia, Equidae)," *Palaeovertebrata*, vol. 36, no. 1-4, pp. 159–233, 2008.

[44] H. Alimen, "Remarques sur Equus hydruntinus Regalia," *Bull. Soc. Géol. France*, vol. 5, pp. 585–595, 1946.

[45] V. Eisenmann and M. Mashkour, "The Small Equids (Perissodactyla, Mammalia) of the Pleistocen of Binagady (Azerbaijan) and Qazvin (Iran): E. hemionus binagadensis nov. subsp. and E. Hydruntinus," *Geobios*, vol. 32 (1), pp. pp. 105–122, 1999. Publisher: Elsevier.

[46] A. Burke, V. Eisenmann, and G. K. Ambler, "The systematic position of Equus hydruntinus, an extinct species of Pleistocene equid," *Quaternary Research*, vol. 59, pp. 459–469, May 2003.

[47] A. Forsten, "A review of Equus stenonis Cocchi (Perissodactyla, Equidae) and related forms," *Quaternary Science Reviews*, vol. 18, pp. 1373–1408, Oct. 1999.

[48] S. J. Davis, "Late Pleistocene and Holocene equid remains from Israel," *Zoological Journal of the Linnean Society*, vol. 70, pp. 289–312, Nov. 1980.

[49] C. Nores, A. M. Muñiz, L. L. Rodríguez, E. A. Bennett, and E.-M. Geigl, "The Iberian zebro: what kind of a beast was it?," *Anthropozoologica*, vol. 50, pp. 21–32, June 2015. Publisher: Muséum national d'Histoire naturelle, Paris.

[50] A. Forsten and R. Ziegler, *The horses (Mammalia, Equidae) from the early Wuermian of Villa Seckendorff, Stuttgart-Bad Cannstatt, Germany*. Staatliches Museum für Naturkunde, 1995. Google-Books-ID: XGBrmQEACAAJ.

[51] C. N. Quesada and C. Liesau, "La Zoología Histórica como complemento de

la Arqueozoología. El caso de Zebro," *Archaeofauna*, pp. 61–71, Oct. 1992. Number: 1.

[52] R. Higuchi, B. Bowman, M. Freiberger, O. A. Ryder, and A. C. Wilson, "DNA sequences from the quagga, an extinct member of the horse family," *Nature*, vol. 312, pp. 282–284, Nov. 1984. Number: 5991 Publisher: Nature Publishing Group.

[53] A. W. Briggs, U. Stenzel, P. L. F. Johnson, R. E. Green, J. Kelso, K. Prüfer, M. Meyer, J. Krause, M. T. Ronan, M. Lachmann, and S. Pääbo, "Patterns of damage in genomic DNA sequences from a Neandertal," *Proceedings of the National Academy of Sciences*, vol. 104, pp. 14616–14621, Sept. 2007. Publisher: National Academy of Sciences Section: Biological Sciences.

[54] D. Y. Yang, B. Eng, J. S. Waye, J. C. Dudar, and S. R. Saunders, "Improved DNA extraction from ancient bones using silica-based spin columns," *American Journal of Physical Anthropology*, vol. 105, no. 4, pp. 539–543, 1998. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/%28SICI%291096-8644%28199804%29105%3A4%3C539%3A%3AAID-AJPA10%3E3.0.CO%3B2-1.

[55] J. Dabney, M. Knapp, I. Glocke, M.-T. Gansauge, A. Weihmann, B. Nickel, C. Valdiosera, N. García, S. Pääbo, J.-L. Arsuaga, and M. Meyer, "Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments," *Proceedings of the National Academy of Sciences*, vol. 110, pp. 15758–15763, Sept. 2013. Publisher: National Academy of Sciences Section: Biological Sciences.

[56] N. Rohland and M. Hofreiter, "Comparison and optimization of ancient DNA extraction," *BioTechniques*, vol. 42, pp. 343–352, Mar. 2007.

[57] N. Rohland, I. Glocke, A. Aximu-Petri, and M. Meyer, "Extraction of highly degraded DNA from ancient bones, teeth and sediments for high-throughput sequencing," *Nature Protocols*, vol. 13, pp. 2447–2461, Nov. 2018. Number: 11 Publisher: Nature Publishing Group.

[58] L. Orlando, R. Allaby, P. Skoglund, C. Der Sarkissian, P. W. Stockhammer,

M. C. Ávila Arcos, Q. Fu, J. Krause, E. Willerslev, A. C. Stone, and C. Warinner, "Ancient DNA analysis," *Nature Reviews Methods Primers*, vol. 1, pp. 1–26, Feb. 2021. Number: 1 Publisher: Nature Publishing Group.

[59] N. Rohland, E. Harney, S. Mallick, S. Nordenfelt, and D. Reich, "Partial uracil–DNA–glycosylase treatment for screening of ancient DNA," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 370, p. 20130624, Jan. 2015. Publisher: Royal Society.

[60] J. Krause, A. W. Briggs, M. Kircher, T. Maricic, N. Zwyns, A. Derevianko, and S. Pääbo, "A Complete mtDNA Genome of an Early Modern Human from Kostenki, Russia," *Current Biology*, vol. 20, pp. 231–236, Feb. 2010.

[61] L. Orlando, V. Eisenmann, F. Reynier, P. Sondaar, and C. Hänni, "Morphological Convergence in Hippidion and Equus (Amerhippus) South American Equids Elucidated by Ancient DNA Analysis," *Journal of Molecular Evolution*, vol. 57, pp. S29–S40, Aug. 2003.

[62] J. Weinstock, E. Willerslev, A. Sher, W. Tong, S. Y. W. Ho, D. Rubenstein, J. Storer, J. Burns, L. Martin, C. Bravi, A. Prieto, D. Froese, E. Scott, L. Xulong, and A. Cooper, "Evolution, Systematics, and Phylogeography of Pleistocene Horses in the New World: A Molecular Perspective," *PLOS Biology*, vol. 3, p. e241, June 2005. Publisher: Public Library of Science.

[63] L. Orlando, A. Ginolhac, G. Zhang, D. Froese, A. Albrechtsen, M. Stiller, M. Schubert, E. Cappellini, B. Petersen, I. Moltke, P. L. F. Johnson, M. Fumagalli, J. T. Vilstrup, M. Raghavan, T. Korneliussen, A.-S. Malaspinas, J. Vogt, D. Szklarczyk, C. D. Kelstrup, J. Vinther, A. Dolocan, J. Stenderup, A. M. V. Velazquez, J. Cahill, M. Rasmussen, X. Wang, J. Min, G. D. Zazula, A. Seguin-Orlando, C. Mortensen, K. Magnussen, J. F. Thompson, J. Weinstock, K. Gregersen, K. H. Røed, V. Eisenmann, C. J. Rubin, D. C. Miller, D. F. Antczak, M. F. Bertelsen, S. Brunak, K. A. S. Al-Rasheid, O. Ryder, L. Andersson, J. Mundy, A. Krogh, M. T. P. Gilbert, K. Kjær, T. Sicheritz-Ponten, L. J. Jensen, J. V. Olsen, M. Hofreiter, R. Nielsen, B. Shapiro, J. Wang, and E. Willerslev, "Recalibrating Equus evolution using the genome sequence of an early Middle Pleistocene horse," *Nature*, vol. 499, pp. 74–78, July

2013. Bandiera_abtest: a Cg_type: Nature Research Journals Number: 7456 Primary_atype: Research Publisher: Nature Publishing Group Subject_term: Evolutionary genetics Subject_term_id: evolutionary-genetics.

[64] C. Der Sarkissian, L. Ermini, M. Schubert, M. A. Yang, P. Librado, M. Fumagalli, H. Jónsson, G. K. Bar-Gal, A. Albrechtsen, F. G. Vieira, B. Petersen, A. Ginolhac, A. Seguin-Orlando, K. Magnussen, A. Fages, C. Gamba, B. Lorente-Galdos, S. Polani, C. Steiner, M. Neuditschko, V. Jagannathan, C. Feh, C. L. Greenblatt, A. Ludwig, N. I. Abramson, W. Zimmermann, R. Schafberg, A. Tikhonov, T. Sicheritz-Ponten, E. Willerslev, T. Marques-Bonet, O. A. Ryder, M. McCue, S. Rieder, T. Leeb, M. Slatkin, and L. Orlando, "Evolutionary Genomics and Conservation of the Endangered Przewalski's Horse," *Current Biology*, vol. 25, pp. 2577–2583, Oct. 2015. Publisher: Elsevier.

[65] M. Schubert, H. Jónsson, D. Chang, C. Der Sarkissian, L. Ermini, A. Ginolhac, A. Albrechtsen, I. Dupanloup, A. Foucal, B. Petersen, M. Fumagalli, M. Raghavan, A. Seguin-Orlando, T. S. Korneliussen, A. M. V. Velazquez, J. Stenderup, C. A. Hoover, C.-J. Rubin, A. H. Alfarhan, S. A. Alquraishi, K. A. S. Al-Rasheid, D. E. MacHugh, T. Kalbfleisch, J. N. MacLeod, E. M. Rubin, T. Sicheritz-Ponten, L. Andersson, M. Hofreiter, T. Marques-Bonet, M. T. P. Gilbert, R. Nielsen, L. Excoffier, E. Willerslev, B. Shapiro, and L. Orlando, "Prehistoric genomes reveal the genetic foundation and cost of horse domestication," *Proceedings of the National Academy of Sciences*, vol. 111, pp. E5661–E5669, Dec. 2014. Publisher: Proceedings of the National Academy of Sciences.

[66] A. Ludwig, M. Reissmann, N. Benecke, R. Bellone, E. Sandoval-Castellanos, M. Cieslak, G. G. Fortes, A. Morales-Muñiz, M. Hofreiter, and M. Pruvost, "Twenty-five thousand years of fluctuating selection on leopard complex spotting and congenital night blindness in horses," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 370, p. 20130386, Jan. 2015. Publisher: Royal Society.

[67] T. Jansen, P. Forster, M. A. Levine, H. Oelke, M. Hurles, C. Renfrew, J. We-

ber, and K. Olek, "Mitochondrial DNA and the origins of the domestic horse," *Proceedings of the National Academy of Sciences*, vol. 99, pp. 10905–10910, Aug. 2002. Publisher: Proceedings of the National Academy of Sciences.

[68] J. Lira, A. Linderholm, C. Olaria, M. Brandström Durling, M. T. P. Gilbert, H. Ellegren, E. Willerslev, K. Lidén, J. L. Arsuaga, and A. Götherström, "Ancient DNA reveals traces of Iberian Neolithic and Bronze Age lineages in modern Iberian horses," *Molecular Ecology*, vol. 19, no. 1, pp. 64–78, 2010. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-294X.2009.04430.x.

[69] D. W. Anthony and D. Ringe, "The Indo-European Homeland from Linguistic and Archaeological Perspectives," *Annual Review of Linguistics*, vol. 1, no. 1, pp. 199–219, 2015. _eprint: https://doi.org/10.1146/annurev-linguist-030514-124812.

[70] E. T. Shev, "THE INTRODUCTION OF THE DOMESTICATED HORSE IN SOUTHWEST ASIA," *Archaeology, Ethnology & Anthropology of Eurasia*, vol. 44, pp. 123–136, Mar. 2016. Number: 1.

[71] M. A. Zeder, "Domestication and early agriculture in the Mediterranean Basin: Origins, diffusion, and impact," *Proceedings of the National Academy of Sciences*, vol. 105, pp. 11597–11604, Aug. 2008. Publisher: Proceedings of the National Academy of Sciences.

[72] B. S. Arbuckle, "Animals in the ancient world," in *A Companion to the Archaeology of the Ancient Near East* (D. T. Potts, ed.), pp. 201–219, Wiley, 2012.

[73] S. Guimaraes, B. S. Arbuckle, J. Peters, S. E. Adcock, H. Buitenhuis, H. Chazin, N. Manaseryan, H.-P. Uerpmann, T. Grange, and E.-M. Geigl, "Ancient DNA shows domestic horses were introduced in the southern Caucasus and Anatolia during the Bronze Age," *Science Advances*, vol. 6, p. eabb0030, Sept. 2020. Publisher: American Association for the Advancement of Science.

[74] A. K. Outram, N. A. Stear, R. Bendrey, S. Olsen, A. Kasparov, V. Zaibert, N. Thorpe, and R. P. Evershed, "The Earliest Horse Harnessing and Milking,"

Science, vol. 323, pp. 1332–1335, Mar. 2009. Publisher: American Association for the Advancement of Science.

[75] C. Gaunitz, A. Fages, K. Hanghøj, A. Albrechtsen, N. Khan, M. Schubert, A. Seguin-Orlando, I. J. Owens, S. Felkel, O. Bignon-Lau, P. de Barros Damgaard, A. Mittnik, A. F. Mohaseb, H. Davoudi, S. Alquraishi, A. H. Alfarhan, K. A. S. Al-Rasheid, E. Crubézy, N. Benecke, S. Olsen, D. Brown, D. Anthony, K. Massy, V. Pitulko, A. Kasparov, G. Brem, M. Hofreiter, G. Mukhtarova, N. Baimukhanov, L. Lõugas, V. Onar, P. W. Stockhammer, J. Krause, B. Boldgiv, S. Undrakhbold, D. Erdenebaatar, S. Lepetz, M. Mashkour, A. Ludwig, B. Wallner, V. Merz, I. Merz, V. Zaibert, E. Willerslev, P. Librado, A. K. Outram, and L. Orlando, "Ancient genomes revisit the ancestry of domestic and Przewalski's horses," *Science*, vol. 360, pp. 111–114, Apr. 2018. Publisher: American Association for the Advancement of Science.

[76] P. Librado, N. Khan, A. Fages, M. A. Kusliy, T. Suchan, L. Tonasso-Calvière, S. Schiavinato, D. Alioglu, A. Fromentier, A. Perdereau, J.-M. Aury, C. Gaunitz, L. Chauvey, A. Seguin-Orlando, C. Der Sarkissian, J. Southon, B. Shapiro, A. A. Tishkin, A. A. Kovalev, S. Alquraishi, A. H. Alfarhan, K. A. S. Al-Rasheid, T. Seregély, L. Klassen, R. Iversen, O. Bignon-Lau, P. Bodu, M. Olive, J.-C. Castel, M. Boudadi-Maligne, N. Alvarez, M. Germonpré, M. Moskal-del Hoyo, J. Wilczyński, S. Pospuła, A. Lasota-Kuś, K. Tunia, M. Nowak, E. Rannamäe, U. Saarma, G. Boeskorov, L. Lõugas, R. Kyselý, L. Peške, A. Bălăşescu, V. Dumitraşcu, R. Dobrescu, D. Gerber, V. Kiss, A. Szécsényi-Nagy, B. G. Mende, Z. Gallina, K. Somogyi, G. Kulcsár, E. Gál, R. Bendrey, M. E. Allentoft, G. Sirbu, V. Dergachev, H. Shephard, N. Tomadini, S. Grouard, A. Kasparov, A. E. Basilyan, M. A. Anisimov, P. A. Nikolskiy, E. Y. Pavlova, V. Pitulko, G. Brem, B. Wallner, C. Schwall, M. Keller, K. Kitagawa, A. N. Bessudnov, A. Bessudnov, W. Taylor, J. Magail, J.-O. Gantulga, J. Bayarsaikhan, D. Erdenebaatar, K. Tabaldiev, E. Mijiddorj, B. Boldgiv, T. Tsagaan, M. Pruvost, S. Olsen, C. A. Makarewicz, S. Valenzuela Lamas, S. Albizuri Canadell, A. Nieto Espinet, M. P. Iborra, J. Lira Garrido, E. Rodríguez González, S. Celestino, C. Olària, J. L. Arsuaga, N. Kotova, A. Pryor, P. Crabtree, R. Zhumatayev, A. Toleubaev, N. L. Morgunova, T. Kuznetsova,

D. Lordkipanize, M. Marzullo, O. Prato, G. Bagnasco Gianni, U. Tecchiati, B. Clavel, S. Lepetz, H. Davoudi, M. Mashkour, N. Y. Berezina, P. W. Stockhammer, J. Krause, W. Haak, A. Morales-Muñiz, N. Benecke, M. Hofreiter, A. Ludwig, A. S. Graphodatsky, J. Peters, K. Y. Kiryushin, T.-O. Iderkhangai, N. A. Bokovenko, S. K. Vasiliev, N. N. Seregin, K. V. Chugunov, N. A. Plasteeva, G. F. Baryshnikov, E. Petrova, M. Sablin, E. Ananyevskaya, A. Logvin, I. Shevnina, V. Logvin, S. Kalieva, V. Loman, I. Kukushkin, I. Merz, V. Merz, S. Sakenov, V. Varfolomeyev, E. Usmanova, V. Zaibert, B. Arbuckle, A. B. Belinskiy, A. Kalmykov, S. Reinhold, S. Hansen, A. I. Yudin, A. A. Vybornov, A. Epimakhov, N. S. Berezina, N. Roslyakova, P. A. Kosintsev, P. F. Kuznetsov, D. Anthony, G. J. Kroonen, K. Kristiansen, P. Wincker, A. Outram, and L. Orlando, "The origins and spread of domestic horses from the Western Eurasian steppes," *Nature*, pp. 1–7, Oct. 2021. Bandiera_abtest: a Cc_license_type: cc_by Cg_type: Nature Research Journals Primary_atype: Research Publisher: Nature Publishing Group Subject_term: Evolutionary genetics;Population genetics Subject_term_id: evolutionary-genetics;population-genetics.

[77] A. Beja-Pereira, P. R. England, N. Ferrand, S. Jordan, A. O. Bakhiet, M. A. Abdalla, M. Mashkour, J. Jordana, P. Taberlet, and G. Luikart, "African Origins of the Domestic Donkey," *Science*, vol. 304, pp. 1781–1781, June 2004. Number: 5678.

[78] B. Kimura, F. B. Marshall, S. Chen, S. Rosenbom, P. D. Moehlman, N. Tuross, R. C. Sabin, J. Peters, B. Barich, H. Yohannes, F. Kebede, R. Teclai, A. Beja-Pereira, and C. J. Mulligan, "Ancient DNA from Nubian and Somali wild ass provides insights into donkey ancestry and domestication," *Proceedings of the Royal Society B: Biological Sciences*, vol. 278, pp. 50–57, Jan. 2011. Number: 1702.

[79] E. T. Todd, L. Tonasso-Calvière, L. Chauvey, S. Schiavinato, A. Fages, A. Seguin-Orlando, P. Clavel, N. Khan, L. Pérez Pardal, L. Patterson Rosa, P. Librado, H. Ringbauer, M. Verdugo, J. Southon, J.-M. Aury, A. Perdereau, E. Vila, M. Marzullo, O. Prato, U. Tecchiati, G. Bagnasco Gianni, A. Tagliacozzo, V. Tinè, F. Alhaique, J. L. Cardoso, M. J. Valente, M. Telles Antunes, L. Frantz, B. Shapiro, D. G. Bradley, N. Boulbes, A. Gardeisen, L. K. Hor-

witz, A. Öztan, B. S. Arbuckle, V. Onar, B. Clavel, S. Lepetz, A. A. Vahdati, H. Davoudi, A. Mohaseb, M. Mashkour, O. Bouchez, C. Donnadieu, P. Wincker, S. A. Brooks, A. Beja-Pereira, D.-D. Wu, and L. Orlando, "The genomic history and global expansion of domestic donkeys," *Science*, vol. 377, pp. 1172–1180, Sept. 2022. Publisher: American Association for the Advancement of Science.

[80] S. Rosenbom, V. Costa, N. Al-Araimi, E. Kefena, A. S. Abdel-Moneim, M. A. Abdalla, A. Bakhiet, and A. Beja-Pereira, "Genetic diversity of donkey populations from the putative centers of domestication," *Animal Genetics*, vol. 46, no. 1, pp. 30–36, 2015. Number: 1.

[81] E. Hempel, F. Bibi, J. T. Faith, K.-P. Koepfli, A. M. Klittich, D. A. Duchêne, J. S. Brink, D. C. Kalthoff, L. Dalén, M. Hofreiter, and M. V. Westbury, "Blue turns to grey - Palaeogenomic insights into the evolutionary history and extinction of the blue antelope (Hippotragus leucophaeus)," *Molecular Biology and Evolution*, p. msac241, Nov. 2022.

[82] M. M. Ciucani, J. K. Jensen, M.-H. S. Sinding, O. Smith, S. B. Lucenti, E. Rosengren, L. Rook, C. Tuveri, M. Arca, E. Cappellini, M. Galaverni, E. Randi, C. Guo, G. Zhang, T. Sicheritz-Pontén, L. Dalén, M. T. P. Gilbert, and S. Gopalakrishnan, "Evolutionary history of the extinct Sardinian dhole," *Current Biology*, vol. 0, Oct. 2021. Publisher: Elsevier.

[83] J. T. Vilstrup, A. Seguin-Orlando, M. Stiller, A. Ginolhac, M. Raghavan, S. C. A. Nielsen, J. Weinstock, D. Froese, S. K. Vasiliev, N. D. Ovodov, J. Clary, K. M. Helgen, R. C. Fleischer, A. Cooper, B. Shapiro, and L. Orlando, "Mitochondrial Phylogenomics of Modern and Ancient Equids," *PLOS ONE*, vol. 8, p. e55950, Feb. 2013. Publisher: Public Library of Science.

[84] D. Cai, S. Zhu, M. Gong, N. Zhang, J. Wen, Q. Liang, W. Sun, X. Shao, Y. Guo, Y. Cai, Z. Zheng, W. Zhang, S. Hu, X. Wang, H. Tian, Y. Li, W. Liu, M. Yang, J. Yang, D. Wu, L. Orlando, and Y. Jiang, "Radiocarbon and genomic evidence for the survival of Equus Sussemionus until the late Holocene," *eLife*, vol. 11, p. e73346, May 2022. Publisher: eLife Sciences Publications, Ltd.

[85] T. van der Valk, P. Pečnerová, D. Díez-del Molino, A. Bergström, J. Oppenheimer, S. Hartmann, G. Xenikoudakis, J. A. Thomas, M. Dehasque, E. Sağlıcan, F. R. Fidan, I. Barnes, S. Liu, M. Somel, P. D. Heintzman, P. Nikolskiy, B. Shapiro, P. Skoglund, M. Hofreiter, A. M. Lister, A. Götherström, and L. Dalén, "Million-year-old DNA sheds light on the genomic history of mammoths," *Nature*, vol. 591, pp. 265–269, Mar. 2021. Number: 7849 Publisher: Nature Publishing Group.

[86] R. L. Rogers and M. Slatkin, "Excess of genomic defects in a woolly mammoth on Wrangel island," *PLOS Genetics*, vol. 13, p. e1006601, Mar. 2017. Publisher: Public Library of Science.

[87] D. P. L. Toews and A. Brelsford, "The biogeography of mitochondrial and nuclear discordance in animals," *Molecular Ecology*, vol. 21, no. 16, pp. 3907–3930, 2012. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-294X.2012.05664.x.

[88] K. Pawłowska and L.-M. Shillito, "An Integrated Zooarchaeological and Micromorphological Perspective on Midden Taphonomy at Late Neolithic Çatalhöyük," *Open Archaeology*, vol. 8, pp. 436–459, Jan. 2022. Publisher: De Gruyter Open Access.

[89] S. R. Steadman, L. D. Hackley, S. Selover, B. Yıldırım, M. von Baeyer, B. S. Arbuckle, R. Robinson, and A. Smith, "Early Lives: The Late Chalcolithic and Early Bronze Age at Çadır Höyük," *Journal of Eastern Mediterranean Archaeology and Heritage Studies*, vol. 7, pp. 271–298, Sept. 2019. Publisher: Duke University Press.

[90] S. R. Steadman, B. S. Arbuckle, and G. McMahon, "Pivoting East: Çadır Höyük, Transcaucasia, and Complex Connectivity in the Late Chalcolithic," *Documenta Praehistorica*, vol. 45, pp. 64–85, Dec. 2018.

[91] S. Steadman, G. McMahon, E. Şerifoğlu, M. Cassis, A. Lauricella, L. Hackley, S. Selover, B. Yıldırım, B. Arbuckle, M. Von Baeyer, K. Tardio, S. Adcock, E. Dinç, G. Özger, B. Selvi, S. Offutt, and A. Hartley, "The 2017-2018 Seasons at Çadir Höyük on the North Central Plateau," *Anatolica*, vol. 45, pp. 77–119, 2019. Num Pages: 42 Publisher: Peeters Publishers.

[92] S. R. Steadman and G. McMahon, "The 2015-2016 Seasons at ÇadırHöyük on the North Central Anatolian Plateau," in *The Archaeology of Anatolia Volume II: Recent Discoveries (2015-2016)* (S. R. Steadman and G. McMahon, eds.), pp. 94–117, Cambridge Scholars Publishing, 2017.

[93] J. C. Ross, G. McMahon, Y. Heffron, S. E. Adcock, S. R. Steadman, B. S. Arbuckle, A. Smith, and M. von Baeyer, "Anatolian Empires: Local Experiences from Hittites to Phrygians at Çadır Höyük," *Journal of Eastern Mediterranean Archaeology and Heritage Studies*, vol. 7, pp. 299–320, Sept. 2019. Publisher: Duke University Press.

[94] J. C. Ross, S. R. Steadman, G. McMahon, S. E. Adcock, and J. W. Cannon, "When the Giant Falls: Endurance and Adaptation at Çadır Höyük in the Context of the Hittite Empire and Its Collapse," *Journal of Field Archaeology*, vol. 44, pp. 19–39, Jan. 2019. Publisher: Routledge _eprint: https://doi.org/10.1080/00934690.2018.1558906.

[95] W. F. Libby, "Atmospheric Helium Three and Radiocarbon from Cosmic Radiation," *Physical Review*, vol. 69, pp. 671–672, June 1946. Publisher: American Physical Society.

[96] P. J. Reimer, W. E. N. Austin, E. Bard, A. Bayliss, P. G. Blackwell, C. B. Ramsey, M. Butzin, H. Cheng, R. L. Edwards, M. Friedrich, P. M. Grootes, T. P. Guilderson, I. Hajdas, T. J. Heaton, A. G. Hogg, K. A. Hughen, B. Kromer, S. W. Manning, R. Muscheler, J. G. Palmer, C. Pearson, J. v. d. Plicht, R. W. Reimer, D. A. Richards, E. M. Scott, J. R. Southon, C. S. M. Turney, L. Wacker, F. Adolphi, U. Büntgen, M. Capano, S. M. Fahrni, A. Fogtmann-Schulz, R. Friedrich, P. Köhler, S. Kudsk, F. Miyake, J. Olsen, F. Reinig, M. Sakamoto, A. Sookdeo, and S. Talamo, "The IntCal20 Northern Hemisphere Radiocarbon Age Calibration Curve (0–55 cal kBP)," *Radiocarbon*, vol. 62, pp. 725–757, Aug. 2020. Publisher: Cambridge University Press.

[97] M. Meyer and M. Kircher, "Illumina Sequencing Library Preparation for Highly Multiplexed Target Capture and Sequencing," *Cold Spring Harbor Protocols*, vol. 2010, p. pdb.prot5448, June 2010.

[98] M. Schubert, S. Lindgreen, and L. Orlando, "AdapterRemoval v2: rapid adapter trimming, identification, and read merging," *BMC Research Notes*, vol. 9, p. 88, Feb. 2016.

[99] G. M. Kılınç, A. Omrak, F. Özer, T. Günther, A. M. Büyükkarakaya, E. Bıçakçı, D. Baird, H. M. Dönertaş, A. Ghalichi, R. Yaka, D. Koptekin, S. C. Açan, P. Parvizi, M. Krzewińska, E. A. Daskalaki, E. Yüncü, N. D. Dağtaş, A. Fairbairn, J. Pearson, G. Mustafaoğlu, Y. S. Erdal, Y. G. Çakan, Togan, M. Somel, J. Storå, M. Jakobsson, and A. Götherström, "The Demographic Development of the First Farmers in Anatolia," *Current Biology*, vol. 26, pp. 2659–2666, Oct. 2016. Publisher: Elsevier.

[100] T. Günther, C. Valdiosera, H. Malmström, I. Ureña, R. Rodriguez-Varela, O. Sverrisdóttir, E. A. Daskalaki, P. Skoglund, T. Naidoo, E. M. Svensson, J. M. B. d. Castro, E. Carbonell, M. Dunn, J. Storå, E. Iriarte, J. L. Arsuaga, J.-M. Carretero, A. Götherström, and M. Jakobsson, "Ancient genomes link early farmers from Atapuerca in Spain to modern-day Basques," *Proceedings of the National Academy of Sciences*, vol. 112, pp. 11917–11922, Sept. 2015. Publisher: National Academy of Sciences Section: Biological Sciences.

[101] C. M. Wade, E. Giulotto, S. Sigurdsson, M. Zoli, S. Gnerre, F. Imsland, T. L. Lear, D. L. Adelson, E. Bailey, R. R. Bellone, H. Blöcker, O. Distl, R. C. Edgar, M. Garber, T. Leeb, E. Mauceli, J. N. MacLeod, M. C. T. Penedo, J. M. Raison, T. Sharpe, J. Vogel, L. Andersson, D. F. Antczak, T. Biagi, M. M. Binns, B. P. Chowdhary, S. J. Coleman, G. Della Valle, S. Fryc, G. Guérin, T. Hasegawa, E. W. Hill, J. Jurka, A. Kiialainen, G. Lindgren, J. Liu, E. Magnani, J. R. Mickelson, J. Murray, S. G. Nergadze, R. Onofrio, S. Pedroni, M. F. Piras, T. Raudsepp, M. Rocchi, K. H. Røed, O. A. Ryder, S. Searle, L. Skow, J. E. Swinburne, A. C. Syvänen, T. Tozaki, S. J. Valberg, M. Vaudin, J. R. White, M. C. Zody, Broad Institute Genome Sequencing Platform, Broad Institute Whole Genome Assembly Team, E. S. Lander, and K. Lindblad-Toh, "Genome Sequence, Comparative Analysis, and Population Genetics of the Domestic Horse," *Science*, vol. 326, pp. 865–867, Nov. 2009. Publisher: American Association for the Advancement of Science.

[102] H. Li and R. Durbin, "Fast and accurate short read alignment with Burrows–Wheeler transform," *Bioinformatics*, vol. 25, pp. 1754–1760, July 2009.

[103] W. Xu, Y. Lin, K. Zhao, H. Li, Y. Tian, J. N. Ngatia, Y. Ma, S. K. Sahu, H. Guo, X. Guo, Y. C. Xu, H. Liu, K. Kristiansen, T. Lan, and X. Zhou, "An efficient pipeline for ancient DNA mapping and recovery of endogenous ancient DNA from whole-genome sequencing data," *Ecology and Evolution*, vol. 11, pp. 390–401, Jan. 2021.

[104] A. Oliva, R. Tobler, B. Llamas, and Y. Souilmi, "BWA-mem is not the best aligner for ancient DNA short reads," tech. rep., Aug. 2021. Company: Cold Spring Harbor Laboratory Distributor: Cold Spring Harbor Laboratory Label: Cold Spring Harbor Laboratory Section: Contradictory Results Type: article.

[105] H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, and 1000 Genome Project Data Processing Subgroup, "The Sequence Alignment/Map format and SAMtools," *Bioinformatics*, vol. 25, pp. 2078–2079, Aug. 2009.

[106] P. Skoglund, B. H. Northoff, M. V. Shunkov, A. P. Derevianko, S. Pääbo, J. Krause, and M. Jakobsson, "Separating endogenous ancient DNA from modern day contamination in a Siberian Neandertal," *Proceedings of the National Academy of Sciences*, vol. 111, pp. 2229–2234, Feb. 2014. Publisher: National Academy of Sciences Section: Biological Sciences.

[107] J. S. Pedersen, E. Valen, A. M. V. Velazquez, B. J. Parker, M. Rasmussen, S. Lindgreen, B. Lilje, D. J. Tobin, T. K. Kelly, S. Vang, R. Andersson, P. A. Jones, C. A. Hoover, A. Tikhonov, E. Prokhortchouk, E. M. Rubin, A. Sandelin, M. T. P. Gilbert, A. Krogh, E. Willerslev, and L. Orlando, "Genome-wide nucleosome map and cytosine methylation levels of an ancient human genome," *Genome Research*, vol. 24, pp. 454–466, Mar. 2014. Company: Cold Spring Harbor Laboratory Press Distributor: Cold Spring Harbor Laboratory Press Institution: Cold Spring Harbor Laboratory Press Label: Cold Spring Harbor Laboratory Press Publisher: Cold Spring Harbor Lab.

[108] G. Jun, M. K. Wing, G. R. Abecasis, and H. M. Kang, "An efficient and scalable analysis framework for variant extraction and refinement from population-

scale DNA sequence data," *Genome Research*, vol. 25, pp. 918–925, June 2015. Company: Cold Spring Harbor Laboratory Press Distributor: Cold Spring Harbor Laboratory Press Institution: Cold Spring Harbor Laboratory Press Label: Cold Spring Harbor Laboratory Press Publisher: Cold Spring Harbor Lab.

[109] M. Schubert, M. Mashkour, C. Gaunitz, A. Fages, A. Seguin-Orlando, S. Sheikhi, A. H. Alfarhan, S. A. Alquraishi, K. A. S. Al-Rasheid, R. Chuang, L. Ermini, C. Gamba, J. Weinstock, O. Vedat, and L. Orlando, "Zonkey: A simple, accurate and sensitive pipeline to genetically identify equine F1-hybrids in archaeological assemblages," *Journal of Archaeological Science*, vol. 78, pp. 147–157, Feb. 2017.

[110] N. Patterson, A. L. Price, and D. Reich, "Population Structure and Eigenanalysis," *PLOS Genetics*, vol. 2, p. e190, Dec. 2006. Publisher: Public Library of Science.

[111] J. K. Pickrell and J. K. Pritchard, "Inference of Population Splits and Mixtures from Genome-Wide Allele Frequency Data," *PLOS Genetics*, vol. 8, p. e1002967, Nov. 2012. Publisher: Public Library of Science.

[112] T. S. Korneliussen, A. Albrechtsen, and R. Nielsen, "ANGSD: Analysis of Next Generation Sequencing Data," *BMC Bioinformatics*, vol. 15, p. 356, Nov. 2014.

[113] S. Kumar, G. Stecher, M. Li, C. Knyaz, and K. Tamura, "MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms," *Molecular Biology and Evolution*, vol. 35, pp. 1547–1549, June 2018.

[114] Y. Tateno, N. Takezaki, and M. Nei, "Relative efficiencies of the maximum-likelihood, neighbor-joining, and maximum-parsimony methods when substitution rate varies with site.," *Molecular Biology and Evolution*, vol. 11, pp. 261–277, Mar. 1994.

[115] S. Guindon and O. Gascuel, "Efficient Biased Estimation of Evolutionary Distances When Substitution Rates Vary Across Sites," *Molecular Biology and Evolution*, vol. 19, pp. 534–543, Apr. 2002.

[116] R. Yoshida and M. Nei, "Efficiencies of the NJp, Maximum Likelihood, and Bayesian Methods of Phylogenetic Construction for Compositional and Non-compositional Genes," *Molecular Biology and Evolution*, vol. 33, pp. 1618–1624, June 2016.

[117] S. Guindon and O. Gascuel, "A Simple, Fast, and Accurate Algorithm to Estimate Large Phylogenies by Maximum Likelihood," *Systematic Biology*, vol. 52, pp. 696–704, Oct. 2003.

[118] D. Posada, "jModelTest: Phylogenetic Model Averaging," *Molecular Biology and Evolution*, vol. 25, pp. 1253–1256, July 2008.

[119] M. Hasegawa, H. Kishino, and T. Yano, "Dating of the human-ape splitting by a molecular clock of mitochondrial DNA," *Journal of Molecular Evolution*, vol. 22, no. 2, pp. 160–174, 1985.

[120] F. Rodríguez, J. L. Oliver, A. Marín, and J. R. Medina, "The general stochastic model of nucleotide substitution," *Journal of Theoretical Biology*, vol. 142, pp. 485–501, Feb. 1990.

[121] A. J. Drummond, M. A. Suchard, D. Xie, and A. Rambaut, "Bayesian Phylogenetics with BEAUti and the BEAST 1.7," *Molecular Biology and Evolution*, vol. 29, pp. 1969–1973, Aug. 2012.

[122] A. Rambaut, "FigTree," 2014.

[123] J. Huang, Y. Zhao, D. Bai, W. Shiraigol, B. Li, L. Yang, J. Wu, W. Bao, X. Ren, B. Jin, Q. Zhao, A. Li, S. Bao, W. Bao, Z. Xing, A. An, Y. Gao, R. Wei, Y. Bao, T. Bao, H. Han, H. Bai, Y. Bao, Y. Zhang, D. Daidiikhuu, W. Zhao, S. Liu, J. Ding, W. Ye, F. Ding, Z. Sun, Y. Shi, Y. Zhang, H. Meng, and M. Dugarjaviin, "Donkey genome and insight into the imprinting of fast karyotype evolution," *Scientific Reports*, vol. 5, p. 14106, Sept. 2015. Number: 1.

[124] G. Renaud, B. Petersen, A. Seguin-Orlando, M. F. Bertelsen, A. Waller, R. Newton, R. Paillot, N. Bryant, M. Vaudin, P. Librado, and L. Orlando, "Improved de novo genomic assembly for the domestic donkey," *Science Advances*, vol. 4, p. eaaq0392, Apr. 2018. Number: 4.

[125] C. Wang, H. Li, Y. Guo, J. Huang, Y. Sun, J. Min, J. Wang, X. Fang, Z. Zhao, S. Wang, Y. Zhang, Q. Liu, Q. Jiang, X. Wang, Y. Guo, C. Yang, Y. Wang, F. Tian, G. Zhuang, Y. Fan, Q. Gao, Y. Li, Z. Ju, J. Li, R. Li, M. Hou, G. Yang, G. Liu, W. Liu, J. Guo, S. Pan, G. Fan, W. Zhang, R. Zhang, J. Yu, X. Zhang, Q. Yin, C. Ji, Y. Jin, G. Yue, M. Liu, J. Xu, S. Liu, J. Jordana, A. Noce, M. Amills, D. D. Wu, S. Li, X. Zhou, and J. Zhong, "Donkey genomes provide new insights into domestication and selection for coat color," *Nature Communications*, vol. 11, p. 6014, Dec. 2020. Number: 1.

[126] S. Purcell, B. Neale, K. Todd-Brown, L. Thomas, M. A. R. Ferreira, D. Bender, J. Maller, P. Sklar, P. I. W. d. Bakker, M. J. Daly, and P. C. Sham, "PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses," *The American Journal of Human Genetics*, vol. 81, pp. 559–575, Sept. 2007. Publisher: Elsevier.

[127] D. Reich, K. Thangaraj, N. Patterson, A. L. Price, and L. Singh, "Reconstructing Indian population history," *Nature*, vol. 461, pp. 489–494, Sept. 2009. Number: 7263 Publisher: Nature Publishing Group.

[128] N. Patterson, P. Moorjani, Y. Luo, S. Mallick, N. Rohland, Y. Zhan, T. Genschoreck, T. Webster, and D. Reich, "Ancient Admixture in Human History," *Genetics*, vol. 192, pp. 1065–1093, Nov. 2012.

[129] P. Skoglund, S. Mallick, M. C. Bortolini, N. Chennagiri, T. Hünemeier, M. L. Petzl-Erler, F. M. Salzano, N. Patterson, and D. Reich, "Genetic evidence for two founding populations of the Americas," *Nature*, vol. 525, pp. 104–108, Sept. 2015. Bandiera_abtest: a Cg_type: Nature Research Journals Number: 7567 Primary_atype: Research Publisher: Nature Publishing Group Subject_term: Anthropology;Population genetics Subject_term_id: anthropology;population-genetics.

[130] H. Wickham, *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2016.

[131] R. E. Green, J. Krause, A. W. Briggs, T. Maricic, U. Stenzel, M. Kircher, N. Patterson, H. Li, W. Zhai, M. H.-Y. Fritz, N. F. Hansen, E. Y. Durand, A.-S.

Malaspinas, J. D. Jensen, T. Marques-Bonet, C. Alkan, K. Prüfer, M. Meyer, H. A. Burbano, J. M. Good, R. Schultz, A. Aximu-Petri, A. Butthof, B. Höber, B. Höffner, M. Siegemund, A. Weihmann, C. Nusbaum, E. S. Lander, C. Russ, N. Novod, J. Affourtit, M. Egholm, C. Verna, P. Rudan, D. Brajkovic, Kucan, I. Gušic, V. B. Doronichev, L. V. Golovanova, C. Lalueza-Fox, M. de la Rasilla, J. Fortea, A. Rosas, R. W. Schmitz, P. L. F. Johnson, E. E. Eichler, D. Falush, E. Birney, J. C. Mullikin, M. Slatkin, R. Nielsen, J. Kelso, M. Lachmann, D. Reich, and S. Pääbo, "A Draft Sequence of the Neandertal Genome," *Science*, vol. 328, pp. 710–722, May 2010. Publisher: American Association for the Advancement of Science.

[132] C. M. Schlebusch, H. Malmström, T. Günther, P. Sjödin, A. Coutinho, H. Edlund, A. R. Munters, M. Vicente, M. Steyn, H. Soodyall, M. Lombard, and M. Jakobsson, "Southern African ancient genomes estimate modern human divergence to 350,000 to 260,000 years ago," *Science*, vol. 358, pp. 652–655, Nov. 2017. Publisher: American Association for the Advancement of Science.

[133] K. Mualim, C. Theunert, and M. Slatkin, "Estimation of coalescence probabilities and population divergence times from SNP data," *Heredity*, vol. 127, pp. 1–9, July 2021. Number: 1 Publisher: Nature Publishing Group.

[134] F. Baumdicker, G. Bisschop, D. Goldstein, G. Gower, A. P. Ragsdale, G. Tsambos, S. Zhu, B. Eldon, E. C. Ellerman, J. G. Galloway, A. L. Gladstein, G. Gorjanc, B. Guo, B. Jeffery, W. W. Kretzschumar, K. Lohse, M. Matschiner, D. Nelson, N. S. Pope, C. D. Quinto-Cortés, M. F. Rodrigues, K. Saunack, T. Sellinger, K. Thornton, H. van Kemenade, A. W. Wohns, Y. Wong, S. Gravel, A. D. Kern, J. Koskela, P. L. Ralph, and J. Kelleher, "Efficient ancestry and mutation simulation with msprime 1.0," *Genetics*, vol. 220, p. iyab229, Mar. 2022.

[135] F. C. Ceballos, K. Gürün, N. E. Altınışık, H. C. Gemici, C. Karamurat, D. Koptekin, K. B. Vural, I. Mapelli, E. Sağlıcan, E. Sürer, Y. S. Erdal, A. Götherström, F. Özer, Atakuman, and M. Somel, "Human inbreeding has decreased in time through the Holocene," *Current Biology*, vol. 31, pp. 3925–3934.e8, Sept. 2021.

[136] F. C. Ceballos, P. K. Joshi, D. W. Clark, M. Ramsay, and J. F. Wilson, "Runs of homozygosity: windows into population history and trait architecture," *Nature Reviews Genetics*, vol. 19, pp. 220–234, Apr. 2018. Number: 4 Publisher: Nature Publishing Group.

[137] C. Kleinhans and S. Willows-Munro, "Low genetic diversity and shallow population structure in the endangered vulture, Gyps coprotheres," *Scientific Reports*, vol. 9, p. 5536, Apr. 2019. Number: 1 Publisher: Nature Publishing Group.

[138] E. Yurtman, O. Özer, E. Yüncü, N. D. Dağtaş, D. Koptekin, Y. G. Çakan, M. Özkan, A. Akbaba, D. Kaptan, G. Atağ, K. B. Vural, C. Y. Gündem, L. Martin, G. M. Kılınç, A. Ghalichi, S. C. Açan, R. Yaka, E. Sağlıcan, V. K. Lagerholm, M. Krzewińska, T. Günther, P. Morell Miranda, E. Pişkin, M. Şevketoğlu, C. C. Bilgin, Atakuman, Y. S. Erdal, E. Sürer, N. E. Altınışık, J. A. Lenstra, S. Yorulmaz, M. F. Abazari, J. Hoseinzadeh, D. Baird, E. Bıçakçı, Çevik, F. Gerritsen, R. Özbal, A. Götherström, M. Somel, Togan, and F. Özer, "Archaeogenetic analysis of Neolithic sheep from Anatolia suggests a complex demographic history since domestication," *Communications Biology*, vol. 4, pp. 1–11, Nov. 2021. Number: 1 Publisher: Nature Publishing Group.

[139] P. Pečnerová, D. Díez-del Molino, N. Dussex, T. Feuerborn, J. von Seth, J. van der Plicht, P. Nikolskiy, A. Tikhonov, S. Vartanyan, and L. Dalén, "Genome-Based Sexing Provides Clues about Behavior and Social Structure in the Woolly Mammoth," *Current Biology*, vol. 27, pp. 3505–3510.e3, Nov. 2017.

[140] G. Gower, L. E. Fenderson, A. T. Salis, K. M. Helgen, A. L. van Loenen, H. Heiniger, E. Hofman-Kamińska, R. Kowalczyk, K. J. Mitchell, B. Llamas, and A. Cooper, "Widespread male sex bias in mammal fossil and museum collections," *Proceedings of the National Academy of Sciences*, vol. 116, pp. 19019–19024, Sept. 2019. Publisher: Proceedings of the National Academy of Sciences.

[141] A. Fages, A. Seguin-Orlando, M. Germonpré, and L. Orlando, "Horse males became over-represented in archaeological assemblages during the Bronze

Age," *Journal of Archaeological Science: Reports*, vol. 31, p. 102364, June 2020.

[142] J. Weinstock, "Osteometry as a Source of Refined Demographic Information: Sex-Ratios of Reindeer, Hunting Strategies, and Herd Control in the Late Glacial site of Stellmoor, Northern Germany," *Journal of Archaeological Science*, vol. 27, pp. 1187–1195, Dec. 2000.

[143] C. A. Hagen, J. E. Sedinger, and C. E. Braun, "Estimating sex-ratio, survival, and harvest susceptibility in greater sage-grouse: making the most of hunter harvests," *Wildlife Biology*, vol. 2018, Feb. 2018. Publisher: Nordic Board for Wildlife Research.

[144] P. Moehlman and F. Kebede, "IUCN Red List of Threatened Species: Equus africanus," *IUCN Red List of Threatened Species*, Sept. 2014.

[145] R. Adavoudi and M. Pilot, "Consequences of Hybridization in Mammals: A Systematic Review," *Genes*, vol. 13, p. 50, Dec. 2021.

APPENDIX A

Table A.1: Bronze and Iron Age archaeological sites where Equids have been reported among faunal remains in modern-day Turkey. EBA: Early Bronze Age; MBA: Middle Bronze Age; LBA: Late Bronze Age; BA: Bronze Age; IA: Iron Age.

Site	**Period**	**E. hydruntinus**	**E. h. hemionus**	**E. h. onager**	**E. a. asinus**	**E. caballus**	**Equus sp.**	**Reference**
Küllüoba	EBA	**?**					**X**	Efe and Fidan, 2008
Acemhöyük	EBA	**X**			**X**	**X**		Arbuckle, 2013
Büyükkaya (Boğazköy)	EBA				**X**		**X**	von den Driesch and Pöllath, 2004
Kaman Kalehöyük	EBA						**X**	Atıcı, 2003

Tilbeşar Höyük	EBA			**X**			**X**	Berthon and Mashkour, 2008
Karataş Semayük	EBA						**X**	Hesse and Perkins, 1974
Büyüktepe	EBA		**X**		**X**	**X**		Howell-Meurs, 2001b
İkiztepe	EBA						**X**	Yakar, 2007
Titriş Höyük	EBA						**X**	Greenfield, 2002
Gritille	EBA				**X**			Stein, 1987
Sos Höyük	EBA				**X**	**X**		Howell-Meurs, 2001a
Altıntepe	EBA				**X**	**X**	**X**	Satar, 2006
Çadır Höyük	MBA						**X**	Adcock 2020
Giricano	MBA					**X**	**X**	Berthon, 2013
Kenan Tepe	MBA				**X**		**X**	Berthon, 2009
Hibermerdon	MBA				**X**	**X**	**X**	Berthon, 2009

Başur Höyük	MBA					**X**	**X**	Berthon, 2013
Ziyaret Tepe	MBA				**X**	**X**	**X**	Green-field - Jongsma, 2013
Kilisetepe	MBA						**X**	Bursa, 2007
Horum Höyük	MBA						**X**	Bar-tosiewicz, 2005
Ke-sikkaya North-west (Boğazköy)	MBA						**X**	Adcock, 2020
Kültepe	MBA				**X**			Atıcı, 2014b
Demirci Höyük	MBA	**X**						Uerp-mann, 1987
Büyükkaya (Boğazköy)	LBA				**X**	**X**	**X**	von den Driesch and Pöllath, 2004
Boğazköy – Lower Town	LBA			**X**	**X**	**X**	**X**	von den Driesch and Boess-neck, 1981

Hiber-medon	LBA				**X**	**X**	**X**	Laneri et al, 2008
Şarhöyük	LBA				**X**			Satar, 2004
Çadır Höyük	LBA						**X**	Ar-buckle, 2009; Adcock 2020
Kavuşan Höyük	LBA		**X**		**X**	**X**	**X**	Berthon, 2013
Giri-cano	LBA				**X**	**X**		Berthon, 2013
Müslü-mantepe	LBA				**X**		**X**	Berthon, 2013
Kilisetepe	LBA						**X**	Gündem, 2015
Sarıkale (Boğazköy)	LBA				**X**		**X**	Hollstein and Middea 2014
Çatal-höyük	BA	**X**				**X**	**X**	Pawłowska, in press; Russell et al., 2003, 2004; Twiss etal., 2005

Çatal-höyük	IA						**X**	Pawłowska, in press; Russell et al., 2003, 2004; Twiss etal., 2005
Panaztepe	IA						**X**	Tekkaya, 1992
Sos Höyük	IA	**X?**			**X**	**X**		Sagona, 1998
Büyükkaya (Boğazköy)	IA				**X**	**X**	**X**	von den Driesch and Pöllath, 2004
Büyük-tepe	IA		**X**		**X**	**X**		Howell - Meurs, 2001b
Çadır Höyük	IA						**X**	Ar-buckle, 2009; Adcock 2020
Gordion	IA						**X**	Miller, 2009
Kaman Kale-höyük	IA						**X**	Hongo, 1997

Kilisetepe	IA				**X**	**X**		Baker, 2008
Maşat Höyük	IA				**X**	**X**	**X**	Dönmez, 2010
Kavuşan Höyük	IA		**X**		**X**	**X**	**X**	Berthon, 2013
Hiber-medon	IA		**X**	**X**	**X**	**X**	**X**	Laneri, 2008
Büyükardıç	IA					**X**		Açıkkol and Yılmaz, 2005
Altın-tepe	IA				**X**	**X**	**X**	Satar, 2006
Boğazköy	IA				**X**	**X**		Deniz, 2015
Ke-sikkaya South (Boğazköy)	IA						**X**	Adcock 2020
Koru-cutepe					**X**	**X**		Boess-neck and Von Den Driesch, 1974

APPENDIX B

Table B.1: Pre-screening statistics of 15 equid samples. Samples which were sequenced further (see Supplementary Table 2) are indicated in bold. 5' damage and 3' damage indicate the proportion of C->T and G->A mismatches at first and last positions across all reads (Undet: Undetermined)

Sample	Library ID	Raw reads	Equid reads	Equid proportion (%)	Genomic coverage	mtDNA coverage	Average read length (bp)	5' Damage	3' Damage	mtDNA haplogroup	Genomic sex	Osteological material	Zonkey taxon assignment
CDH008	**CDH008_e1b1l1_T**	**36,289,711**	**2,243,532**	**11.35**	**0.1112x**	**2.4976x**	**74.04**	**0.31**	**0.32**	**H1**	**Female**	**Bone (Phalanx)**	**Wild ass**
CDH010	**CDH010_e1b1l1_C**	**15,188,290**	**747,329**	**9.15**	**0.0384x**	**0.8740x**	**75.43**	**0.29**	**0.29**	**H1**	**Female**	**Bone (Calcaneus)**	**Wild ass**
CDH013	CDH013_e1b1l1_T	47,073	27	0.06	0.0000x	0.0000x	76.42	0.29	0.29	Undet.	Undet.	Bone (Phalanx)	Wild ass
CDH022	CDH022_e1b1l1_C	24,997,323	2,342	0.02	0.0001x	0.0121x	71.03	0.13	0.12	Undet.	Undet.	Tooth	Wild ass
CHH001	CHH001_e1b1l1_G	17,409,522	919	0.03	0.0000x	0.0000x	67.15	0.06	0.07	Undet.	Undet.	Tooth	Wild ass
CHH002	CHH002_e1b1l1_A	13,407,962	1,646	0.03	0.0000x	0.0000x	61.76	0.11	0.10	Undet.	Female	Tooth	Wild ass
CHH003	**CHH003_e1b1l1_C**	**33,679,870**	**1,679,031**	**9.12**	**0.0743x**	**5.4064x**	**66.95**	**0.36**	**0.36**	**H1**	**Male**	**Tooth**	**Wild ass**
CHH004	CHH004_e1b1l1_G	15,946,006	1,330	0.03	0.0000x	0.0155x	68.41	0.08	0.07	Undet.	Female	Tooth	Wild ass
CHH005	CHH005_e1b1l1_C	65,462,468	5,433	0.08	0.0001x	0.0124x	66.11	0.04	0.04	Undet.	Female	Tooth	Wild ass
CHH006	CHH006_e1b1l1_A	20,835,763	1,199	0.07	0.0000x	0.0065x	68.18	0.04	0.04	Undet.	Female	Tooth	Wild ass
CHH007	CHH007_e1b1l1_A	33,101,944	4,424	0.08	0.0001x	0.0087x	62.24	0.06	0.06	Undet.	Male	Tooth	Wild ass
CHH008	CHH008_e1b1l1_C	12,666,348	1,329	0.07	0.0000x	0.0000x	59.14	0.05	0.06	Undet.	Female	Tooth	Wild ass
CHH009	CHH009_e1b1l1_A	27,733,217	124,087	0.48	0.0030x	0.1672x	71.36	0.27	0.27	Undet.	Female	Tooth	Wild ass
CHH010	CHH010_e1b1l1_T	14,660,685	1,298	0.08	0.0000x	0.0000x	63.85	0.07	0.06	Undet.	Female	Tooth	Wild ass
CHH011	CHH011_e1b1l1_G	79,752,572	5,537	0.06	0.0001x	0.0138x	65.42	0.04	0.03	Undet.	Female	Tooth	Wild ass

APPENDIX C

Table C.1: Deep sequencing statistics of selected individuals. Merged library (screening + deep sequencing) statistics are given in bold.

Sample	Library ID	Genomic Coverage (x)	mtDNA Coverage (x)	Clonality (%)	5' Damage	3' Damage	Proportion of reads filtered in identity check (%)
CDH008	CDH8_TGGAATA_HN	1.5002	1.5002	11.38	0.32	0.32	0.8600
CDH008	cdh008_TGGAATA_	3.0175	3.0175	15.56	0.32	0.32	0.8574
CDH008	cdh008_b1e1l1ds_L001	1.4737	1.4737	17.90	0.31	0.32	0.8700
CDH008	cdh008_b1e1l1ds_L002	1.2743	1.2743	16.41	0.31	0.32	0.8542
CDH008	**cdh008_merged**	**6.3847**	**6.3847**	**26.97**	**0.32**	**0.32**	**-**
CDH010	CDH10_e1b1l1p1d	0.6891	0.6891	18.76	0.29	0.29	1.1707
CDH010	**cdh010_merged**	**0.7217**	**0.7217**	**19.02**	**0.29**	**0.29**	**-**
CHH003	CHH3_e1b1l1p1ds	0.4963	0.4963	16.57	0.36	0.36	1.4283
CHH003	**chh003_merged**	**0.5667**	**0.5667**	**16.52**	**0.36**	**0.36**	**-**

APPENDIX D

This appendix includes information about the publicly available genomic data that was utilized in several aspects of this thesis, such as dataset construction, mitochondrial analyses and heterozygosity estimation.

This appendix is available online through following link:

https://doi.org/10.5281/zenodo.8220514

APPENDIX E

This appendix includes information about outgroup-f3 statistics calculated on autosomal chromosomes using the main dataset.

This appendix is available online through following link:

https://doi.org/10.5281/zenodo.8220514

APPENDIX F

This appendix includes information about all possible D-statistic comparisons calculated on autosomal chromosomes using the main and the secondary datasets.

This appendix is available online through following link:

https://doi.org/10.5281/zenodo.8220514

CURRICULUM VITAE

Personal Information

Surname, Name: Özkan, Mustafa

Education

Degree	Institution	Year of Graduation
M.Sc.	Ankara University, Ankara, Turkey	2016
B.Sc.	Ankara University, Ankara, Turkey	2012

Grants and Fellowships

2023 Society for Molecular Biology and Evolution meeting, Young Investigator Travel Award, Ferrara, Italy

2023 The Scientific and Technological Research Council of Turkiye, 2218-National Postdoctoral Research Fellowship Program, Archaeogenomic Investigation of Equids of Troy

Publications

Altınışık, N. E., Kazancı, D. D., Aydoğan, A., Gemici, H. C., Erdal, Ö. D., Sarıaltun, S., Vural, K. B., Koptekin, D., Gürün, K., Sağlıcan, E., Fernandes, D., Çakan, G., Koruyucu, M. M., Lagerholm, V. K., Karamurat, C., **Özkan, M.**, Kılınç, G. M., Sevkar, A., Sürer, E., ... Somel, M. (2022). A genomic snapshot of demographic and cultural dynamism in Upper Mesopotamia during the Neolithic Transition. Science Advances, 8(44).

Yurtman, E., Özer, O., Yüncü, E., Dağtaş, N. D., Koptekin, D., Çakan, Y. G., **Özkan, M.**, Akbaba, A., Kaptan, D., Atağ, G., Vural, K. B., Gündem, C. Y., Martin, L., Kılınç, G. M., Ghalichi, A., Açan, S. C., Yaka, R., Sağlıcan, E., Lagerholm, V. K., … Özer, F. (2021). Archaeogenetic analysis of Neolithic sheep from Anatolia suggests a complex demographic history since domestication. Communications Biology, 4(1), Article 1.

Derkus, B., **Özkan, M.**, Emregul, K. C., Emregul, E. (2015). Single frequency analysis for clinical immunosensor design. RSC Advances, 6(1), 281–289.

Presentations

2023 Mustafa Özkan, Kanat Gürün, Eren Yüncü, Kıvılcım Başak Vural, Gözde Atağ, Ali Akbaba, Fatma Rabia Fidan, Ekin Sağlıcan, N. Ezgi Altınışık, Dilek Koptekin, Kamilla Pawłowska, Ian Hodder, Sarah E. Adcock, Benjamin S. Arbuckle, Sharon R. Steadman, Gregory McMahon, Yılmaz Selim Erdal, Cemal Can Bilgin, Eva-Maria Geigl, Thierry Grange, Anders Götherstrom, Füsun Özer, Mehmet Somel. The first complete genome of the European wild ass (*Equus hydruntinus*) reveals its reticulate phylogenetic history, dramatic population decline, and its latest survivors. (Poster) Society for Molecular Biology and Evolution. Ferrara, Italy.

2022 Mustafa Özkan, Kıvılcım Başak Vural, Duygu Deniz Kazancı, Marcella Frangipane, Francesca Balossi Restelli, Giovanni Siracusano, Laszlo Bartosiewicz, Yılmaz Selim Erdal, Elif Genç, Derya Silibolatlaz Baykara, Anders Götherström, Mehmet Somel, Füsun Özer. Spread of Donkeys into Asia Minor: Archaeogenomic Evidence. (Poster) European Society for Evolutionary Biology. Prague, Czechia.

2022 Mustafa Özkan. Archaeogenomic Evaluation of Extinct Wild Asses of Anatolia. (Oral Presentation) 42nd Excavations, Research and Archaeometry Symposium. Denizli, Turkey.

2022 Mustafa Özkan, Kanat Gürün, Eren Yüncü, Kıvılcım Başak Vural, Gözde Atağ, Ali Akbaba, Fatma Rabia Fidan, Ekin Sağlıcan, N. Ezgi Altınışık, Dilek Koptekin, Kamilla Pawłowska, Ian Hodder, Sarah E. Adcock, Benjamin S. Arbuckle, Sharon

R. Steadman, Gregory McMahon, Yılmaz Selim Erdal, Cemal Can Bilgin, Anders Götherstrom, Füsun Özer, Mehmet Somel. Speciation status, complex population history and extinction mechanisms of European Wild Ass. (Oral Presentation) 8th Ecology and Evolutionary Biology Symposium Turkey. Ankara, Turkey.

2019 Mustafa Özkan. Horse Archaeogenomics in Anatolia. (Oral Presentation) 41st Excavations, Research and Archaeometry Symposium. Diyarbakır, Turkey.

2018 Mustafa Özkan, Ali Akbaba, Eren Yüncü, Atilla Engin, Derya Baykara, Mehmet Somel, İnci Togan, Füsun Özer. Preliminary Results on mtDNA Haplogroups of Ancient Goat Samples from Oylum Höyük. (Oral Presentation) International Council of Zooarchaeology meeting. Ankara, Turkey.

2015 Mustafa Özkan, Emel Emregül, Kaan C. Emregül. Developement of Polyethylene Glycol Functionalized Carbon Nanotube Based Anti Myelin Basic Protein Immunosensor. (Poster) BIOMED. Istanbul, Turkey.

2014 Burak Derkuş, **Mustafa Özkan**, Ecem Boztuna, Emel Emregül. Alginate and Alginate - Titanium Dioxide Nanocomposite as Electrode Materials for Anti-Myelin Basic Protein Immunosensing. (Poster) 10th Nanoscience and Nanotechnology Conference. Antalya, Turkey.

Computer Skills

R, Python, Unix environment (bash, awk)

Language Skills

Turkish (Native), English (Advanced)