Drilling fossils and reconstructing genomes: paleogenomics uncovers domestication, dispersal and admixture in *Bos taurus*

Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

Deborah Diquelou

A thesis submitted for the degree of Doctor of Philosophy

2026

Department of Genetics

Trinity College Dublin, the University of Dublin

I declare that this thesis has not been submitted as an exercise for a degree at this or any other university and it is entirely my own work. I confirm that the work has been completed in full compliance with relevant university policies, including the Academic Integrity policy, the Trinity Policy on Good Research Practice, and the guidance on AI and GenAI in Teaching, Learning, Assessment and Research.

I agree to deposit this thesis in the University's open access institutional repository or allow the Library to do so on my behalf, subject to Irish Copyright Legislation and Trinity College Library conditions of use and acknowledgement.

I consent to the examiner retaining a copy of the thesis beyond the examining period, should they so wish (EU GDPR May 2018).

Signed _____ Date _____

Thesis Summary

Paleogenomics allows the study of population structure and evolution. In the case of ancient DNA, it is a marvelous tool to depict the past. In this thesis, *Bos* individuals from Europe, Asia, and Africa are described and used to infer migrations and admixture events with local populations of wild or domesticated cattle.

The first chapter is dedicated to screening results of ancient samples and explores DNA conservation through time and geography. The success or failure of DNA retrieval in 180 samples is analysed to assess DNA extraction protocols. DNA retrieval is then processed in depth in the subsequent chapters, with dedicated population genomics analyses specialised for the study of ancient DNA.

The second chapter focuses on mitochondrial data. New samples were added to an existing database of wild and domesticated samples, leading to 215 individuals included in the chapter and used to build a mitochondrial phylogeny. The main haplogroups found in domesticated cattle, T1, T2, and T3, are identified in most of the samples, but others are present as well, such as Q, E, and K, and even the wild cattle haplogroups R and P. A particular sample is described, coming from Neolithic Saudi Arabia; it represents one of the earliest traces of the T1 haplogroup.

Chapter 3 is the biggest part of the project, with 373 samples gathered for a deep study of Eurasian cattle, using pseudohaploid data to circumvent missing information inherent to ancient DNA. This strategy allows for the projection of samples on a PCA and the estimation of introgression from European and Eastern aurochsen. Population modelling is performed in this chapter in order to investigate the timing of wild Eastern introgression in Europe.

Finally, Chapter 4 focuses on the Middle East, describing newly sequenced data from this difficult environment, characterised by arid conditions known to prevent DNA preservation, resulting in fewer published samples in this region compared to Europe. Thanks to the new samples, and to the successful recovery of DNA, early populations and the timing of indicine introgression in the region are investigated.

Acknowledgements

What an incredible journey this thesis represents. From Brittany, France, to Dublin, Ireland, I have finally achieved my goal, set a very long time ago. It took me to the south of France and back to Brittany before moving to Dublin, and I met many people along the road who encouraged me, from friends to teachers and internship supervisors. After four years, I can finally submit this work and think about the future. My time in Dublin has been full of discoveries, not just scientific, but culturally as well, and I hope I leave with a better English accent than when I arrived. But first things first, I need to thank the people who advised me and helped me to complete my PhD.

This position was offered to me by Professor Daniel Bradley, in the beautiful Smurfit Institute of Trinity College Dublin. His knowledge and experience are tremendously valuable for any student, and I was lucky to be part of his team. I had the chance to have an additional supervisor, Dr. Victoria Elizabeth Mullin, who showed a lot of patience and trained me to be a geneticist. The Bradley group is made up of the nicest people, guided by a talented and attentive lab manager, Dr. Valeria Mattiangeli, who taught us all how to drill our samples and how to perform molecular extractions; nothing could have been done without her guidance. Of course, this project was only possible thanks to the collaboration of many archaeologists who trusted me with precious samples, and this international teamwork has been a pleasure to be a part of.

Present and past PhD students of the lab were great friends, and how could I not mention Áine Halpin in particular, my friend and former housemate. We started at the same time, sharing the same space in our corner of the lab, and you have been my role model for these four years, giving me courage when I was in doubt. From one Celtic woman to another, I wish you the best!

To my family, who endured my obsession with paleogenomics, even though they did not know what it was. To my father who, when asked by a coworker about my studies in Biological Anthropology, answered: “she looks at skin cells”. Though unsure about the specifics of my work, I know my parents have been proud of me, and that my brother, my friend, is impressed by this outcome. To my grandma, Yvonne, you make me laugh so much, and I truly believe you are immortal; I will not accept any other options.

Lastly, to David, what can be said. You have been by my side for six years now; you witnessed my beginnings in Dublin (you even carried my bags there) and ended up following me permanently to this far, far country. I don’t know anyone more skilled than you, and you definitely played a role beyond moral support.

Finally, I want to remember two people who are not here anymore, my grandpa, Fañch, and my aunt, Chantal. I wonder every day what you would be telling me, how you would make me laugh, and tease me (I get it from you!).

To you all,

Bon vent, bonne mer

Kenavo

Table of Contents

1. Introduction to animal domestication and NGS techniques 1
1.1. The Neolithic Revolution 1
1.2. The onset of domestication 1
1.3. Cattle through time 4
1.3.1. The advent of taurine cattle 4
1.3.2. The domestication of indicine cattle: a separate event 5
1.3.3. Geographical distribution of cattle 6
1.4. Steppe migrations and their impact on livestock diffusion 7
1.5. Paleogenomics in the era of NGS 7
1.5.1. Next generation sequencing and ancient genomics 7
1.5.2. Studying cattle domestication via paleogenomics 9
2. DNA preservation and extraction in ancient samples 11
2.1. Introduction 11
2.2. Material and Methods 11
2.2.1. Archaeological samples 11
2.2.2. Drilling strategy 11
2.2.3. DNA extractions and libraries 12
2.2.3.1. DNA extraction 12
2.2.3.2. Double stranded libraries 13
2.2.3.2.1. USER treatment 13
2.2.3.2.2. Blunt-end repair 14
2.2.3.2.3. Adapter ligation 15
2.2.3.2.4. Amplification 16
2.2.4. Bioinformatics pipeline 16
2.2.4.1. Raw files 16
2.2.4.2. Species assignment 17
2.2.4.3. Alignment 18
2.2.4.4. Endogenous content 18
2.2.4.5. Deeper sequencing 18
2.2.4.5.1. Strategy 18
2.2.4.5.2. Bioinformatics extra steps 19
2.2.4.5.3. Molecular sex 20
2.2.4.5.4. Damage patterns 20
2.3. Results 20
2.3.1. Description of the dataset 20
2.3.2. Species identified 23
2.3.3. Endogenous Content 24
2.3.3.1. Per skeletal element 25
2.3.3.2. Across time 27
2.3.4. Evaluation of endogenous content with a new protocol 30
2.3.4.1. Rescuing samples 30
2.3.4.2. Comparisons in new samples 32
2.3.5. Deeper sequencing 34
2.4. Discussion 36
2.4.1. A broad origin of samples 36
2.4.1.1. From a geographic point of view 36
2.4.1.2. From a wide range of periods 37
2.4.2. Endogenous content through space and time 38
2.4.2.1. Bone drilling and DNA preservation 38
2.4.2.2. DNA preservation through time 40
2.4.2.3. Inequalities in DNA content between regions 41

2.4.3. Improvement of the ancient cattle database for future analyses 41
2.4.3.1. Testing the bleach protocol 41
2.4.3.2. Final coverage 42
3. Mitochondrial DNA diversity in ancient cattle 51
3.1. Introduction 51
3.2. Material and Methods 52
3.2.1. Archaeological samples 52
3.2.2. Extraction of the mitochondrial genome from whole genome sequencing 53
3.2.2.1. Raw data 53
3.2.2.2. Alignment and read groups 53
3.2.2.3. Merging several PCRs and quality filters 54
3.2.2.4. Statistics 54
3.2.3. Final mitochondrial genome 54
3.2.3.1. Fasta conversion 54
3.2.3.2. Indels correction and circularisation 55
3.2.4. Mitochondrial Capture 55
3.2.5. Phylogeny and haplogroup 55
3.2.5.1. Maximum likelihood tree 55
3.2.5.2. BEAST 56
3.3. Results 56
3.3.1. Mitochondrial pipeline output 56
3.3.2. Mitochondrial phylogeny 57
3.3.2.1. Maximum likelihood tree 57
3.3.2.2. Haplogroup distribution through time 62
3.3.3. A rare sample from Saudi Arabia 65
3.3.3.1. Capture output 65
3.3.3.2. BEAST phylogeny 65
3.4. Discussion 69
3.4.1. Mitochondrial DNA is more easily extracted from ancient samples than nuclear data 69
3.4.2. Haplogroups became structured over time 69
3.4.2.1. A reduced haplogroup diversity in younger samples 69
3.4.2.2. Specific haplogroups became representative of regions 70
4. A window into the structure and history of past populations of domesticated cattle 77
4.1. Introduction 77
4.2. Material and Methods 79
4.2.1. Samples 79
4.2.2. Projection PCA 79
4.2.3. D-statistics 80
4.2.4. Pseudohaploid dataset 80
4.2.4.1. Admixture 81
4.2.4.2. Admixtools2 81
4.2.4.2.1. F4 ratio 81
4.2.4.2.2. qpAdm 82
4.3. Results 82
4.3.1. Principal Component Analysis (PCA) 82
4.3.1.1. Projection on a modern reference dataset of both taurine and indicine cattle 83
4.3.1.2. Projection on a taurine only reference dataset 86
4.3.2. D-statistics 90
4.3.2.1. Indicine introgression in taurine samples 90
4.3.2.2. Hybridization with wild *Bos* 92

4.3.2.3. Autosomes vs. X 94
4.3.3. Ancestries assessment with Admixture 98
4.3.4. Eastern aurochs proportions estimation via f4-ratios 103
4.3.5. Ancestry modelling with qpAdm 107
4.4. Discussion 110
4.4.1. A variety of ancestries in ancient cattle 110
4.4.1.1. The origin of domestication in the Near-East 110
4.4.1.2. A European domesticates component 111
4.4.1.3. Introgression from local aurochsen in Europe 111
4.4.1.4. Eastern aurochsen ancestry can also be found in domesticates of Eurasia 112
4.4.1.5. The indicine arrival in South-Western Asia 121
4.4.2. African cattle ancestry 121
5. Ancient cattle populations in the Middle East 123
5.1. Introduction 123
5.2. Material and methods 123
5.2.1. Samples 123
5.2.1.1. The Middle East and Central Asia region 123
5.2.1.2. Early populations 124
5.2.2. Projection PCA 126
5.2.3. D-statistics 126
5.2.4. Admixture 127
5.2.5. qpWave 127
5.3. Results 128
5.3.1. Principal Component Analysis (PCA) 128
5.3.2. D-Statistics for indicine introgression assessment 130
5.3.3. Admixture analysis in South-Western Asia 133
5.3.4. Early Middle East structure estimated via qpWave 135
5.4. Discussion 141
5.4.1. Early indicine introgression in South-Western Asia 141
5.4.2. Aurochs introgression in early domesticates 144
5.4.3. T1 mtDNA haplogroup distribution and early introgression in the Middle East145
6. Conclusions 148
6.1. Ancient DNA and paleogenomics 148
6.2. Screening and deeper sequencing 148
6.2.1. New samples 148
6.2.2. DNA preservation 149
6.2.3. Future work 149
6.3. The mitochondrial haplogroup distribution in ancient cattle populations 149
6.3.1. Haplogroup distribution 149
6.3.2. Hybridization with aurochsen 150
6.3.3. New data on haplogroup T1 150
6.4. Admixture events highlighted by whole-genome data 151
6.4.1. Population structure 151
6.4.2. Local introgression 151
6.4.3. Eastern migrations: the Yamnaya 151
6.5. Early populations in the Middle East 152
6.5.1. The timing of indicine introgression 152
6.5.2. Early populations of the Middle East 152
6.6. Conclusion and future directions 153
7. Appendix 155
8. Bibliography 210

1. Introduction to animal domestication and NGS techniques

1.1. The Neolithic Revolution

The Neolithic was one of the most impactful periods in the history of humans, as the emergence of plant and animal domestication would forever change human society. With arable farming and animal herding, people became sedentary. This shift from nomadic communities of hunter-gatherers to sedentary farmers and pastoralists is known as the Neolithic Revolution, a term that reflects the importance of these changes.

Many theories have been suggested to explain why communities decided that a life of farming was more attractive than a life of hunting and gathering, as reviewed by Jacob L. Weisdorf (2005). Among popular theories, the end of the late Glacial period could have provided a warmer environment allowing farming or a food shortage or a growth in population could also have pressured people into finding new food sources, at a time when large mammals such as mammoths went extinct, depriving hunters from a good food resource. Another explanation is simply random discovery that seeds can germinate; the appearance of farming would have been an accident. It is possible that, as highlighted by Bar-Yosef and Belfer-Cohen (1989), hunter-gatherers first adopted sedentarism as a response to climatic changes which narrowed their territories leading to overexploitation of resources. In a smaller territory with less hunting resources, cultivation became a beneficial strategy, and would later provide fodder for domesticated animals.

1.2. The onset of domestication

Sedentarism and cereal farming are a characteristic of the Pre-Pottery Neolithic A (PPNA) archaeological period (12,000 to 10,800 years ago). Thanks to this new way of life, food became more abundant, and populations expanded. However, farming only yields food at specific times in a year. To ensure its more constant availability a way to control

animals had to be developed. This led to animal domestication during the Pre-Pottery Neolithic B (PPNB) (10,800 to 8,500 years ago) (Bar-Yosef, 1998; Hole, 1984).

One exception to this PPNB and later timing of animal domestication is that of dogs, who were domesticated before the Neolithic by hunter-gatherers, at least 12,000 years ago, perhaps even maybe 35,000 years ago (Thalmann and Perri, 2019). An important difference is that dogs were not domesticated in an attempt to develop a new food source, as opposed to other species, but rather as a help for hunting and camp protection. The domestication of species for food or as a way to obtain specific products is then a characteristic of the PPNB.

Domestication enables selection, and thus genomic alteration that will eventually give rise to new phenotypes. This process is unique to humans, and its impact on the environment and on our own history is heavily studied (Vigne, 2015). We know that it started around 10.5k years ago in the Fertile Crescent, a region of southwest Asia, and focused on four different taxa: cattle, sheep, goat and pigs (Hongo et al., 2009). The wild progenitors of these domesticated species were abundant in the region, and were progressively tamed by early pastoralists.

Traces of goats (*Capra hircus*) are found in South Anatolia and the Zagros Mountains, where bezoar (*Capra aegagrus*) were domesticated. It has been proposed that this species had two centres for its domestication, although Anatolia seemed to be the main source of modern mtDNA haplogroups, with the Zagros mountains potentially providing less genetic diversity, as argued by Naderi et al. (2008). Ancient whole genome data has demonstrated a multiple origin for domesticated goats in the Fertile Crescent, but contradicts the idea that Anatolia was the main origin of domesticated goats. We consider the Zagros mountains in Iran as the core centre for goat domestication, but the process would have involved several communities of herders interacting with each other, leading to the emergence of domesticated goats (Daly et al., 2018 and 2021). Similar to goats, sheep (*Ovis aries*) are small ruminants, easily controlled by man, and can provide resources during their life (wool/hair and milk). Sheep were also domesticated in the Fertile Crescent, within the region spanning from Anatolia to the Zagros mountains

approximately 10.5 kya (Zeder, 2008), before their introduction into Europe and Central Asia 6 kya (Taylor et al., 2021).

The Fertile Crescent also saw the domestication of pigs (*Sus domesticus*) around the same time as goats and sheep. Their domestication was more gradual, involving several admixture events from wild boars (Groenen, 2016), and a second separate event occurred in China 8 kya. Wild boar (*Sus scrofa*) were widespread in several continents, and were likely attracted to settlements for food access. This and the widespread distribution of wild boar through Eurasia allowed for admixture with domesticated animals. In the course of Neolithic migrations, pigs reached Europe and in an extraordinary process of wild introgression, European boar genetics overwhelmed the initial pioneer Eastern genomes over the course of three millennia (Larson et al., 2007; Frantz et al., 2019).

Later human mobility was facilitated by the domestication of horses (*Equus caballus*). They provided traction and locomotion in addition to being a meat supplier. Archeological sites containing horse remains point to a different centre of domestication than the Fertile Crescent species. Wild horses (*Equus ferus*) were domesticated by the Botai culture, in the Central Asian steppes 5.5 kya, for riding and milking (Outram et al., 2009). Surprisingly, this initial domesticated population seems to be the ancestor of modern Przewalski horses, a central Asian group supposed to be the last wild horses, but have no ancestry relating to modern domesticated horses (Gaunitz et al., 2018). Other populations of horses existed at the time of domestication (Iberian and Siberian lineages), but did not contribute significantly to the ancestry of the DOM2 lineage, which gave rise to modern horses (Fages et al., 2019). This population seems to have appeared in the Pontic-Caspian Steppe, where oldest remains from this lineage were found, dating from 3,500 to 2,900 BC, and became widespread throughout Eurasia after 2,200 BC, replacing existing populations (Librado et al., 2021).

Domestication took place in several regions, with different goals (such as food and textile production, locomotion, security). However, in each case it is a proof of human abilities to adapt and develop new skills. Cattle prehistory is no exception, and as one of the four major Near Eastern domesticates (with goats, sheep and pigs), it is compelling to study its domestication.

1.3. Cattle through time

1.3.1. The advent of taurine cattle

Modern cattle are found in most regions of the world. Two major groups can be identified: the Asian indicine cattle (also called Zebu, or humped cattle, *Bos indicus*), and Eurasian and African taurine cattle (humpless cattle, *Bos taurus*). As shown by Chen et al. (2018), cattle history is complex. In Asia for example, several migrations of Eurasian *Bos taurus* must have happened, in addition to admixture of several lineages of Indian *Bos indicus*, with wild aurochs (*Bos primigenius*) and banteng (*Bos javanicus*). All these processes make it difficult to unravel the full story behind cattle domestication and dispersal.

However, thanks to zooarchaeology, we can pinpoint the domestication centre of cattle to the northern part of the Fertile Crescent, as described above for the other main domesticates. Hongo et al. (2009) studied *Bos* remains from South-Eastern Turkey dating from the PPNB, while Helmer et al. (2005) described several archaeological sites in the same region, along the Euphrates. These results and other zooarchaeological researches are reviewed by Arbuckle et al. (2016). Indications of cattle management are found in these PPNB sites, including size reduction and kill-off patterns (younger animals) and confirm a domestication of cattle around 10.5kya.

The discovery of ancient DNA methodologies (extraction and amplification of nuclear and/or mitochondrial genomes) gave a valuable means to add detail to archeological findings, and to formulate new hypotheses. In 2001, Troy et al. worked on mitochondrial DNA from 383 *B. taurus* samples, and confirmed a Near-Eastern origin for modern cattle, from a small aurochs matrilineal population. Simulations based on mitochondrial DNA inferred that the original population that gave rise to the first domesticated cattle was very small, featuring only around 80 females aurochs (Bollongino et al., 2012). This was probably due to the aggressiveness of the progenitor, which limited the opportunity for domestication. Probably once more easily manageable animals emerged, there was reduced incentive to add wild animals to the herd, and domesticated

animals were transported outside of the Fertile Crescent. One possible exception is the early arrival of cattle in Cyprus, but this awaits molecular investigation (Vigne et al., 2011).

Once pastoralists had domesticated goats, sheep and pigs, why did they progress to taming a large and dangerous animal such as the aurochs? Why put up the necessary large enclosures to contain a herd? Of course, a cow provides much more meat than a smaller mammal, but also requires more resources to grow. In addition to meat, cattle can provide secondary products during their life, which perhaps justified the effort to raise them. They produce milk, which was processed and consumed as early as 8 kya in Anatolia (Evershed et al., 2008), and 4,8 kya in the Pontic Steppes (Scott et al., 2022). Additionally, the higher strength of these animals was taken advantage of by farmers, who probably used them as help for labour in fields or for transport (Helmer et al., 2018; Pigiere and Smyth, 2023).

1.3.2. The domestication of indicine cattle: a separate event

The benefits of cattle domestication were clearly apparent to past populations of pastoralists, given the investment in *B. taurus* domestication from the wild and dangerous aurochs (*B. primigenius*). Another proof of cattle importance lies in the fact that we can find in Asia another subspecies of cattle: *Bos indicus*. This taxon is widespread in the Indian subcontinent, adapted to its aridity, and was domesticated in the Indus Valley 8 kya (Chen et al., 2010). Given its later date and geographical adjacency it is reasonable to assume that this process was not entirely separate from the *B. taurus* domestication in the Fertile Crescent. However, both types emerged at different times in different regions, from different wild progenitors of *B. primigenius* species (Loftus et al., 1994a; MacHugh et al., 1997). *B. indicus* (zebu or humped cattle) is better suited to arid environments than *B. taurus*, humpless cattle. Traces of hybridization are present in modern cattle genomes and those retrieved from archaeological sites, and indicate admixture of *B. indicus* and *B. taurus* via male zebus, along an axis from East to West, likely as a response to climate change, motivation for improvement of *B. taurus* herds (Bradley et al., 1998; Verdugo et al., 2019).

1.3.3. Geographical distribution of cattle

Analysis of modern breeds of cattle worldwide indicate three main pools: European taurine, African taurine and Asian indicine cattle (Decker et al., 2014). Indeed, when plotted on a PCA, the three of them are placed on extremes of the graph, with hybrid samples intermediate (Verdugo et al., 2019). Mitochondrial studies demonstrated that all taurine cattle originated from the Near East, since the three main haplogroups, T1, T2 and T3, are found in this region (Troy et al., 2001). Nowadays, T1 is mainly found in Africa, T3 in Europe, with most diversity in the Near East (Teasdale and Bradley, 2012).

Although in Europe, taurine cattle are clearly of Near-Eastern origin (Bailey et al., 1997, Bollongino et al., 2006), hybridization with local aurochs has also been detected (Park et al., 2015).

Taurine cattle matrilines in Africa (haplogroup T1) likely also originate in the Near-East given this haplogroup's presence in ancient Levantine samples (Verdugo et al., 2019). As in Europe, African taurine cattle admixed with local aurochs, who may have been adapted to their environment, and this ancestry has also influenced South European cattle (Decker et al., 2014). This may be traced, for example, in several Italian breeds today (McTavish et al., 2013). First depictions of cattle in Africa show humpless animals, and it is only later that indicine cattle were imported, probably as adaptive breeding to the arid environment of Africa. Anatomically indicine cattle in Africa did not originate from this continent, but rather were the result of hybridization of Asian indicine who were brought there either directly or via the Near East (Loftus et al., 1994b; Hanotte et al., 2002).

Asian indicine cattle had extensive influence outside of their original subcontinent. They were domesticated 8 kya, their arrival in Western Asia is detected 4 kya, possibly due to climatic changes that motivated herders to seek better adapted animals (Verdugo et al., 2019). Indicine cattle also contribute ancestry to East Asia. In China, several studies inferred that cattle were of Near-Eastern taurine origin, probably brought around 4 kya (Cai et al., 2014; Xia et al., 2021; Zhang et al., 2021). However, the history of cattle in this region is complex and involved admixtures with several different species of the genus *Bos*, including

taurine and indicine cattle, but also local aurochs, banteng and other unspecified taxa (Chen et al., 2018).

1.4. Steppe migrations and their impact on livestock diffusion

Cattle genomics and their dispersal may have been greatly impacted by the many events of migration in human history. Particularly, the Yamnaya were mobile pastoralists living in the Eurasian steppes, and started a major migration towards the west in the Bronze Age, around 4.5 kya. They arrived in Eastern Europe and admixed with local Europeans resulting in the Corded Ware Culture, *via* which Yamnaya genomic characteristics spread into Western Europe (Haak et al., 2015; Lazaridis et al., 2022; Schmid and Schiffels, 2023). Because the Pontic steppes are close to the geographic point of origin of cattle, the Fertile Crescent, and extend to Eastern Asia where *B. indicus* was located, the Yamnaya also had the possibility to be in contact with people from different part of the Eurasian continent, and trade taurine and/or indicine cattle with them. The timing of Yamnaya migrations correspond to the timing of *B. indicus*' arrival in the Near East (Verdugo et al., 2019), they could have been the instigators of the spread of indicine cattle westwards from the Indian subcontinent. Traces of milk on dental calculus in remains show that the Yamnaya were familiar with domesticated animals, notably cattle and horses, which provided the crucial element for migration across the wide steppes: horse riding and wagon traction by cattle (Wilkin et al., 2021; Scott et al., 2022). It is unknown, however, what impact their migrations had on the livestock populations of Europe.

1.5. Paleogenomics in the era of NGS

1.5.1. Next generation sequencing and ancient genomics

Until recently, scientists had to rely on archaeology alone (or more particularly zooarchaeology in this case) to study the past. With the first DNA extractions in the eighties (Higuchi et al., 1984), a fantastic opportunity appeared for scientists to deepen and test archaeological theories. The field of ancient DNA (aDNA) had appeared, and would lead decades later to major discoveries in the field such as the genomes of *Homo neanderthalensis* and *Homo denisova* (Green et al., 2010; Reich et al., 2010). However, this

field took time to develop. Limited at first by the technology of their time, geneticists started by studying mitochondrial DNA (mtDNA); present in a high number of copies in cells, it is easier to retrieve than nuclear sequences. Regarding cattle domestication, it allowed researchers to compare the main populations of Europe, Africa and Asia (Loftus et al., 1994b), and to compare the two species *B. taurus* and *B. indicus* (Loftus et al., 1994a). In 1997, mtDNA from taurine cattle and aurochs highlighted the relationship between both, aurochs being the wild ancestor of taurine cattle (Bailey et al., 1997). Mitochondrial DNA was used many times in the next decade, to study cattle domestication and links between species (Bradley et al., 1998; Troy et al., 2001; Bollongino et al., 2006).

However, mtDNA limits the hypotheses one can test. Since it is exclusively maternally inherited, we can only infer about the maternal lineage of an individual. A way to study the entire genome had to be developed, and it is in 2005 that the first technology of what will be described as next-generation sequencing (NGS) was published (Margulies et al., 2005). The "NGS revolution" would continue to produce new amplification and sequencing technologies throughout the years, each time decreasing the cost and time of sequencing, and increasing its quality (Besser et al., 2018).

In the world of aDNA, NGS greatly improved results obtained from DNA extractions from fossils, whose preservation will impact the quantity of DNA that has been preserved. It enabled the genomes of Neanderthals and Denisovans to be sequenced. It was also clear even from the first sequenced prehistoric *Homo sapiens* genome that this technology would deeply influence understanding of the human past (Keller et al., 2012). Early domestication studies also benefited from NGS. In just a few years, we managed to elaborate on the processes leading to domestication, including time and place for species such as dogs, horses, goats, sheep, pigs, and of course, cattle (MacHugh et al., 2017; Irving-Pease et al., 2019; Frantz et al., 2020; Orlando et al., 2021). This is still a work in progress, and with technologies being improved constantly, new protocols being developed, the field of aDNA continues to be the source of new discoveries about our past.

1.5.2. Studying cattle domestication via paleogenomics

Animal domestication fascinates, and remains a source of unanswered questions. How past populations managed to tame wild animals, and how they even started to imagine doing this, is incompletely understood. Cattle domestication has been serially studied, and the conclusions are that several events of admixture and migrations helped to shape cattle genomes. Their distribution in the world generated modern patterns of diversity. In order to better understand the spreading of domesticated cattle from South-West Asia, in this project more samples from this area will be sequenced and added to a dataset of ancient cattle genomes from varied locations.

The role of Yamanaya migrations in shaping cattle prehistory is also a main question of my research; whether or not they helped the spread of indicine cattle through Western Asia, and if their migration into Eastern Europe caused an impact on European taurine cattle. This will be elucidated by studying and comparing the population structure of ancient cattle in Eastern Europe and South-Western Asia from different periods.

Displacements of cattle happened in both directions along the Western-Eastern axis in Asia during the early history of regions such as Pakistan. This country, though part of the site of origin of *B. indicus*, is poorly studied using paleogenomics. Several Pakistani samples dating from the Bronze Age are targeted in this work to try to decipher the balance of taurine or indicine cattle at that juncture in this key region.

The lack of aDNA analyses from Pakistan can be partly explained by its harsh environment, which makes it rare to find a well preserved bone; DNA extractions are difficult for archaeological samples found in arid environments. Saudi Arabia provides a similar challenge. Next to the Fertile Crescent, and in between the Eurasian and African continent, the history of domesticates there is interesting, but unknown. It would be logical to find early domesticated cattle, imported from the Levant, but the location of the country is also in the middle of the path indicine cattle would have taken (by land or sea) to enter Africa. It is possible that cattle in Saudi Arabia were purely taurine, or maybe they were hybrids, perhaps even, indicine cattle entirely replaced the population, given their

advantage in arid environments. Ancient Arabian samples were targeted in my dataset, and will hopefully help to characterise cattle from the country.

2. DNA preservation and extraction in ancient samples

2.1. Introduction

One aim of this project was to build on the prior database of ancient cattle samples, in order to more deeply characterise ancient populations in Eurasia, following results described in (Verdugo et al., 2019). Also, following the addition of a bleach wash as a possible improvement in the extraction protocol (Mattiangeli et al., 2023), a few samples which were previously drilled (some of which were published in Verdugo et al., 2019) were extracted again in the hope of improving their endogenous DNA content and thus sequence yield. This chapter will describe the new samples added to the database, those that were rescued by the bleach extraction, and indicate which were selected for specific analyses in the next chapters, as well as the genome coverage achieved for each.

2.2. Material and Methods

2.2.1. Archaeological samples

This project involved the screening of 148 new samples, which were assessed for endogenous DNA, and the re-extraction and/or improvement of 32 samples screened prior to the start of this project. The list of the samples is described in Table 1.

Before drilling, each sample was photographed from various angles with labels provided by the archaeologist. A 3D scanner was acquired a couple months after the start of the project, which allowed the scanning of every subsequent sample processed.

2.2.2. Drilling strategy

For the new bones processed during this project, drilling was performed in a dedicated room following a protocol that depends on the type of bone. Different approaches are considered and consist of adapting the drilling site to the sample, targeting the densest part of the bone in petrous and long bones, while in the case of teeth, it is the root extremity that is used (see Figure 1 below).

Drilling is done under a lateral flow hood which is cleaned in between each sample with absolute ethanol and irradiated with a built-in UV lamp when not in use. Equipment is washed with 5% sodium hypochlorite solution and DNA-ExitusPlus before being irradiated in a crosslinker for 30 minutes before it is used for another sample.

The bone sample is then powderised with a mixer mill, to provide between 120 to 140 mg of bone powder for DNA extraction. Additional powder is frozen for future extractions. The mixer mill was used at a frequency of 30 Hz for 10 to 30 seconds. Smaller times were preferred to avoid overheating the powder and hence DNA destruction.

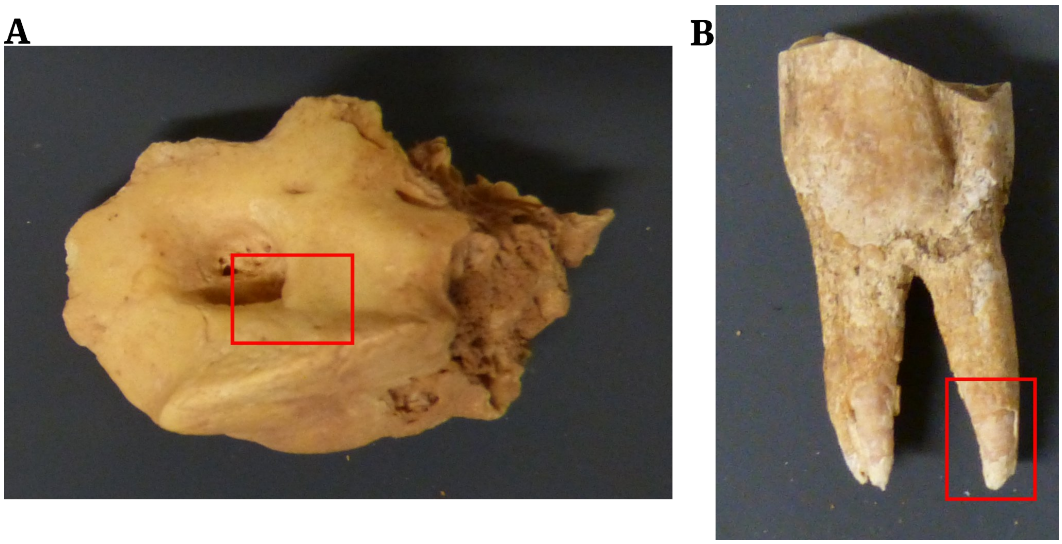

Fig. 1. Drilling sites indicated in red in the case of two different bones. A) In a left petrous, targeting the inner side of the bone surrounding the ear canal; and B) in a tooth, a closed root is selected and cut at the extremity.

2.2.3. DNA extractions and libraries

2.2.3.1. DNA extraction

The protocol for ancient DNA extractions (Mattiangeli et al., 2023) is based on that described in Dabney and Meyer 2019, with modifications (Boessenkool et al., 2017). The modification is a bleach wash before the extraction, aimed at reducing contaminating DNA which is less embedded in the bone matrix. This step was not implemented in previous projects, that is why a few samples were re-extracted and new libraries made.

A 0.5% bleach solution (990uL) was added to the bone powder (with one empty tube acting as an extraction control) and the sample vortexed and then incubated at room temperature for 15 minutes on a rotator. To remove the bleach solution as much as possible, the sample was centrifuged at maximum speed for 10 seconds, and the supernatant removed and replaced by purified water. This step was repeated three times in total.

The bone powder was then treated with 1ml EDTA (0.5M) to dissolve the inorganic bone fraction. Samples were vortexed and incubated at 37°C for 30 minutes on a rotator, centrifuged at max speed for 10 seconds to remove the supernatant which was stored separately (it can be used to extract DNA). An extraction buffer was made with Tris-HCL (1M), N-laurylsarcosine and EDTA, and irradiated for 30 minutes before adding proteinase K. 1ml of the buffer was added to each sample before an overnight incubation at 37°C on a rotator.

The next day, 13 ml of binding buffer (Qiagen) was added to reservoir columns (High Pure Viral Nucleic Acid Large Volume Kit, Roche). Samples were centrifuged for 10 minutes at max speed and the supernatant removed and transferred into the reservoir column containing the binding buffer. The columns were centrifuged in a swing bucket centrifuge for 2 minutes at 700xg, rotated 90° and centrifuged a second time for 1 minute at 700xg. The column was then transferred in a clean collection tube and dry spun for 1 minute at 3,000xg. 750µl of PE buffer (Qiagen) was added to the column which was then centrifuged for 1 minute at 3,000xg and the flow-through discarded. This step was repeated a second time. 25µl of EBT (Elution Buffer + 0.05% Tween, heated at 65°C) was added, and after a 2 minutes incubation at room temperature, the samples were centrifuged for 1 minute at max speed (17,000xg). This step is repeated another time to reach a final DNA extract volume of 50µl.

2.2.3.2. Double stranded libraries

2.2.3.2.1. USER treatment

Double-stranded libraries are made following Meyer and Kircher (2010), after a treatment is performed to reduce the signal from damaged DNA, which accumulates

deamination through time. These damages lead to a change from cytosine to uracil and ultimately typical ancient DNA patterns after sequencing, with misincorporation of thymine and adenine instead of cytosine and guanine respectively. The use of a USER treatment aims at reducing sequencing errors and improving downstream analyses.

To do so, 5µl of USER enzyme (1,000U/ml; Uracil-Specific Excision Reagent, New England BioLabs Inc.) was added to 16.25µl of DNA extract (or 16.25µl of H_2O, representing the library control), and the sample incubated for 1hr at 37°C. The USER enzyme contains both Uracil-DNA glycosylase (UDG) and Endonuclease VIII to perform Uracil excision and open a gap for later repair with the correct nucleotide. Enzymes were removed during the next step of the protocol, through purification via binding and wash buffers.

2.2.3.2.2. Blunt-end repair

UDG-treated samples and a control tube (21,25µl of H_2O) are treated for blunt-end repair using NEB Next End Repair Module (New England BioLabs Inc.). The final volume being 70µl (including the DNA extract), reagents are mixed as follows:

Reagents	Volume (µl)
ddH_2O	38.25
NEBNext End Repair Reaction Buffer	8
NEBNext End Repair Enzyme Mix	3.5

Table 2. Protocol for blunt-end repair mix preparation.

Samples are incubated for 15 minutes at 25°C followed by 5 minutes at 12°C. Purification is done using QIAQuick MinElute purification kit (Qiagen) :

- 350µl of Binding Buffer is added to the samples.
- The total volume is transferred into a QIAquick column which is centrifuged at 3,000xg for 1 minute. Flow-through is discarded.
- 750µl of Wash Buffer is added to the column and another centrifugation at 3,000xg for 1 minute is done. Flow-through is discarded and the samples centrifuged for another minute.

- The column is transferred to a new tube and DNA is eluted in 20µl of EBT, incubated for 2 minutes at room temperature and centrifuged at max speed for 1 minute.

2.2.3.2.3. Adapter ligation

The next step is aimed at attaching ILLUMINA adapters to the DNA and completing any potential gap to make sure the final DNA molecule is double-stranded. First, we use a mix of T4 DNA ligase:

Reagents	Volume (µl)
ddH_2O	10
T4 DNA ligase buffer (10X)	4
PEG-4000 (50%)	4
Adapter mix (50µM each)	1
T4 DNA ligase (5 U/µl)	1

Table 3. Protocol for adapter ligation mix preparation.

20µl of the mix is added to the 20µl of blunt-end repaired DNA. Samples are incubated for 30 minutes at 22°C. The clean-up is performed as previously with the QIAQuick MinElute purification kit, at the difference of 200µl instead of 350µl of Binding Buffer used in the first step (due to differences in final volumes). To ensure a double-stranded molecule is obtained, a mix is prepared as follows:

Reagents	Volume (µl)
ddH_2O	13.5
Thermopol reaction buffer (10X)	4
dNTPs (10mM each)	1
Bst polymerase, large fragment (8 U/µl)	1.5

Table 4. Protocol for adapter fill-in mix preparation.

20µl of the mix is added to the 20µl of DNA and incubation is done for 30 minutes at 37°C followed by 20 minutes at 80°C. Samples are usually immediately amplified or stored at -20°C.

2.2.3.2.4. Amplification

A set of P5 and P7 indexes was chosen for each reaction (in case of several PCRs performed on the same library, the P5 index was reused to be shared by the same sample ; P7 indexes were always unique for each PCRs).

Reagents	Volume (µl)
H_2O	17.5
10X reaction Buffer	2.5
P5 index (10µM)	0.5
Accuprime polymerase	0.5
P7 index (5µM)	1

Table 5. Protocol for amplification reagents volumes.

The mix (22µl) is put into PCR tubes, and 3µl of DNA library added. For screening, 11 cycles of amplification are performed (5 seconds at 95°C, 30 seconds at 60°C and 30 seconds at 68°C) before a final step (5 minutes at 68°C). Samples are purified using the QIAQuick MinElute purification kit as in previous steps, with an elution volume of 10µl. DNA concentration was assessed via the Agilent Tapestation 2200 (Agilent) using a DNA 1000 Kit (Agilent) according to the manufacturer's instructions. Most of the samples were then sequenced at TrinSeq (Trinity Genome Sequencing Laboratory, Trinity College Dublin, Ireland) on a Illumina NovaSeq. A small portion of the samples (39 out of 180) were screened on a MiSeq platform, using a kit v2 (50 cycles= 50 bp single-end) at ELDA Biotech (Dublin, Ireland).

2.2.4. Bioinformatics pipeline

2.2.4.1. Raw files

Reads are received in fastq.gz format and are filtered to keep the correct index combination attributed to the sample during the library amplification. Quality control was done using FastQC (Andrews, 2010).

For paired-end samples, sequenced on the NovaSeq platform), Illumina adapters are removed and read trimmed and collapsed at the same time with AdapterRemoval v2

(Schubert et al., 2016) using parameters to trim consecutive Ns at 5' and 3' extremities, and keep reads that are at least 30 bp long:

```
-AdapterRemoval --file1 file1.fa.gz --file2 file1.fa.gz --mm 3 --
minlength 30 --trimns --trimqualities --collapse --adapter1
AGATCGGAAGAGCACACGTCTGAACTCCAGTCAC --adapter2
AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATCTCGGTGGTCGCCGTATCATT
```

A quarter of the samples were screened via the miSeq platform, which outputs single-end reads. In this case, trimming was done using cutAdapt to remove adapters with an overlap of 1 bp with reads, and keep reads at least 30 bp long:

```
cutadapt -a AGATCGGAAGAGCACACGTCTGAACTCCAGTCAC -O 1 -m 30
```

2.2.4.2. Species assignment

A set of reference genomes was parametered with fastQ Screen (Wingett and Andrews 2018) to test several species that could be mistaken for a cow via morphology identification only:

- Bos taurus = UMD3.1.1
- Canis familiaris = CanFam3.1
- Ovis aries = oviAri3.1
- Capra hircus = ARS1.2
- Sus scrofa = Sscrofa11.1
- Equus caballus = EquCab3.0
- Equus asinus = ASM1607732v2
- Cervus elaphus = mCerEla1.1
- Camelus dromedarius = CamDro3
- Homo sapiens = GRCh38

The software randomly selects 100,000 reads and aligns them using BWA (Li and Durbin 2009) to each reference genome. The output is a table summarising the number of reads that align to only one of the reference genomes at one position only ("one hit, one genome"), to one genome but in multiple locations ("multiple hits, one genome"), to several genomes in one position each time ("one hit, multiple genomes"), to several genomes and in several locations for each ("multiple hits, multiple genomes"), or nowhere

at all (“No hits”). In order to decide which species we are processing, we choose the reference genome that shows the highest number of “one hit, one genome” reads.

2.2.4.3. Alignment

Trimmed and collapsed files were aligned to the reference genome of the species identified with fastQ Screen using BWA v0.7.13 aln (`bwa aln -l 1024 -n 0.01 -o 2`), allowing a 2 gap opening and some mismatches with the reference genome. The sai file obtained from the alignment is converted to sam format using BWA samse which is itself converted to bam file using samtools view (Li et al., 2009). The file is then sorted with samtools sort. Duplicates are removed using MarkDuplicates from Picard v2.20.3 (Broad Institute 2019):

```
picard.jar  MarkDuplicates REMOVE_DUPLICATES=true
```

Reads were finally filtered to remove those with a mapping quality < 25:

```
samtools view -b -q 25 -F 4
```

2.2.4.4. Endogenous content

At each step, a file with basic statistics (number of reads passing each filter) is made with the `samtools flagstat` command. Numbers of reads passing each step are recorded in a spreadsheet, to allow the calculation of the proportions of raw reads aligned and duplicates relative to the total number of raw reads (trimmed and collapsed). Finally, the number of final reads (post-duplicates removal and mapping quality filters) is used to calculate the endogenous DNA via the following:

$$\left(\frac{Total\ number\ of\ final\ reads}{Total\ number\ of\ raw\ reads}\right) * 100$$

2.2.4.5. Deeper sequencing

2.2.4.5.1. Strategy

Samples with both sufficient endogenous DNA (typically > 1%, but a lower threshold of 0.6% was in some cases accepted, due to sample scarcity for a specific site) and high enough read numbers (at least 100,000) were selected for deeper sequencing, as low read

numbers can inflate endogenous content. These samples can be easily identified in Table 1 as they were given a definitive name, instead of the lab name (column “Alias”).

The coverage aimed for the sequencing was estimated based on the endogenous content, from 0.01X for the most poorly preserved samples (< 1% EC) to 1X for the best ones. The average read length obtained during screening, as well as the Bos taurus genome length (2.7 Gb) was used to calculate the number of reads required from the deeper sequencing to achieve the coverage goal, following the following formula:

$$\left(\frac{(genome\ length/avg\ read\ length)}{Endogenous\ proportion}\right) * \text{coverage goal}$$

An additional 10% of reads needed was added to the final number to take into account sequencing variability, and the resulting number used to estimate the sequencing capacity allocated to each sample. Library concentrations obtained from the Tapestation allowed for a conversion from proportion to volumes for the final pooling.

2.2.4.5.2. Bioinformatics extra steps

Selected samples for deeper sequencing went through at least one additional sequencing cycle, with at least one new PCR made, depending on the coverage aimed at. After regular trimming as described above, the alignment step is modified to include read groups when using bwa samse, via the option `-r`. The read groups include the fields “ID” (sample ID plus sequencing platform and run number), “PL” (sequencing company, eg “ILLUMINA”), “LB” (sample ID plus PCR number and indexes) and “SM” (sample ID).

Most of the time, deeper sequencing involves several PCRs per sample, which are merged after sequencing using MergeSamFile from Picard v2.20.3 (Broad Institute 2019), before the duplicates removal step. Indels are then detected and realigned using RealignerTargetCreator and IndelRealigner from GATK (McKenna et al., 2010).

Reads are filtered to keep those with a mapping quality ≥ 25 (`samtools view -b -q 25 -F 4`), and the file is sorted and indexed to assess coverage with Qualimap (Okonechnikov et al., 2016) and its tool `bamqc`.

2.2.4.5.3. Molecular sex

Since the coverage becomes higher than for simple screening, sex was determined using in-house R script, karyotypeR (https://github.com/teasdalm/karyotypeR), which implements the Skoglund score (Skoglund et al., 2013). It allows the identification of specific chromosomes on which reads are mapped, including sexual chromosomes. The software calculates the ratio of reads mapping to the chromosomes X and Y and deduces the sex of the samples, as females should have twice as many reads mapping to the chromosome X than males.

2.2.4.5.4. Damage patterns

To ensure damage patterns are consistent with ancient DNA, PMDTools (Skoglund et al., 2014) was used since it allows for deamination estimation on USER-treated samples.

2.3. Results

2.3.1. Description of the dataset

A total of 180 samples were tested for endogenous DNA (148 new and 32 previously worked bones). They are detailed in Table 1. Many countries from Eurasia were represented, with more samples coming from Iran and Germany, the Caucasus, Anatolia and Eastern Europe (Figure 2). They date from various time periods, from the Paleolithic to the Middle Age, but the Neolithic and the Bronze Age are more highly represented in the dataset, as this time period is of particular interest for the project (Figure 3).

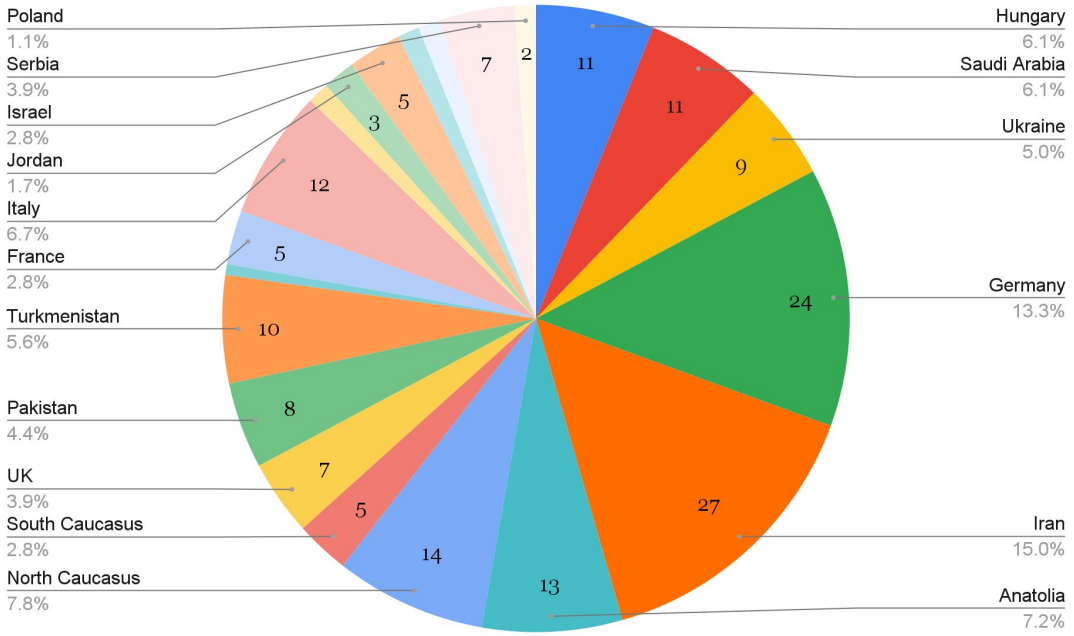

Fig. 2. Geographic origin of the 180 samples screened during the project.

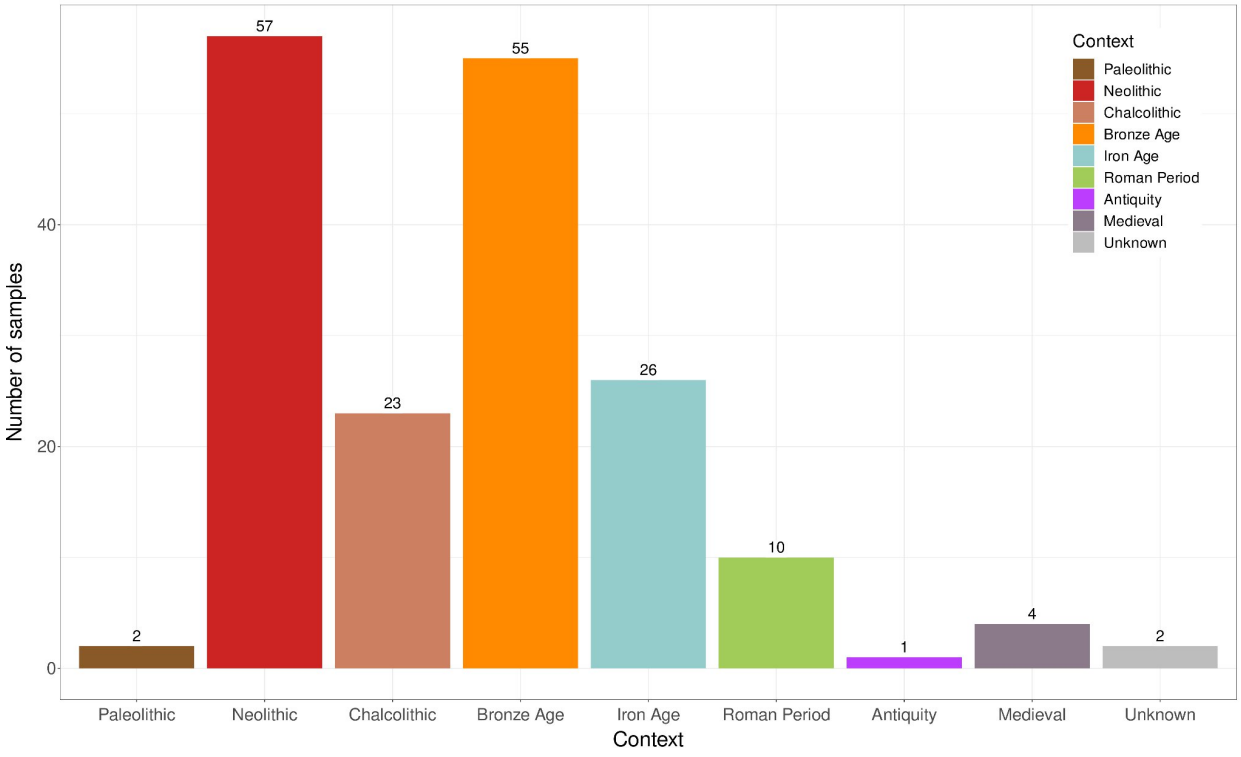

Fig. 3. Time period distribution in the dataset. The two "Unknown" are from Germany, probably Neolithic according to the archaeologist. In the following analyses every sample from the Roman Period, Antiquity and Medieval times were grouped in "Post-Iron Age".

Samples were plotted on a map (Figure 4) below, to give a better representation of the sample distribution according to both age and time.

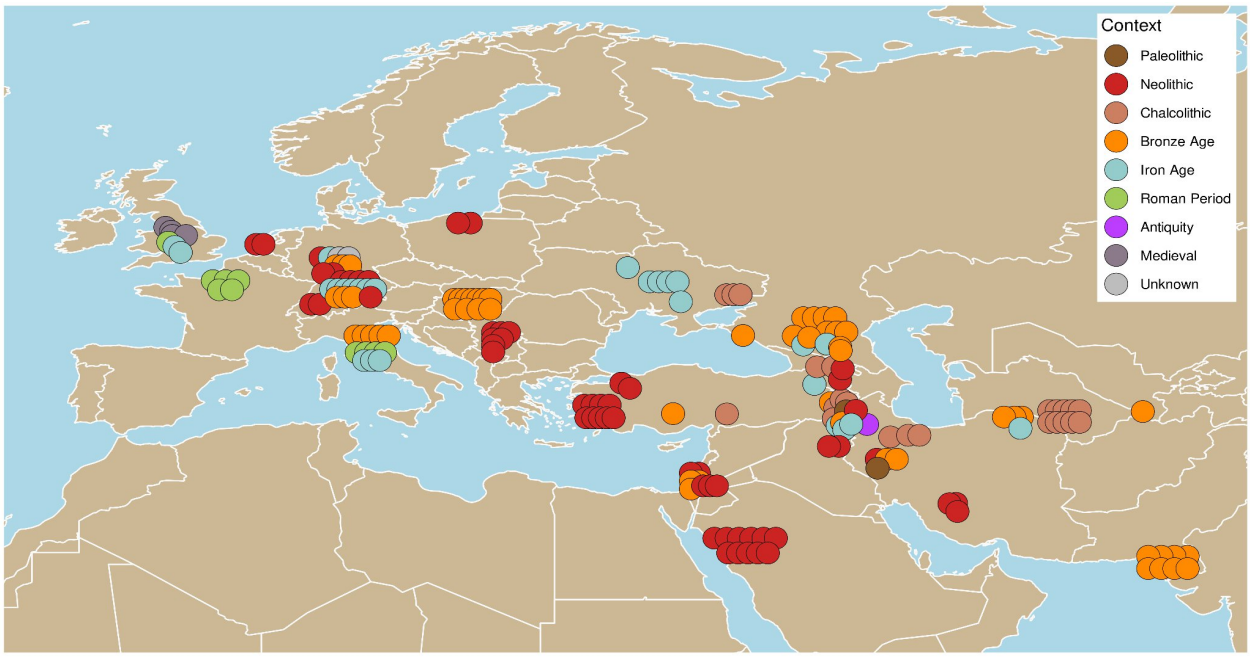

Fig. 4. Geographic distribution of the 180 screened samples. Colours represent the age of the samples. The map was made using Rstudio with jittered coordinates but dots were manually rearranged when too much overlapping was present.

As seen on Figure 4, some countries are better represented in the dataset than others. Germany, Iran and the North Caucasus show the highest representation according to the number of samples and time periods. Some sites have a high number of samples from the same period, such as Szazhalombatta in Hungary (11 Bronze Age samples), or Monjukli Depe in Turkmenistan (10 Chalcolithic samples). 11 Neolithic samples from the city of Al'Ula in Saudi Arabia were attempted for DNA extraction. Despite an unequal distribution, many locations are covered by this new dataset, from North Western Europe to Central Asia, allowing an assessment of DNA preservation in various regions.

Another characteristic of this dataset is the skeletal element chosen for drilling. Indeed, specific bones were targeted (Figure 5). Hence, 105 out of the 180 samples were petrous bones, the second best represented category are the teeth, with 45 samples. The 30 remaining samples come from various parts of the animal's skeleton. 7 of them are not identified further than "long bone", which is a category encompassing femur, tibia or humerus-like bones.

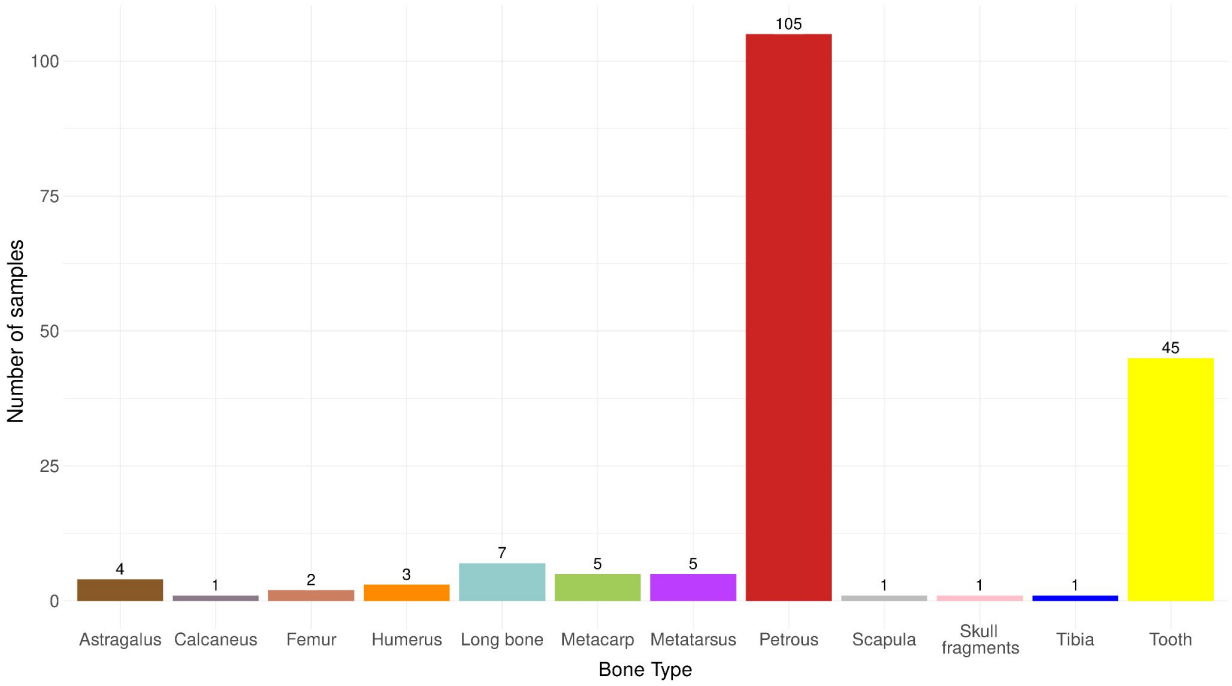

Fig. 5. Number of samples per bone type drilled. Petrous bones and teeth were preferred when available, but other types were tested for DNA extraction.

2.3.2. Species identified

When read numbers were high enough to allow a multiple alignment via fastQ Screen, most of the samples screened were identified as *Bos taurus*. Only 3 samples turned out differently. We discovered one sheep (DD018 - Hasanlu, Iran), one horse (DD095 - Orchosani, South Caucasus) and one donkey (DD123 - Naderi Tepe, Iran). These samples were realigned to their corresponding reference genome to calculate their true endogenous content.

Other samples yield too low DNA content to be able to identify with certainty the species. These samples come from South Western Asian countries such as Saudi Arabia (DD004, DD005, DD006, DD026, DD126, LL040, LL041, LL042), Iran (LL006, LL055, DD089, DD090), Iraq (LL033, LL037), and Georgia (DD094). All these samples were morphologically identified as cattle bones by a zooarcheologist; endogenous content could not confirm this.

2.3.3. Endogenous Content

Extraction yielded different results in terms of DNA content. Most of the dataset (92 samples) falls between 0.000 and 5.000% of endogenous content (Figure 6). 60 samples resulted in less than 1.000% while others went as high as 72.000 %, though this highest result belongs to a sample identified as donkey, the second highest value was from a horse. The highest value for a *Bos taurus* specimen is 62.134%. As seen in Figure 6, a minority but significant number of samples gave between 5.000 and 60.000 % of endogenous DNA.

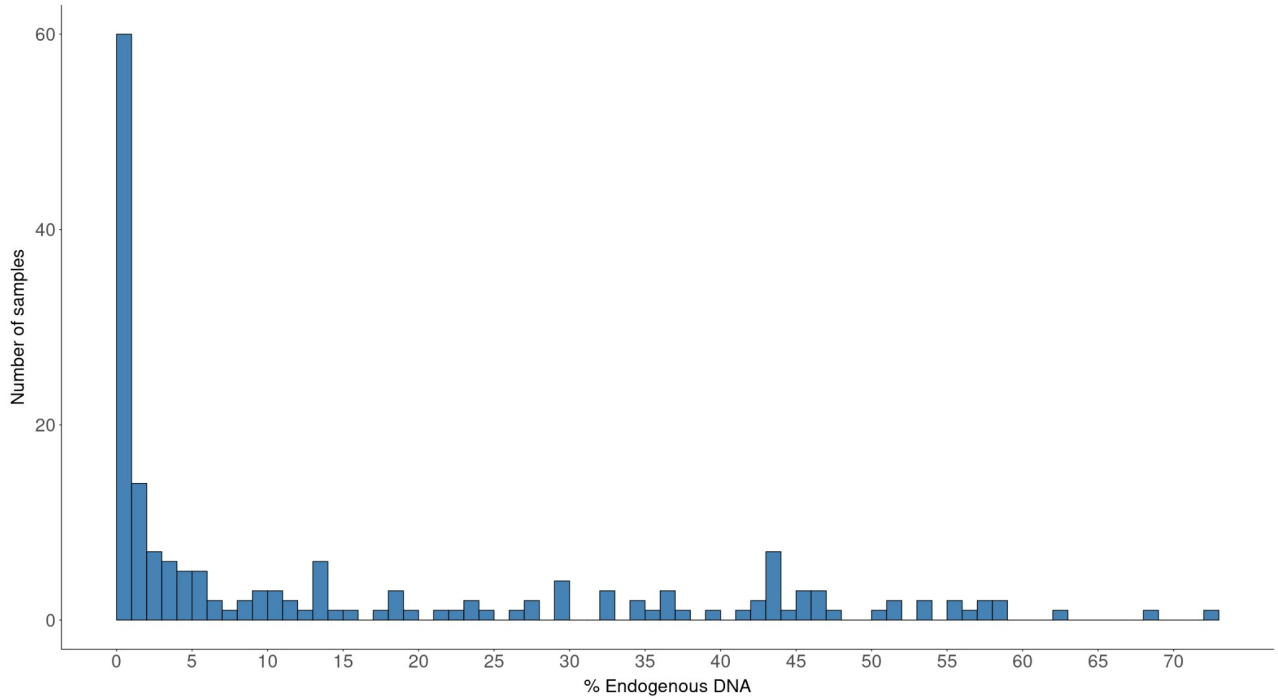

Fig. 6. Histogram of endogenous content distribution throughout the dataset. Endogenous content is represented in 1% intervals on the horizontal axis and the number of samples counted in each interval is represented by the vertical axis.

Endogenous content is represented on the map of Figure 7, ranging from low amounts of DNA in red to medium in blue and high in green. European samples most of the time yield a high endogenous content of DNA (21 samples with EC > 10%) or at least a medium value (n=15, EC ≥ 1%). Only 9 European samples contained less than 1%. In Anatolia, ancient samples have a medium (n=8) or high (n=4) EC but one sample had a poor result. In the Caucasus, more low EC are observed (8 samples), with 7 medium and only 2 samples presenting a high EC. Similarly, the Levant has 3 samples with a medium EC but 5

with a low EC. Every sample from Saudi Arabia showed less than 1% EC, as did the two samples from Iraq. Iran has either very bad DNA preservation (6 samples) or very good (14 samples). Turkmenistan is represented by 3 medium samples and 7 high EC samples, while the only screened sample from Uzbekistan had high preservation. Finally, 2 samples from Pakistan were medium in DNA content and 6 were low.

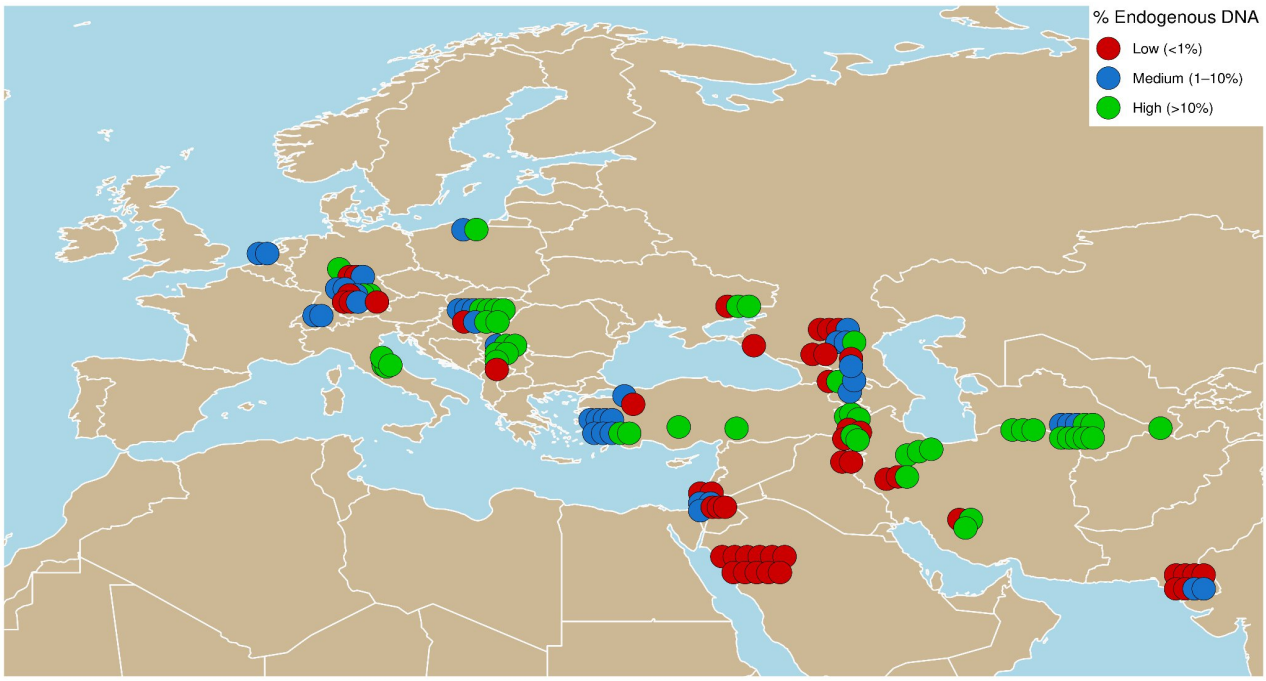

Fig. 7. Distribution of endogenous content from the Neolithic to the Bronze Age in Eurasia. Samples are colored according to their DNA content, low in red (EC < 1%), medium in blue (1% ≤ EC ≤ 10%) and high in green (> 10%).

2.3.3.1. Per skeletal element

Endogenous contents displayed in Figure 6 were compared with the type of bone from which DNA was extracted (Figure 8). The best results are found in petrous bones and teeth, but boxplots show that despite the high values obtained from a subset of these bones, many of them gave poor DNA preservation, especially for teeth for which the median is very low (1.086%, against 17.245% for petrous bones). Similarly, long bones can result in both bad or good values, but the median is low, 3.142%. This value can be improved if we merge the values of long bones with those of the tibia, femur and humerus; endogenous content then has a median of 9.147%. Other bones were screened in small numbers that prevented median assessments and comparisons with the more studied

bones, with the exception of the metatarsus (4.196% of median endogenous DNA) which included a single exceptional sample with 42.036% endogenous DNA.

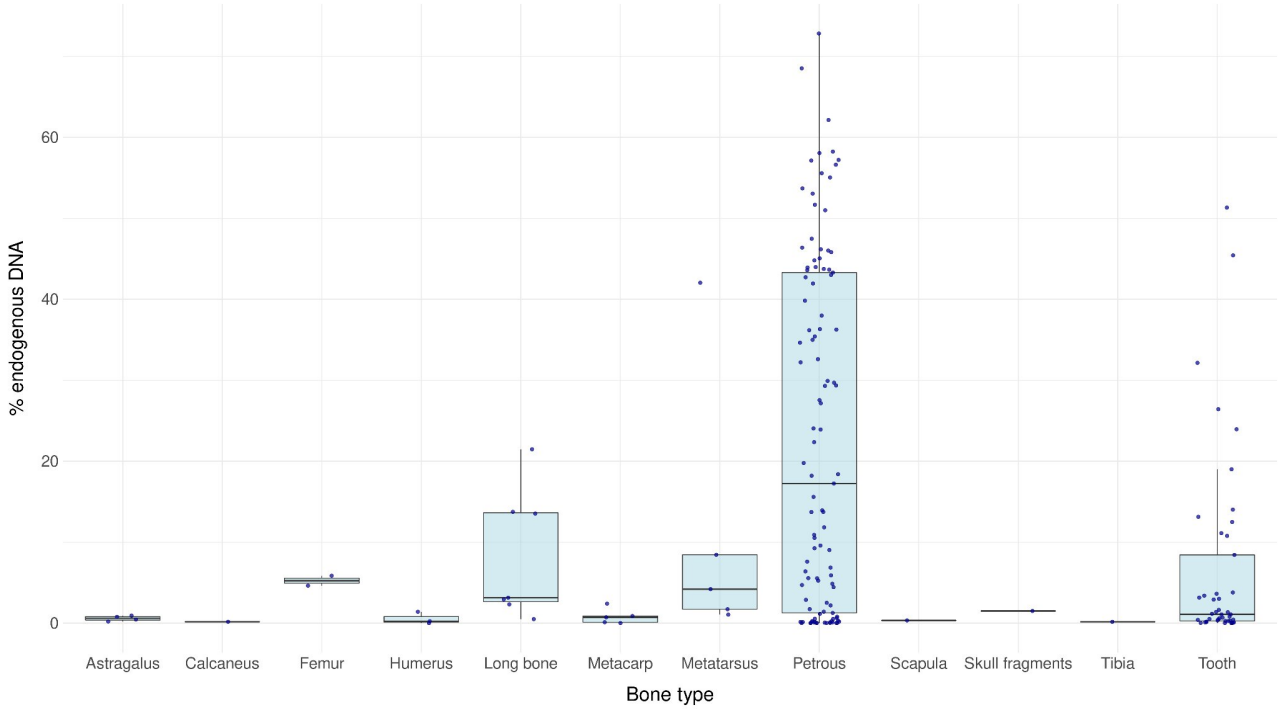

Fig. 8. Boxplots of endogenous content according **to the type of bone.** All 180 samples are represented as described in Table 1.

Teeth available for drilling came in various preservation states. To assess the association of perceived physical preservation with endogenous content, teeth were divided into two groups, a “good” one and a “bad” one in Figure 9 below. This partition was decided according to visual factors. Broken teeth, opened roots, or any other kind of damage led them to be classified a “bad”. A “good” tooth is an intact one, with the entire roots present.

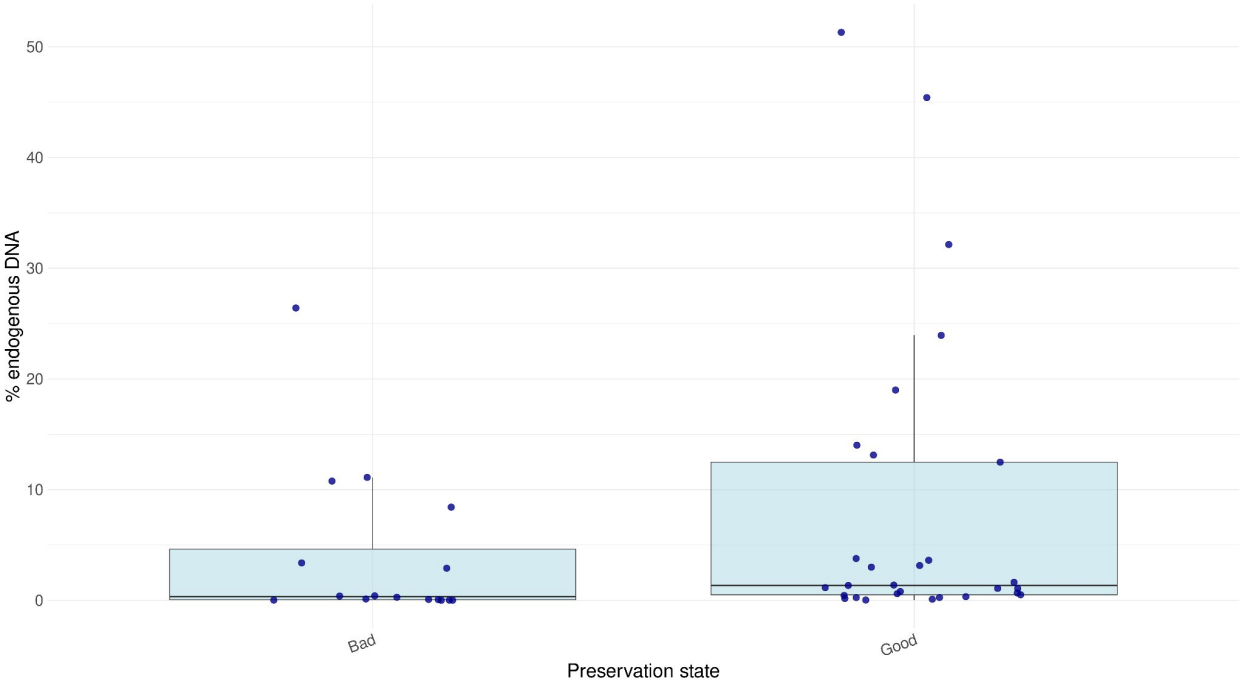

Fig. 9. Boxplots of endogenous content in teeth according to their preservation. A "good" tooth is intact and complete, a "bad" tooth is broken, has open roots or suddenly broke while drilling due to bad density.

"Bad" teeth have a maximum of 26.415% of Endogenous content (DD012, Chalcolithic Ukraine), and a median of 0.344% EC. DD012 is an outlier, the second best sample, DD034 (Iron Age, Caucasus), only reaches 11.112% EC. In the "good" group, EC can go as high as 51.314%, as seen in DD127 (Germany, probably Neolithic). The second best value is close, with DD010-1 (Iron Age Ukraine) showing 45.418% EC. This group has a median at 1.351% of endogenous content. These groups are statistically different, a Wilcoxon test finds a p-value of 0.042.

2.3.3.2. Across time

Another factor related to DNA preservation is the archaeological period the sample is dated from. In Figure 10, a scatterplot represents endogenous content through time, with older samples on the left and more recent ones on the right (with the exception of 2 "Unknown", which were not assigned a definitive time period, but were probably late Neolithic). Colors represent the time period, from Paleolithic to Medieval times.

As seen earlier, many samples have low values for endogenous content, close to 0%. As time progresses towards the present, the number of samples resulting in higher values improves. The DNA preserved individual samples are found in the Chalcolithic, Bronze Age and Iron Ages, but more consistent clusters of samples located at the top of the plot are found in more recent periods: Iron Age, Roman Period and Medieval times.

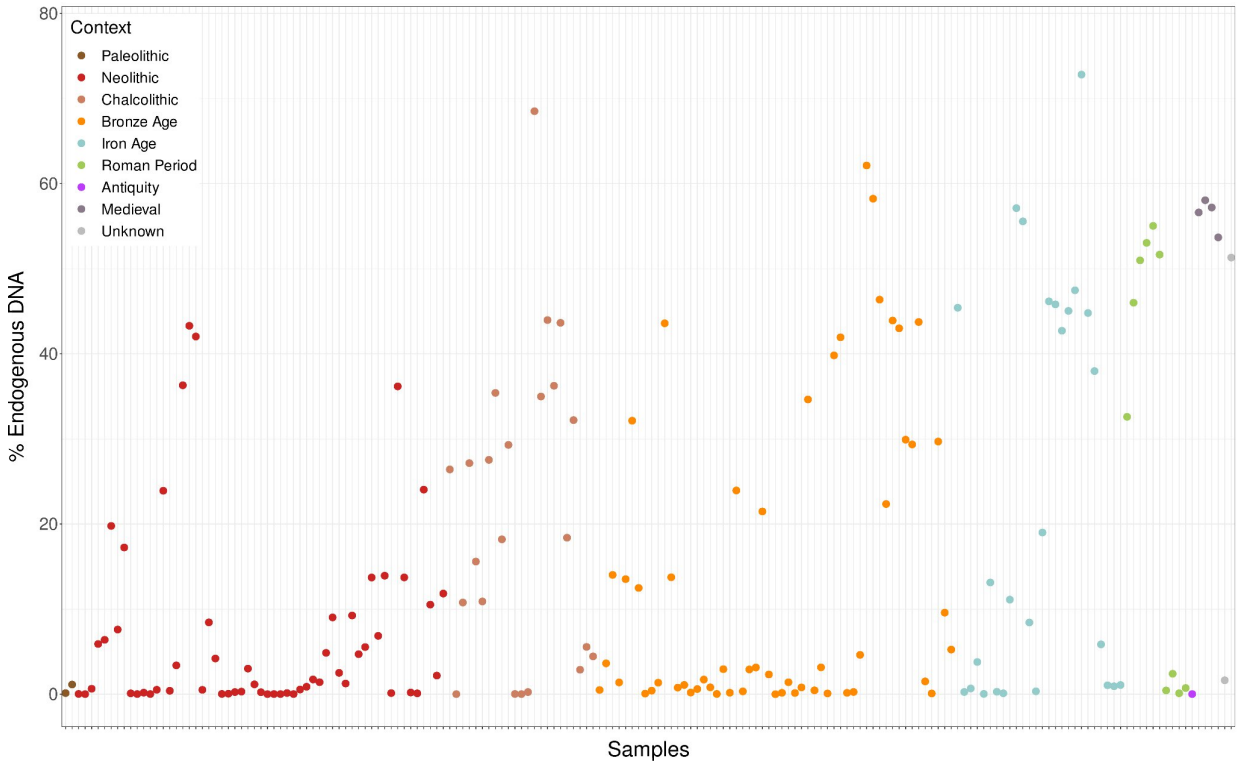

Fig. 10. Variation of endogenous content through time periods. The plot represents each 180 screened samples colored according to their time period. $R^2 = 0.153$.

The correlation between time period and endogenous content was investigated in Figure 11. Samples are divided in three time periods, early, middle and late, and regression lines and coefficients of determination R^2 are calculated for each. The three lines indicate positive regressions but with different strength. Early samples have a coefficient $R^2 = 0.093$, middle samples have $R^2 = 0.053$ and later samples have $R^2 = 0.568$. This indicates that the strongest trend is a substantial improvement in recent centuries.

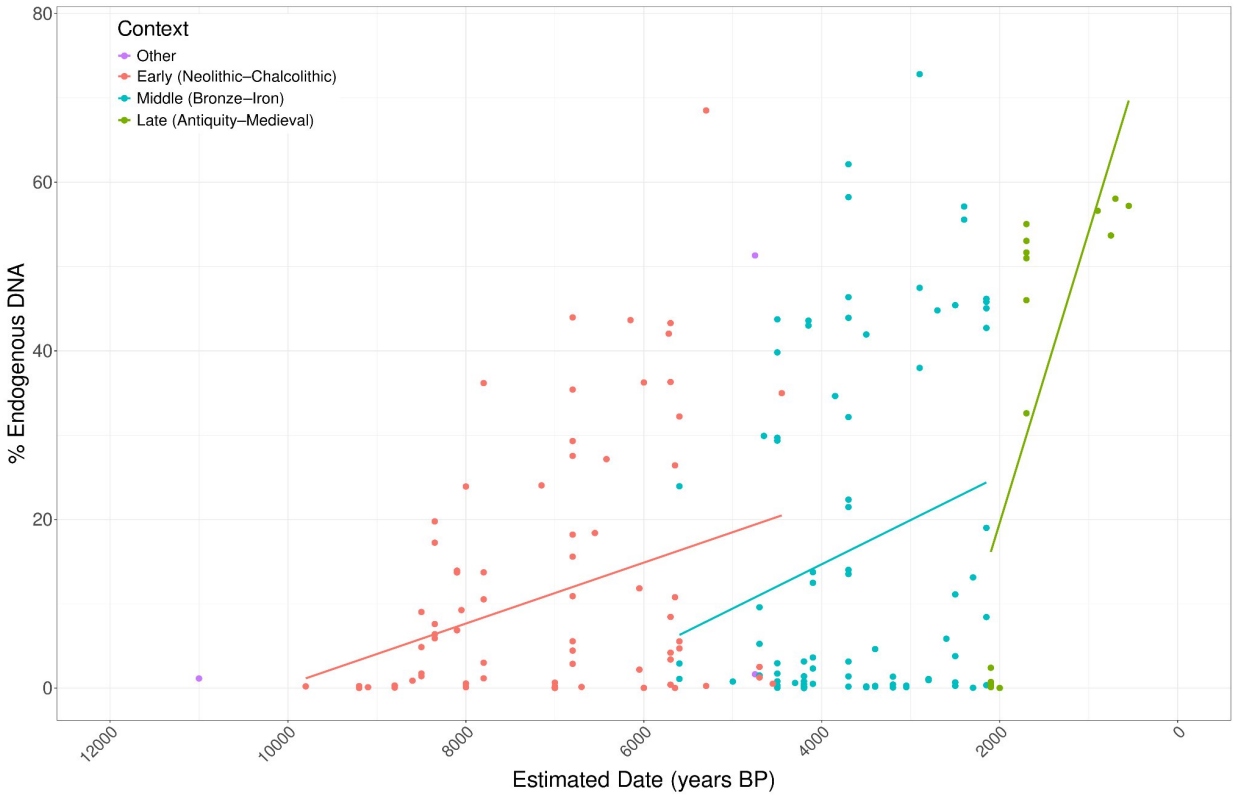

Fig. 11. Variation of endogenous content through time. Samples are plotted according to their estimated age (BP), 3 regression lines are plotted for 3 groups of periods: early (Neolithic + Chalcolithic, R^2 = 0.093), middle (Bronze and Iron Age, R^2 = 0.053) and late (Antiquity, Roman Period and Medieval times, R^2 = 0.568).

The evolution of endogenous content through time observed in Figures 10 and 11 above can also be seen in averages and medians calculated in Table 6 for each time period. Paleolithic samples were excluded due to their small number (2), as well as the two German samples that are for now listed with an "Unknown" time period. The unique sample from Antiquity has been included in the Roman Period group. This leads to a similar trend as before, with smaller average and median values in the Neolithic compared to later periods. The Chalcolithic does not fit into this progression however, as it shows the highest values for average and median DNA content before the Middle Ages. DNA is best preserved in the more recent samples from the Medieval times, when endogenous content does not fall below 32.597%. In between the Chalcolithic and the Middle Ages, the Bronze Age and Iron Age show gradual improvement of DNA preservation compared to the Neolithic, before a small relapse during the Antiquity - Roman Period.

Table 6. Evolution of endogenous content through time (%). Average and median values of endogenous content were calculated at each period (except for the Paleolithic). For each time period, the number n of samples included in the calculations is indicated.

	Neolithic (n=60)	Chalcolithic (n=23)	Bronze Age (n=55)	Iron Age (n=26)	Antiquity and Roman Period (n=11)	Medieval times (n=4)
Min	0.011	0.011	0.000	0.035	0.015	53.684
Max	51.314	68.516	62.134	72.816	55.037	58.049
Average	7.664	21.409	13.658	23.345	26.639	56.387
Median	1.732	18.393	2.944	12.121	32.597	56.907

2.3.4. Evaluation of endogenous content with a new protocol

2.3.4.1. Rescuing samples

23 samples were drilled prior to the beginning of this project, and yielded estimates for endogenous DNA content. Some were published (see Table 7), others were discarded due to poor DNA preservation. After their original screening, a new bleach step was added to the extraction protocol, aimed at reducing contamination and thus improving endogenous content. This new protocol was tested with these 23 samples and the new endogenous content value is compared to the previous one (non-bleach) in Table 7.

Table 7. Comparisons of endogenous content after a bleach or a non bleach extraction protocol. Samples with an alias followed by an * (n=5) were published in Verdugo et al. (2019).

Sample	Alias	Country	Site	Period	% endogenous DNA	Non-bleach %	Improvement?
KD159	Tqa3*	Israel	Tel es-Qashish	Bronze Age	1.501	4.170	No
MV015		Anatolia	Barcin Hoyuk	Neolithic	0.875	1.530	No
MV020		Anatolia	Çukuriçi Höyük	Neolithic	1.732	1.680	Yes
MV021		Anatolia	Çukuriçi Höyük	Neolithic	1.410	2.540	No
MV023	Cuk1	Anatolia	Çukuriçi Höyük	Neolithic	4.860	1.070	Yes
MV024	Cuk2	Anatolia	Çukuriçi Höyük	Neolithic	9.025	1.830	Yes
MV026		Switzerland	Sutz-Lattrigen	Neolithic	2.509	1.880	Yes
MV028		Switzerland	Sutz-Lattrigen	Neolithic	1.256	1.610	No
MV045	Men3	Anatolia	Mentese	Neolithic	9.249	5.570	Yes
MV052		Netherlands	Schipluiden	Neolithic	4.704	4.070	Yes
MV053	Sch3	Netherlands	Schipluiden	Neolithic	5.540	13.850	No
MV072	Bla3	Serbia	Blagotin	Neolithic	6.856	7.210	No
MV076	Bla2*	Serbia	Blagotin	Early Neolithic	13.927	13.590	Yes
MV079		Serbia	Stubline	Neolithic	0.131	4.600	No
MV087	Bub2	Serbia	Bubanj	Neolithic	2.473	10.630	No
MV155	Tqa2*	Israel	Tel es-Qashish	Bronze Age	9.584	3.840	Yes
MV159	Plo6*	Serbia	Plocnik	Neolithic	24.040	4.170	Yes
MV184	Ruc1	Poland	Rzucewo	Neolithic	2.187	2.880	No
MV185	Ruc2	Poland	Rzucewo	Neolithic	11.831	8.000	Yes
MV209	Tes1	Israel	Tel-el-Safi	Early Bronze Age	5.244	4.430	Yes
MV234	Mon1*	Turkmenistan	Monjukli Depe	Chalcolithic	2.875	7.780	No

MV235	Mon2	Turkmenistan	Monjukli Depe	Chalcolithic	5.557	3.240	Yes
MV236	Mon3	Turkmenistan	Monjukli Depe	Chalcolithic	4.442	1.930	Yes

Out of the 23 samples that went for DNA extraction with the new protocol, 13 resulted in a better endogenous content than from their previous non-bleach extraction, but 10 did not. The improvement ranges from an additional 0.052% to 19.870% in endogenous DNA content. Samples that possess an Alias name without any asterisk were further sequenced using the newly made library after bleach extraction (11 samples). No alias name but an improvement in endogenous content indicates that the read number was still too low (which was the case for 3 samples).

2.3.4.2. Comparisons in new samples

Most of the newly screened samples were treated with bleach during DNA extraction. However, when dealing with samples with which we expect difficulty extracting DNA (because they are very old, or from arid environments where DNA degrades faster), it was sometimes decided to use the traditional non-bleach protocol to prevent the harsher chemical treatment destroying endogenous DNA. Several samples from Neolithic Saudi Arabia and Bronze Age Pakistan were hence extracted via both protocols.

Results in terms of final number of reads and endogenous content from both protocols are shown in Table 8. The first samples from Saudi Arabia, DD004, DD005 and DD006, were drilled to provide enough bone powder to test for both protocols, trying to extract a second time from each using the EDTA wash. Because DD006 provided more bone powder, a second non-bleach extraction (plus EDTA wash) was performed. DD046 and DD048 were only processed with the bleach protocol, as it had become the norm, but a test was still performed on DD047. The low count of read number and poor endogenous content found in Saudi Arabian samples suggested the sole use of the non-bleach approach for the last sample screened from this country, DD126. Both protocols were tried for bones from Pakistan, with an additional extraction from the EDTA wash for most of the bleach extractions.

Table 8. Samples from Saudi Arabia and Pakistan and their endogenous content after a bleach or non-bleach extraction. The final number of reads is also indicated to help inform about the endogenous content meaning.

Sample	Protocol	Number of final reads	Endogenous DNA (%)	Country	Geographic Origin	Period
DD004	Bleach	8	**1.372**	Saudi Arabia	Al'Ula	Neolithic
	Bleach	17	**3.953**	Saudi Arabia	Al'Ula	Neolithic
	Non-bleach	29	**0.020**	Saudi Arabia	Al'Ula	Neolithic
	Non-bleach	69	**0.042**	Saudi Arabia	Al'Ula	Neolithic
DD005	Bleach	404	**0.015**	Saudi Arabia	Al'Ula	Neolithic
	Bleach	257	**0.018**	Saudi Arabia	Al'Ula	Neolithic
	Non-bleach	23	**0.006**	Saudi Arabia	Al'Ula	Neolithic
	Non-bleach	813	**0.008**	Saudi Arabia	Al'Ula	Neolithic
DD006	Bleach	66	**2.505**	Saudi Arabia	Al'Ula	Neolithic
	Bleach	2569	**0.747**	Saudi Arabia	Al'Ula	Neolithic
	Non-bleach	301	**0.028**	Saudi Arabia	Al'Ula	Neolithic
	Non-bleach	777	**0.032**	Saudi Arabia	Al'Ula	Neolithic
	Non-bleach	629	**0.210**	Saudi Arabia	Al'Ula	Neolithic
	Non-bleach	2140	**0.318**	Saudi Arabia	Al'Ula	Neolithic
DD046	Bleach	13	**0.020**	Saudi Arabia	Al'Ula	Neolithic
DD047	Bleach	2058	**0.966**	Saudi Arabia	Al'Ula	Neolithic
	Non-bleach	2549	**1.160**	Saudi Arabia	Al'Ula	Neolithic
DD048	Bleach	142	**0.020**	Saudi Arabia	Al'Ula	Neolithic
DD126	Non-bleach	167	**0.030**	Saudi Arabia	Al'Ula	Neolithic
	Non-bleach	225	**0.033**	Saudi Arabia	Al'Ula	Neolithic
DD056	Bleach	989	**1.053**	Pakistan	Chanhu Daro	Bronze Age
	Bleach	9300	**0.315**	Pakistan	Chanhu Daro	Bronze Age
	Bleach	4087	**0.874**	Pakistan	Chanhu Daro	Bronze Age
	Non-bleach	10247	**0.095**	Pakistan	Chanhu Daro	Bronze Age
DD057	Bleach	716	**2.180**	Pakistan	Chanhu Daro	Bronze Age
	Bleach	4115	**1.179**	Pakistan	Chanhu Daro	Bronze Age

	Bleach	6236	**2.298**	Pakistan	Chanhu Daro	Bronze Age
	Bleach	1427	**0.756**	Pakistan	Chanhu Daro	Bronze Age
	Non-bleach	9145	**0.078**	Pakistan	Chanhu Daro	Bronze Age
DD058	Bleach	1748	**3.472**	Pakistan	Chanhu Daro	Bronze Age
	Bleach	11616	**1.653**	Pakistan	Chanhu Daro	Bronze Age
	Bleach	16056	**4.523**	Pakistan	Chanhu Daro	Bronze Age
	Bleach	12519	**5.791**	Pakistan	Chanhu Daro	Bronze Age
	Non-bleach	14766	**0.560**	Pakistan	Chanhu Daro	Bronze Age
DD059	Bleach	2918	**0.269**	Pakistan	Chanhu Daro	Bronze Age
	Bleach	12918	**0.000**	Pakistan	Chanhu Daro	Bronze Age
	Non-bleach	977	**0.014**	Pakistan	Chanhu Daro	Bronze Age
DD060	Bleach	3162	**0.268**	Pakistan	Chanhu Daro	Bronze Age
	Bleach	44221	**1.202**	Pakistan	Chanhu Daro	Bronze Age
	Non-bleach	352	**0.023**	Pakistan	Chanhu Daro	Bronze Age

Overall, samples from Saudi Arabia did not perform well enough to get a satisfactory number of reads (less than 3,000 every time). A slight increase in this number can be seen from the non-bleach compared to the bleach protocol for DD004, DD056 and DD057. The endogenous content is very low in Saudi Arabia, especially considering the small number of reads, which can inflate endogenous value.

It is to note that due to the multiple tests performed on these samples, and the nature of Figures 6, 8, 10 and 11, an average was taken for representation of endogenous content (DD004 and DD006 had very small numbers of reads associated to their highest endogenous content values, these higher than usual values were thus removed from the average calculations).

2.3.5. Deeper sequencing

Following screening, 179 samples (both newly screened or older samples needing coverage improvement) were sent for higher sequencing, aiming to obtain at least 0.01X coverage. In the case of re-extraction of previously screened samples (Table 7), they were

sent for deeper sequencing only if the improvement was sufficient (and associated with a high number of reads) unless the existing content qualified for the deeper sequencing strategy.

Samples that went for additional sequencing are listed in Table 1, along with the coverage achieved. Coverage distribution is plotted in Figure 12. The majority of the samples are located below the 0.1X limit on the left of the figure. A zoom in the 0 to 1X area is shown in Figure 12.B.

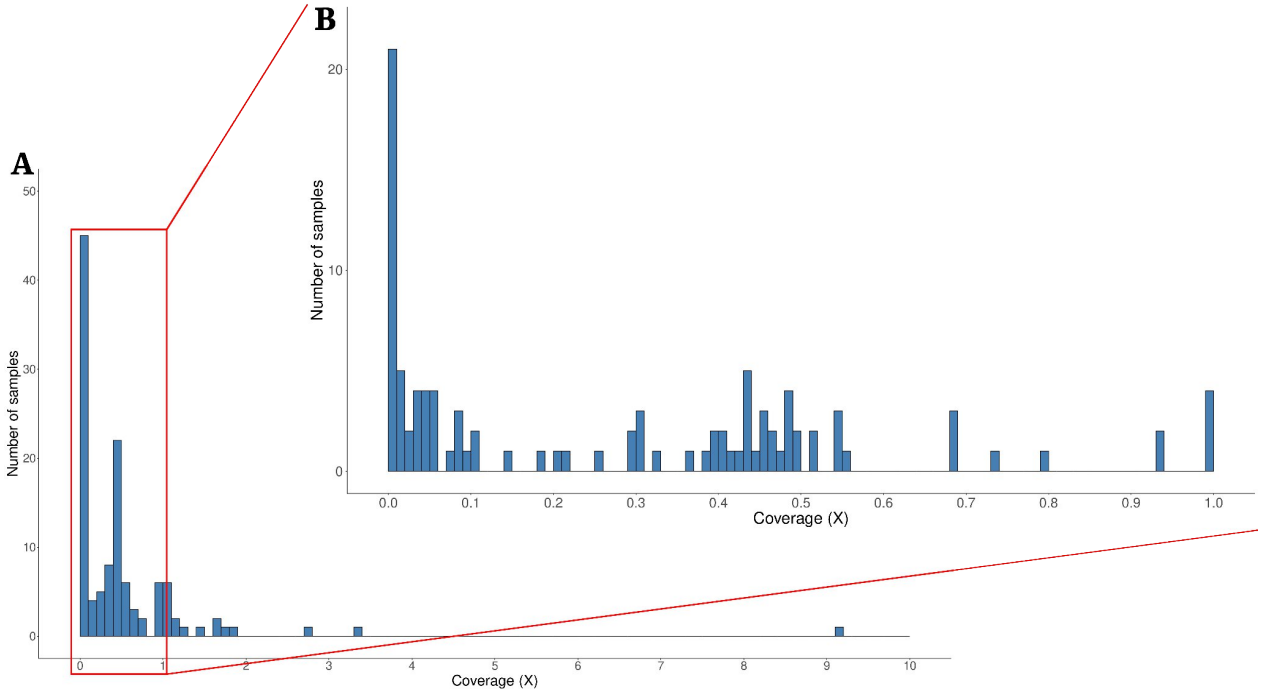

Fig. 12. Coverage distribution in the 179 samples sent for deeper sequencing. A) All samples are represented, with the highest coverage being above 9X. B) A close-up view of coverage distribution below 1X, where most of the samples are located.

10 of the samples on the far left of the figure did not yield sufficient DNA and came back with a 0.00X coverage. 6 other samples did not reach the 0.01X threshold hoped for. As seen in Figure 12.B, two peaks of samples are found for values between 0 and 0.1X (44 samples) and between 0.40 and 0.50X (22 samples). Other samples are distributed in a similar range of numbers between 0.10 and 1X, with a maximum of 8 samples in the same range of coverage. On average, the mean coverage achieved is 0.527X, with a median of

0.380X. The two highest values are found in Medieval England samples, reaching 3.310 and 9.140X, the third is from the Roman Period in France, with 2.790X.

2.4. Discussion

2.4.1. A broad origin of samples

2.4.1.1. From a geographic point of view

The dataset screened allowed the sequencing of new samples from a wide variety of regions and time periods. As seen in Figures 2 and 4, it represents 180 samples from 21 areas, from North-Western Europe to South-Eastern Asia. The geographical sample distribution is not equal, as some countries are over-represented compared to others. Iran, Germany and the Caucasus are the top 3 regions screened during this project. This is due to specific interest for these regions, as the three of them help understand questions linked to their location:

- ➔ What did early domesticated cattle genome diversity look like in Iran, and what was its relationship with the domestication centre in eastern Anatolia?
- ➔ As a corridor between Asia and Europe, can we detect signs of Asian cattle arrival in the Caucasus?
- ➔ Similarly, when did this Asian signal reach Germany (and more broadly Central Europe)?

In Germany and the Caucasus, the most pressing questions concern the possible appearance of Eastern cattle genomic components in Europe, via the human migratory spread of Yamnaya in Eastern Europe and later establishment of Corded Ware populations in Central Europe. This westward human migration has been clearly shown by ancient human genetics (Haak et al., 2015, Juras et al., 2018, Lazaridis et al., 2025), but has not been detected in domesticated cattle. This is investigated in Chapter 4; it would be expected that pastoralists would move with their domesticated animals and indeed such has been suggested in ancient sheep (Daly et al., 2025). 19 samples from the Caucasus were screened, mostly from the Bronze Age, while in Germany, 24 new samples from the Neolithic to Iron Age were added to the project. Adjacent countries such as Hungary (11 samples) and Ukraine (9 samples) were targeted in screening to search for a signal entering

Eastern Europe. In Central Europe, in addition to the German samples, other individuals were selected from Poland, Serbia, Italy, France, Switzerland, and also the UK.

Iran, despite its proximity to the Euphrates valley, seemed to have a low interest in cattle economy until the 6th millennium BC (Arbuckle et al., 2016), and few genomics data are available (only mitochondrial data in Scheu et al., 2015, while several samples were published in Verdugo et al., 2019). A number of new samples, and the improvement of published ones (thanks to the bleach protocol or additional sequencing of older extractions) aim at helping to more deeply characterise ancient Iranian cattle (see Chapter 5). A particularly interesting feature of these newly screened samples is their age, as 21 date from the Bronze Age or earlier. Several Neolithic/Chalcolithic samples were hence processed with the aim of clarifying domesticated cattle history in Iran.

As a side note for Iran, two samples pre-date domestication (DD089 and DD115 are from the Paleolithic); one of them did not contain enough DNA to match a reference genome, while the other had the best-matching result with *Bos taurus*. This reference genome is also used to identify *Bos primigenius*, as no reference exists yet for this species (Park et al., 2015 used UMD3.1, while Rossi et al., 2024 used ARS-UCD1.2, both European taurine reference genomes). Indeed, it is expected that these two samples are aurochsen considering their estimated age. Their sequencing would have helped describe wild populations in Iran, which are for now unknown.

A selection of samples also come from other non-European locations, namely: Pakistan, Saudi Arabia, Anatolia, Jordan, Turkmenistan. These countries, in association with Iran, could help characterise the emergence of animal domestication in the Near-East, and understand its spread in South-Western Asia (Chapter 5).

2.4.1.2. From a wide range of periods

Studying species domestication involves processing bones from the Neolithic, a period during which the domestication of plants and animals emerged. This event is sometimes called the “Neolithic Revolution” (Hole 1984), and is further described in a review by Horwitz et al. (1999), indicating phases of domestication and highlighting its importance for humans. As cattle were domesticated during the Neolithic, investigating

genetic data from this time period is crucial, as the genomes of very early domesticated individuals can inform on the process involved. Figure 3 shows that 57 samples dating from the Neolithic and 23 from the Chalcolithic were processed, along with 55 Bronze Age samples. These periods are best represented in the screening dataset, allowing us to see genetic shifts, if any, from the early set of domesticated cattle to a more established population in the Bronze Age, within the various geographical areas. The map on Figure 4 is a good indicator of how both time and place are sampled variables in this dataset. Both Western and Eastern countries are present, and for both early and later time periods are available in sampling. Germany is particularly well represented with many samples from three time periods, but Europe in general is sampled from Neolithic through Medieval times. The Near-East is also characterized from the Neolithic to the Bronze Age, with a few Iron Age samples to complete this dataset. This project is hence based on a strong screening strategy aimed at extending existing ancient cattle data in both space and time.

2.4.2. Endogenous content through space and time

With the dataset described above, we aimed to appropriately characterise past populations of Europe and South-Western Asia. However, this goal is based on hopes that DNA was preserved well in all samples. We know this is not the case, and endogenous content varied throughout the dataset (from 0 to almost 80%; Figure 6). We find that unfortunately, DNA preservation varies greatly with place of origin, and it is harder to retrieve from older samples, even when drilling the same type of bone.

2.4.2.1. Bone drilling and DNA preservation

The drilling strategy targeted petrous bones and teeth, and also a few long bones were attempted. This choice comes from previous studies highlighting the utility of dense bones for DNA extraction. Indeed, it was shown in Gamba et al., 2014 that petrous bones allowed better DNA preservation than other bones, a result that was confirmed by Pinhasi et al., 2015 and Parker et al., 2020. This is why the dataset screened contains so many petrous bones (105 samples on Figure 5, 45 being teeth while other bone categories have less than 7 samples - 12 if we group every long bone). Our results for endogenous content are in accordance with what has been previously published: petrous bones afford an

enhanced possibility to extract target DNA than other bones. This is visible in Figure 8, with a median endogenous content from petrous bones of 17.245%, the highest median of all bone types. Teeth were the remains of choice when extracting DNA before petrous bones were investigated. But here, the median endogenous content is in the same range as other bones (petrous bones excluded). However, they do show outliers with strong DNA content, reaching as high as 51.314%, a value not shown in any other non-petrous bone. Other papers have shown that teeth should not be overlooked as a targeted element, as they can still be a good DNA reservoir compared to petrous and postcranial bones (Haarkötter et al., 2023). There is also the possibility that a bias was introduced in this dataset. Indeed, Hansen et al., 2017 compared both petrous bones and teeth, and concluded that teeth can be as good if they visually look well preserved (complete and closed roots are present). In our project, sample drilling is dependent on availability, and if no petrous bones were present, teeth were selected as best as possible, but had variable preservation states (see Figure 13 below). Teeth samples were divided in two groups based on preservation, “bad” and “good”, and we found out a statistically significant difference in endogenous content (Figure 9). Teeth in a better preservation state seem to contain a higher quantity of DNA compared to the poorer preserved ones.

Other bone categories, although sampled in lower numbers (with a minority yielding good endogenous content), show DNA preservation in this project and other papers (Parker et al., 2020, Haarkötter et al., 2023) that is not good enough to fully rely on these skeletal elements.

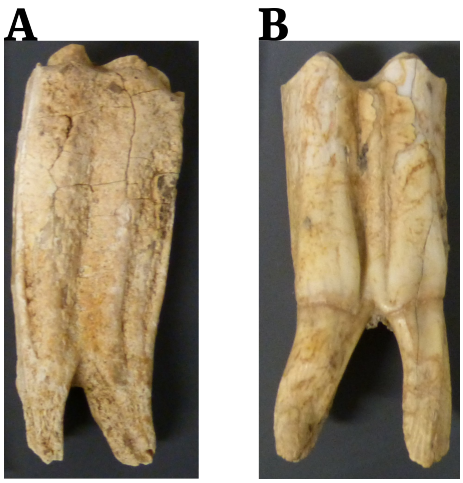

Fig. 13. Comparison of two teeth samples in different conservation states. A) The sample DD013 is a badly preserved tooth showing short and open roots, and an overall damaged structure. It is the worst endogenous content obtained from teeth. B) DD008 is a good sample with intact roots and better integrity. It led to a 32.143% endogenous content, one of the best of its category.

2.4.2.2. DNA preservation through time

The dataset goes back as far as the Paleolithic, but results in endogenous content vary greatly among samples (Figure 10). When looking at the relationship between time and EC, it looks like very old samples have poorer values than more recent ones (Figure 10), which was already suggested by (Lindahl 1993; Allentoft et al., 2012). Average and median values for EC shown in Table 6 indicate that Neolithic is the worst period for a good DNA preservation, but values improve in the Chalcolithic before degrading in the Bronze Age (while being higher than for the Neolithic). EC then gradually improves through the Iron Age, Antiquity and finally Medieval times. However, numbers of samples are not equivalent between all time periods. Antiquity and Medieval times are not as well represented as the Neolithic and Bronze Age, which can influence their results, especially for Medieval times. This can also explain the sudden bump in average EC for the Chalcolithic, which would be expected to be between the average EC of the Neolithic and Bronze Age. A correlation between time period and EC is shown on Figure 11, with positive relationship between three time periods and their corresponding values in DNA content. However, the earlier the period, the less strong is the correlation of time with endogenous

content. Coefficients are very low for “Early” and “Middle” (Neolithic-Chalcolithic, and Bronze Age - Iron Age respectively), at 0.093 and 0.053. Later samples, from Antiquity / Roman Period and Medieval times show a strong relationship between their age and their EC (R^2 = 0.568). This is interesting as it suggests a strong time dependent effect that is focused on recent centuries. This work accords with prior opinion that time does have an impact on DNA degradation, but cannot be the only factor involved.

2.4.2.3. Inequalities in DNA content between regions

As time seems to not be the only one factor impacting our results, geography was investigated as well. We find that European samples have better DNA preservation than samples from the Caucasus, the Levant, Iraq, Saudi Arabia and Pakistan (Figure 7). Turkmenistan and Uzbekistan, in Central Asia, have much better results than these countries, which are more southern. Iran seems to have mixed results depending on sampling location. The worst samples are located in the Western part of the country, close to its border with Iraq. All these areas with poor preservation have in common a hot arid climate, which is known to accelerate DNA decay (Bollongino et al., 2008; Pinhasi et al., 2015). Even when samples from very arid countries (Saudi Arabia and Pakistan) were extracted several times (Table 8), endogenous content remained extremely low, despite the drilling of petrous bones, which was recommended for difficult environments in Pinhasi et al., 2015.

2.4.3. Improvement of the ancient cattle database for future analyses

2.4.3.1. Testing the bleach protocol

Samples screened or published in previous projects (Verdugo et al., 2019) were screened a second time using the new bleach protocol (Table 7). This led to the improvement of 13 samples dating from the Neolithic (n=9), Chalcolithic (n=2) and Bronze Age (n=2), and the deeper sequencing of 11 of them. This shows the efficiency of bleach extraction even on Neolithic samples, which managed to rescue more than half of the samples tested. On samples coming from arid regions however (Saudi Arabia and Pakistan, Table 8), the bleach was not able to retrieve DNA from the bones, and did not improve the

result for the one sample (DD047) that was sent for deeper sequencing (with the non-bleach library). Very old or fragile (due to arid environments) samples hence should not be processed with a protocol that includes bleach.

2.4.3.2. Final coverage

This project led to the addition of 109 samples to the existing database of ancient cattle, 3 other samples (Shr1, Abu1 and Abu2) did not have enough DNA and failed to have a coverage high enough for analyses. The 109 samples have at least 0.010X coverage, but can reach 9.140X. However, as seen in Figure 12, many samples fall below the 1X threshold, and go even as low as 0.050X. These are promising results as these samples can be improved in the future, via more sequencing or after sufficient coverage is attained, genome imputation (Erven et al., 2024). With improved coverage these were included in analyses in the next chapters aimed at studying population structure to uncover admixture in European samples and characterise Middle-Eastern taurine populations.

Table 1. Samples screened for endogenous content and deeper sequencing results. The column "Wet Lab" indicates the person who performed the drilling and DNA extraction/library (DD = Deborah Diquelou, LL = Louis L'Hôte, AMH = Áine Halpin, KD = Kevin Daly, MV = Marta Verdugo). Samples with sufficient coverage to be plotted on a PCA in later analyses are indicated in the last column.

Sample	Alias	Country	Site	Period	Bone type	Sex	% endogenous DNA	Sent for deeper sequencing?	Coverage (X)	Wet Lab	Nuclear analysis
DD001		Hungary	Szazhalombatta-Sancalja	Late Bronze Age	Long bone		0.496			DD	
DD002	Sza11	Hungary	Szazhalombatta-Foldvar	Early Bronze Age	Tooth	Female	3.625	Yes	0.030	DD	X
DD003	Sza6	Hungary	Szazhalombatta-Foldvar	Early Bronze Age	Tooth	Male	14.021	Yes	0.480	DD	X
DD004		Saudi Arabia	Al' Ula	Neolithic	Petrous		0.030			DD	
DD005		Saudi Arabia	Al' Ula	Neolithic	Petrous		0.012			DD	
DD006		Saudi Arabia	Al' Ula	Neolithic	Petrous		0.640			DD	
DD007-1	Sza15	Hungary	Szazhalombatta-	Middle	Tooth	Male	1.385	Yes	0.007	DD	X

			Foldvar	Bronze Age							
DD007-2	Sza7	Hungary	Szazhalombatta-Foldvar	Middle Bronze Age	Long bone	Female	13.525	Yes	0.470	DD	X
DD008	Sza8	Hungary	Szazhalombatta-Foldvar	Middle Bronze Age	Tooth	Female	32.143	Yes	1.020	DD	X
DD009	Sza9	Hungary	Szazhalombatta-Sancalja	Early Bronze Age	Tooth	Female	12.488	Yes	0.430	DD	X
DD010-1	Zha1	Ukraine	Zhabotin Settlement	Iron Age	Tooth	Male	45.418	Yes	0.390	DD	X
DD010-2		Ukraine	Zhabotin Settlement	Iron Age	Tooth		0.261			DD	
DD010-3		Ukraine	Zhabotin Settlement	Iron Age	Tooth		0.674			DD	
DD010-4	Zha2	Ukraine	Zhabotin Settlement	Iron Age	Tooth	Male	3.785	Yes	0.020	DD	X
DD011		Ukraine	Katerinovka cemetery	Iron Age	Tooth		0.035			DD	
DD012	Sha1	Ukraine	Shakter	Chalcolithic	Tooth	Male	26.415	Yes	0.370	DD	X
DD013		Ukraine	Shakter	Chalcolithic/Early Bronze Age	Tooth		0.011			DD	
DD014	Mam1	Ukraine	Mamai-Gora	Iron Age	Tooth	Female	13.129	Yes	0.410	DD	X
DD015		Germany	Sonnenschacht cave	Bronze Age	Tooth		0.070			DD	
DD016		Germany	Sonnenschacht cave	Bronze Age	Tooth		0.416			DD	
DD017	Son1	Germany	Sonnenschacht cave	Bronze Age	Tooth	Female	1.351	Yes	0.010	DD	X
DD018		Iran	Hasanlu	Bronze Age	Petrous		43.587*			DD	
DD019	Sza10	Hungary	Szazhalombatta-Sancalja	Early Bronze Age	Long bone	Male	13.743	Yes	0.470	DD	X
DD020	Ulu4	Anatolia	Ulucak	Neolithic	Petrous	Male	5.901	Yes	0.500	DD	X
DD021	Ulu1	Anatolia	Ulucak	Neolithic	Petrous	Female	6.393	Yes	0.190	DD	X
DD022	Ulu2	Anatolia	Ulucak	Neolithic	Petrous	Female	19.774	Yes	0.460	DD	X
DD023	Ulu5	Anatolia	Ulucak	Neolithic	Petrous	Male	7.597	Yes	0.090	DD	X

DD024	Ulu3	Anatolia	Ulucak	Neolithic	Petrous	Male	17.245	Yes	0.550	DD	X
DD026		Iran	Abdol Hossein	Early Neolithic	Petrous		0.100			DD	
DD027	Ust1	North Caucasus	Ust Dzheguta	Bronze Age	Astragalus	Male	0.772	Yes	0.020	DD	X
DD028		South Caucasus	Otkhara	Iron Age	Tooth		0.288			DD	
DD029	Sto2	North Caucasus	Stoderovskoe 1	Bronze Age	Tooth	Male	1.086	Yes	0.020	DD	X
DD030		North Caucasus	Gumbashi 1	Iron Age	Tooth		0.108			DD	
DD031		North Caucasus	Kabardinka 2	Bronze Age	Astragalus		0.197			DD	
DD032	Dam1	North Caucasus	Damba Kalaus 2	Bronze Age	Tooth	Male	0.608	Yes	0.010	DD	X
DD033	Ipa1	North Caucasus	Ipatovo 3	Bronze Age	Metatarsus	Male	1.720	Yes	0.010	DD	X
DD034	Kom1	North Caucasus	Komsomolec 1	Iron Age	Tooth	Male	11.112	Yes	0.400	DD	X
DD035	Shr1	North Caucasus	Sharakhalsun 2	Bronze Age	Tooth	Male	0.804	Yes	0.000	DD	X
DD036		North Caucasus	Sharakhalsun 2	Bronze Age	Tooth		0.030			DD	
DD037	Shr2	North Caucasus	Sharakhalsun 3	Bronze Age	Long bone	Female	2.944	Yes	0.050	DD	X
DD038		North Caucasus	Ransyrt 1	Iron Age	Tooth		0.175			DD	
DD039	Sto1	North Caucasus	Stoderovskoe 1	Bronze Age	Tooth	Male	23.943	Yes	0.450	DD	X
DD040		North Caucasus	Sharakalsun 3	Bronze Age	Scapula		0.331			DD	
DD041	Sto3	North Caucasus	Stoderovskoe 1	Bronze Age	Tooth	Female	2.908	Yes	0.080	DD	X
DD042	Sza12	Hungary	Szazhalombatta-Foldvar	Bronze Age	Long bone	Male	3.142	Yes	0.010	DD	X
DD043-1	Sza13	Hungary	Szazhalombatta-Foldvar	Bronze Age	Long bone	Male	21.472		0.690	DD	X
DD044-1	Sza14	Hungary	Szazhalombatta-Sancalja	Bronze Age	Long bone	Female	2.319			DD	

DD045	Sha2	Ukraine	Shakter	Chalcolithic	Tooth	Male	10.775	Yes	0.440	DD	X
DD046		Saudi Arabia	Al' Ula	Neolithic	Petrous		0.020			DD	
DD047	Aak1	Saudi Arabia	Al' Ula	Neolithic	Petrous	Female	0.182	Yes	0.002	DD	X
DD048		Saudi Arabia	Al' Ula	Neolithic	Petrous		0.020			DD	
DD049	Asp1	UK	Aspen Gardens, Stotfold	Medieval	Petrous	Female	56.614	Yes	1.850	DD	X
DD050	Red1	UK	Redcliffe	Medieval	Petrous	Male	58.049	Yes	1.680	DD	X
DD051	Red2	UK	Redcliffe	Medieval	Petrous	Male	57.200	Yes	3.310	DD	X
DD052	Wal1	UK	Wallingford	Medieval	Petrous	Female	53.684	Yes	0.940	DD	X
DD053	Cri1	UK	Cribbs Causeway	Roman Period	Petrous	Male	32.597	Yes	0.690	DD	X
DD054	Rug1	UK	Rugby SUE16	Iron Age	Petrous	Female	57.123	Yes	9.140	DD	X
DD055	Wro1	UK	Wroughton	Iron Age	Petrous	Female	55.566	Yes	1.630	DD	X
DD056		Pakistan	Chanhu Daro	Bronze Age	Humerus	Female	0.000	Yes	0.000	DD	
DD057		Pakistan	Chanhu Daro	Bronze Age	Calcaneus	Female	0.167	Yes	0.000	DD	
DD058		Pakistan	Chanhu Daro	Bronze Age	Humerus	Female	1.403	Yes	0.000	DD	
DD059		Pakistan	Chanhu Daro	Bronze Age	Petrous		0.144			DD	
DD060		Pakistan	Chanhu Daro	Bronze Age	Petrous		0.792			DD	
DD061	Mon4	Turkmenistan	Monjukli Depe	Chalcolithic	Petrous	Female	27.156	Yes	0.490	DD	X
DD062	Mon5	Turkmenistan	Monjukli Depe	Chalcolithic	Petrous	Female	15.586	Yes	0.310	DD	X
DD063	Mon6	Turkmenistan	Monjukli Depe	Chalcolithic	Petrous	Female	10.897	Yes	0.220	DD	X
DD064	Mon7	Turkmenistan	Monjukli Depe	Chalcolithic	Petrous	Female	27.537	Yes	0.440	DD	X
DD065	Mon8	Turkmenistan	Monjukli Depe	Chalcolithic	Petrous	Male	35.407	Yes	1.270	DD	X
DD066	Mon9	Turkmenistan	Monjukli Depe	Chalcolithic	Petrous	Male	18.199	Yes	0.460	DD	X
DD067	Mon10	Turkmenistan	Monjukli Depe	Chalcolithic	Petrous	Female	29.299	Yes	0.740	DD	X
DD068		Iran	Darre-ye Bolaghi	Neolithic	Petrous		0.522			DD	

DD069	Bol1	Iran	Darre-ye Bolaghi	Neolithic	Petrous	Female	23.910	Yes	0.690	DD	X
DD070	Bul2	Uzbekistan	Tilla Bulak	Late Bronze Age	Petrous	Female	34.638	Yes	1.090	DD	X
DD071	Brn1	France	Brionne	Roman Period	Petrous (L)	Female	46.008	Yes	1.000	DD	X
DD072	Brn2	France	Brionne	Roman Period	Petrous (L)	Female	50.989	Yes	1.180	DD	X
DD073	Brn3	France	Brionne	Roman Period	Petrous (R)	Female	53.042	Yes	2.790	DD	X
DD074	Brn4	France	Brionne	Roman Period	Petrous (L)	Male	55.037	Yes	1.120	DD	X
DD075	Brn5	France	Brionne	Roman Period	Petrous (L)	Female	51.665	Yes	1.430	DD	X
DD076		Germany	Künzing-Unternberg	Neolithic	Tooth P4		0.400			DD	
DD077	Kun1	Germany	Kunzig-Unternberg	Neolithic	Tooth M1	Female	3.388	Yes	0.050	DD	X
DD078	Kun2	Germany	Kunzig-Unternberg	Neolithic	Petrous	Female	36.306	Yes	1.000	DD	X
DD079	Kun3	Germany	Kunzig-Unternberg	Neolithic	Petrous	Female	43.296	Yes	1.000	DD	X
DD080	Man1	Germany	Manching	Iron Age	Tooth M1	Male	8.425	Yes	0.100	DD	X
DD081		Germany	Manching	Iron Age	Tooth M2		0.339			DD	
DD082	Man2	Germany	Manching	Iron Age	Tooth M1	Male	19.000	Yes	0.440	DD	X
DD083	Man3	Germany	Manching	Iron Age	Petrous	Female	46.166	Yes	1.060	DD	X
DD084	Man4	Germany	Manching	Iron Age	Petrous	Female	45.810	Yes	1.040	DD	X
DD085	Man5	Germany	Manching	Iron Age	Petrous	Male	42.714	Yes	0.520	DD	X
DD086	Man6	Germany	Manching	Iron Age	Petrous	Female	45.047	Yes	1.050	DD	X
DD087		Pakistan	Chanhu Daro	Bronze Age	Tooth		0.451			DD	
DD088		Pakistan	Chanhu Daro	Bronze Age	Tooth		3.153	Yes	0.000	DD	
DD089		Iran	Yafteh	Paleolithic	Tooth		0.130			DD	
DD090		Iran	Tepe Qabrestan	Chalcolithic	Tooth		0.026			DD	

DD091		Pakistan	Chanhu Daro	Bronze Age	Tooth		0.088			DD	
DD092	Qab1	Iran	Tepe Qabrestan	Chalcolithic	Tooth	Male	0.014	Yes	0.040	DD	X
DD093		Iran	Masjed Jameh	Antiquity	Petrous		0.015			DD	
DD094		South Caucasus	Orchosani	Chalcolithic	Petrous		0.252			DD	
DD095		South Caucasus	Orchosani	Chalcolithic	Petrous (R)		68.516			DD	
DD096	Tap1	Iran	Tappeh Hesar Markazi	Bronze Age	Petrous	Male	39.818	Yes	0.940	DD	X
DD097	For1	Iran	Tepe Forudgah	Bronze Age	Petrous	Male	41.942	Yes	0.410	DD	X
DD098	Til1	Anatolia	Tilbeshar	Chalcolithic/ Bronze Age	Petrous	Female	34.986	Yes	0.060	DD	X
DD099	Dav1	Iran	Dava Goz	Chalcolithic	Petrous (R)	Female	43.969	Yes	1.000	DD	X
DD100	Dav2	Iran	Dava Goz	Chalcolithic	Petrous	Male	36.248	Yes	1.730	DD	X
DD101	Cer1	Germany	Cherry Tree Cave	Neolithic?	Tooth	Female	1.635	Yes	0.005	DD	X
DD102	Cer2	Germany	Cherry Tree Cave	Neolithic	Metatarsus	Female	42.036	Yes	0.090	DD	X
DD103		Germany	Kaspauer	Bronze Age	Tibia		0.162			DD	
DD104		Germany	Kaspauer	Bronze Age	Tooth		0.261			DD	
DD105	Kas1	Germany	Kaspauer	Bronze Age	Femur	Female	4.621	Yes	0.040	DD	X
DD106		Germany	Wattendorf-Motzenstein	Neolithic	Tooth		0.512			DD	
DD107	Has6	Iran	Hasanlu	Iron Age	Petrous	Female	47.468	Yes	0.050	DD	X
DD108	Kir5	Germany	Kirschbaumhöhle	Neolithic	Metatarsus	Female	8.440	Yes	0.060	DD	X
DD109	Kir6	Germany	Kirschbaumhöhle	Neolithic	Metatarsus	Female	4.196	Yes	0.030	DD	X
DD110	Gub1	Italy	Gubbio	Bronze Age	Petrous	Female	62.134	Yes	1.050	DD	X
DD111	Gub2	Italy	Gubbio	Bronze Age	Petrous	Female	58.232	Yes	0.550	DD	X
DD112	Gub3	Italy	Gubbio	Bronze Age	Petrous	Female	46.369	Yes	0.420	DD	X
DD113	Gub4	Italy	Gubbio	Bronze Age	Petrous	Female	22.350	Yes	0.440	DD	X

DD114	Gub5	Italy	Gubbio	Bronze Age	Petrous	Male	43.914	Yes	0.490	DD	X
DD115		Iran	Ali Tepe	Epipaleolithic	Petrous		1.138			DD	
DD116	Has7	Iran	Hasanlu	Early Bronze Age	Petrous	Female	43.002	Yes	0.500	DD	X
DD117	Nad1	Iran	Naderi Tepe	Early Bronze Age	Petrous	Female	29.910	Yes	0.300	DD	X
DD118	Nad2	Iran	Naderi Tepe	Bronze Age	Petrous	Male	29.352	Yes	0.310	DD	X
DD119		Iran	Naderi Tepe	Chalcolithic/ Bronze Age	Petrous		43.739	Yes	0.000	DD	
DD120	Soh1	Iran	Soha Chai Tepe	Chalcolithic	Petrous	Male	43.645	Yes	0.490	DD	X
DD121	Zag1	Iran	Zagheh Abyek	Chalcolithic	Petrous	Male	18.393	Yes	0.210	DD	X
DD122	Qaz2	Iran	Alou	Chalcolithic	Petrous	Male	32.210	Yes	0.260	DD	X
DD123		Iran	Naderi Tepe	Iron Age	Petrous		72.816			DD	
DD124	Has8	Iran	Hasanlu	Iron Age	Petrous	Female	44.798	Yes	0.550	DD	X
DD125	Has9	Iran	Hasanlu	Iron Age	Petrous	Male	37.974	Yes	0.400	DD	X
DD126		Saudi Arabia	Al'Ula	Neolithic	Petrous		0.031			DD	
DD127	Cer3	Germany	Cherry Tree Cave	Late Neolithic?	Tooth	Female	51.314	Yes	0.520	DD	X
DD129	Cer4	Germany	Cherry Tree Cave	Iron Age	Femur	Male	5.854	Yes	0.150	DD	X
ITTQ2020		Italy	Tarquinia	Roman Period	Astragalus		0.442			DD	
ITTQ2354	Tar1	Italy	Tarquinia	Roman Period	Metacarp	Male	2.405	Yes	0.009	DD	X
ITTQ2662		Italy	Tarquinia	Roman Period	Metacarp		0.110			DD	
ITTQ2991		Italy	Tarquinia	Late Hellenistic	Metacarp		0.724			DD	
ITTQ3263		Italy	Tarquinia	Villanovan	Metatarsus		1.050			DD	
ITTQ545	Tar2	Italy	Tarquinia	Iron Age	Astragalus	Male	0.935	Yes	0.009	DD	X
ITTQ797		Italy	Tarquinia	Villanovan	Tooth		1.078			DD	
LL006		Iran	Tepe jarali	Bronze Age	Petrous		0.089			LL/DD	

LL033		Iraq	Bestanstur	Neolithic	Humerus		0.235			LL/DD	
LL037		Iraq	Bestanstur	Neolithic	Metacarp		0.011			LL/DD	
LL040		Saudi Arabia	Al'Ula	Neolithic	Petrous		0.018			LL/DD	
LL041		Saudi Arabia	Al'Ula	Neolithic	Petrous		0.015			LL/DD	
LL042		Saudi Arabia	Al'Ula	Neolithic	Petrous		0.138			LL/DD	
LL043		Saudi Arabia	Al'Ula	Neolithic	Petrous		0.017			LL/DD	
LL055		Iran	Ali Tapeh	Neolithic	Petrous		0.548			LL/DD	
AMH65		South Caucasus	Gadachrili Gora	Neolithic	Tooth		3.005			ÁMH	
AMH66		South Caucasus	Gadachrili Gora	Neolithic	Tooth		1.156			ÁMH	
KD380		Jordan	El-Hemmeh	Neolithic	Petrous		0.058			KD/DD	
KD384		Jordan	El-Hemmeh	Neolithic	Petrous		0.252			KD/DD	
KD388		Jordan	El-Hemmeh	Neolithic	Petrous		0.304			KD/DD	
KD159	Tqa3	Israel	Tel es-Qashish	Bronze Age	Skull fragments	Female	1.501	Yes	0.020	KD/DD	X
MV015		Anatolia	Barcin Hoyuk	Neolithic	Metacarp		0.875			MV/DD	
MV020		Anatolia	Çukuriçi Höyük	Neolithic	Petrous		1.732			MV/DD	
MV021		Anatolia	Çukuriçi Höyük	Neolithic	Petrous		1.410			MV/DD	
MV023	Cuk1	Anatolia	Çukuriçi Höyük	Neolithic	Petrous	Male	4.860	Yes	0.090	MV/DD	X
MV024	Cuk2	Anatolia	Çukuriçi Höyük	Neolithic	Petrous	Male	9.025	Yes	0.440	MV/DD	X
MV026		Switzerland	Sutz-Lattrigen	Neolithic	Petrous		2.509	Yes	0.000	MV/DD	
MV028		Switzerland	Sutz-Lattrigen	Neolithic	Petrous		1.256	Yes	0.000	MV/DD	
MV045	Men3	Anatolia	Mentese	Neolithic	Petrous	Male	9.249	Yes	0.330	MV/DD	X
MV052		Netherlands	Schipluiden	Neolithic	Petrous (R)		4.704			MV/DD	
MV053	Sch3	Netherlands	Schipluiden	Neolithic	Petrous (L)	Male	5.540	Yes	0.040	MV/DD	X
MV069	Bla1	Serbia	Blagotin	Neolithic	Petrous	Male	13.716	Yes	0.300	MV/DD	X

MV072	Bla3	Serbia	Blagotin	Neolithic	Petrous (L)	Male	6.856	Yes	0.020	MV/DD	X
MV076	Bla2	Serbia	Blagotin	Early Neolithic	Petrous (L)	Male	13.927	Yes	0.060	MV/DD	X
MV079		Serbia	Stubline	Neolithic	Petrous		0.131			MV/DD	
MV086	Bub1	Serbia	Bubanj	Neolithic	Petrous	Female	34.130	Yes	0.800	MV/DD	X
MV087	Bub2	Serbia	Bubanj	Neolithic	Petrous	Female	13.727	Yes	0.110	MV/DD	X
MV111	Abu1	Israel	Abu Gosh	MPPNB	Petrous	Male	0.200		0.000	MV/DD	
MV135	Ace1	Anatolia	Acemhöyük	Bronze Age	Petrous	Female	29.690	Yes	0.310	MV/DD	X
MV151	Abu2	Israel	Abu Gosh	Mid-Late PPNB	Petrous		0.100		0.000	MV/DD	
MV155	Tqa2	Israel	Tel es-Qashish	Bronze Age	Petrous	Female	9.584	Yes	0.560	MV/DD	X
MV159	Plo6	Serbia	Plocnik	Neolithic	Petrous (L)	Male	24.040	Yes	0.460	MV/DD	X
MV176	Far1	Iran	Fars	Neolithic	Petrous	Female	10.525	Yes	0.110	MV/DD	X
MV184	Ruc1	Poland	Rzucewo	Neolithic	Petrous	Male	2.187	Yes	0.008	MV/DD	X
MV185	Ruc2	Poland	Rzucewo	Neolithic	Petrous	Female	11.831	Yes	0.490	MV/DD	X
MV209	Tes1	Israel	Tel-el-Safi	Early Bronze Age	Petrous	Male	5.244	Yes	0.040	MV/DD	X
MV234	Mon1	Turkmenistan	Monjukli Depe	Chalcolithic	Petrous	Female	2.875		0.001	MV/DD	X
MV235	Mon2	Turkmenistan	Monjukli Depe	Chalcolithic	Petrous	Female	5.557	Yes	0.060	MV/DD	X
MV236	Mon3	Turkmenistan	Monjukli Depe	Chalcolithic	Petrous	Female	4.442	Yes	0.050	MV/DD	X

3. Mitochondrial DNA diversity in ancient cattle

3.1. Introduction

Mitochondrial chromosomes, inherited strictly along the maternal lineage, are a fast mutating, non-recombining and thus valuable source of information with which to build useful phylogenies. Initial ancient DNA studies in cattle and other species were based on the mitochondrial genome which proved easier to retrieve than autosomal DNA (MacHugh et al., 1999; Troy et al., 2001). Additionally, and interestingly, the main haplogroup groupings (haplogroups) in modern domesticated taurine cattle, T1, T2 and T3 more specifically, are particularly well geographically structured (Troy et al., 2001). This structure was initially interpreted as a signal of distinct Near Eastern and African domestications of *Bos taurus* (Bradley et al., 1996) but subsequent work with whole mtDNA sequences negated this inference, pointing toward the shallow coalescence of these haplogroups (Bonfiglio et al., 2012; Achilli et al., 2009). This chapter will extend the description of this structure in ancient samples from that of Verdugo et al. (2019) placing newly sequenced samples into this phylogeny, increasing data for European samples and shedding light in lesser sampled regions of South-Western Asia. In particular, the first retrieved sample from Neolithic Saudi Arabia will be presented.

Several methods are available to study mitochondrial phylogenies, by reconstructing trees following a statistical model that will compare several sequences and build the most probable tree. Typical ones are parsimony or neighbour-joining trees, both are fast methods but are less favoured today due to the risk of inaccuracy. Parsimony trees (Cavalli-Sforza and Edwards, 1967) assume a similar mutation rate across the entire genome, and will choose the best tree explaining the differences observed between sequences with the least evolutionary changes. Neighbour-joining trees (Saitou and Nei, 1987) are based on distance matrices and will group together sequences that are closest to each other. The Maximum Likelihood inference model (Felsenstein, 1981), which has the advantage of taking into account the variability in mutation rate by applying an evolutionary model to the dataset, will be used in this chapter to build a phylogenetic tree from 215 mitogenomes. But mutation rate is not the only factor to be considered in

phylogenies. Bayesian methods and Markov Chain Monte Carlo (MCMC) (Huelsenbeck and Bollback, 2001) have been developed to add to the likelihood of a tree to be real known priors that will complete the model, allowing the computation of a posterior probability distribution of phylogenetic trees. MCMC are used to sample these posterior trees, targeting those with higher probability. BEAST (Drummond and Rambaut, 2007) was developed to gather these methods and output a consensus tree. The key element of BEAST for this chapter is the ability to use sample ages to date the tree. We will use it to confirm the age of Aak1, a Saudi Arabian sample hopefully dating back to the Neolithic (~ 7,000 years old) and assess when taurine haplogroups (T1-T2-T3) diverged from each other.

DNA extraction is very challenging in arid environments (see Chapter 1, and Bollongino et al., 2008). Screening Saudi Arabian samples turned out to be extremely difficult (Chapter 1, section 2.3.4.2 - Table 8) but one sample allowed the retrieval of DNA (Aak1). It came with very low coverage for nuclear data (less than 0.01 X), however the mitochondrial genome has been retrieved with good coverage: 20.13X. This result means the sample can be used for mitochondrial phylogenies, but radiocarbon dating failed, leaving uncertainty regarding its real age. BEAST becomes useful here as it will test if a tree is plausible with that sample when assigned a Neolithic date.

3.2. Material and Methods

3.2.1. Archaeological samples

215 samples are included in this chapter. They represent 181 ancient individuals (*Bos taurus* and *Bos primigenius*) and 34 modern samples (*Bos taurus* and *Bos indicus*, Supplementary Table S2). Out of the 181 ancient samples, 54 were processed in a previous project (Supplementary Table S1). The remaining 127 (Table 9) are either new or were improved in coverage, enabling analysis for this chapter.

DNA extraction, library construction and amplification were performed as described in Chapter 2, section 2.2.3.

3.2.2. Extraction of the mitochondrial genome from whole genome sequencing

3.2.2.1. Raw data

Sample processing was done as described in Chapter 2, sections 2.2.2 and 2.2.3 for the drilling and DNA extraction/library, and in section 2.2.4.1 for the trimming of Illumina adapters after sequencing. Collapsed trimmed files are then used for the mitochondrial pipeline described below.

3.2.2.2. Alignment and read groups

Trimmed and collapsed files were first aligned to the *Bos taurus* reference genome (V00654.1, Anderson et al., 1982), which belongs to the T3 haplogroup, using BWA v.0.7.13 (Li and Durbin, 2009) with the command:

```
bwa aln -l 1024 -n 0.01 -o 2 MT_reference_genome.fa
trimmed_file.fastq.gz > sample_MT.sai
```

After reconstruction of the mitochondrial genome, samples are placed in a maximum likelihood tree to verify their haplogroup. If they do not belong to the T3 haplogroup, trimmed files are aligned to the correct haplogroup:

- T1 = JN817304 (Genbank accession number, Bonfiglio et al., 2012)
- T2 = MZ901441 (Genbank accession number, Cubric-Curik et al., 2022)
- Q = KT184471 (Genbank accession number, Olivieri et al., 2015)
- R = FJ971087 (Genbank accession number, Achilli et al., 2009)
- P = Hxh2, an aurochs published in (Rossi et al., 2024)

For one sample, Sto3, its determined haplogroup E does not have a reference genome at the time of this project. We used another sample with the same haplogroup, Mon4, to re-align Sto3. Mon4 had been used as a reference genome in previous work by Dr. Conor Rossi (Rossi et al., 2024). The R reference genome was used to align the published sample ALT232 (Ginja et al., 2023).

The sai file was converted into a bam file by combining two tools : BWA samse and samtools view (Li et al., 2009). The file is then sorted with samtools sort. The conversion step also includes read groups set up with the option `-r`:

```
bwa samse -r "@RG\tID:sample_ID-
Sequencing_Lane\tPL:ILLUMINA\tLB:sample_ID-PCR-Index_P5-
Index_P7\tSM:sample_ID"
```

Duplicates are then removed using MarkDuplicates from Picard v2.20.3 (Broad Institute, 2019):

```
picard.jar MarkDuplicates REMOVE_DUPLICATES=true
```

3.2.2.3. Merging several PCRs and quality filters

To increase coverage, several PCRs can be made for the same samples, all of which will need to be merged together after alignment. We use samtools merge to do so. Another duplicate removal step can be done at this point as a sanity check.

The final bam file is then filtered to remove reads with a mapping quality lower than 30 (`samtools view -q 30`), and indexed with samtools.

3.2.2.4. Statistics

Each step generates a report used for statistics calculations, as described in the first chapter (`samtools flagstat`). Coverage is computed with qualimap v2.1.3 (`qualimap bamqc`, Okonechnikov et al., 2016).

3.2.3. Final mitochondrial genome

3.2.3.1. Fasta conversion

Bam files are converted into fasta files using ANGSD v0.941 (Korneliussen et al., 2014), using the following command options:

```
angsd -doFasta 2 -doCounts 1 -setMinDepth 3 -minQ 20 -minMapQ 30
```

These flags will instruct ANGSD to make a consensus file, using reads with a minimum mapping quality of 30 and minimum base quality of 20. Each site must be covered 3 times to be kept in the consensus fasta.

3.2.3.2. Indels correction and circularisation

Every fasta file was visualised with IGV (Thorvaldsdóttir et al., 2013) to detect and manually correct eventual insertions or deletions. Finally, to account for the circular shape of the mitochondrial genome, the fasta file is modified to concatenate the first and last 15bp of the genome.

3.2.4. Mitochondrial Capture

17 samples were selected for mitochondrial capture, as whole genome shotgun sequencing did not permit enough mitochondrial DNA to be retrieved, but their endogenous content suggested a possibility for mitochondrial sequence enrichment. 16 samples were amplified with 3 PCR reactions each, the last one from Neolithic Saudi Arabia was amplified with 7 PCR reactions, and processed with the MYcroarray MYbaits® In-Solution Sequence Capture kit, and sequenced on a miSeq Illumina platform. They were then processed with the pipeline presented above.

3.2.5. Phylogeny and haplogroup

3.2.5.1. Maximum likelihood tree

Ancient samples that achieved at least 90% of the genome covered at minimum 3X were selected and files concatenated with the moderns used as reference for the different haplogroups (Supplementary Table S2), and a Water Buffalo sample used as outgroup. Multiple sequence alignment was performed with MUSCLE (Edgar 2004). The fasta file was converted into a Phylip format on SeaView v5.0.5 (Gouy et al., 2021) and used to compute a maximum likelihood tree with 100 bootstraps with PhyML (Guindon et al., 2010). The Smart Model Selection tool (Lefort et al., 2017), implemented in PhyML, was used to choose the HKY85 substitution model. The tree was then visualised with FigTree v1.4.4 (Rambaut, 2018).

3.2.5.2. BEAST

The software BEAST v2.6 (Bouckaert et al., 2019) was used to build a phylogeny with goals of dating key nodes including to confirm the supposed date for Aak1, the domestic cattle from Neolithic Saudi Arabia. This sample was indirectly dated by archaeologists but direct radiocarbon dating failed. A set of ancient samples presenting the haplogroups T1, T2, T3 and Q were aligned using MUSCLE along with published samples belonging to the T1 haplogroup (Bonfiglio et al., 2012, Appendix - Supplementary Table S3). The alignment is saved as a nexus file, and sample dates are added at the end of the file, using a uniform distribution to account for uncertainty. BEAUTi, the graphic interface included with BEAST, was used to parameterize the runs, choosing "BEAST Model Test" to determine the best substitution model and a "Relaxed Clock Log Normal" for the substitution rate. The MCMC had a chain length of 300 million and a pre-burnin of 3 million. The 3 runs were checked for convergence with Tracer v1.7.2 (Rambaut et al., 2018), merged with LogCombiner and Treeannotator (included with BEAST) was then used to build a consensus tree, visualised on FigTree v1.4.4.

3.3. Results

3.3.1. Mitochondrial pipeline output

As described in Chapter 2 section 2.3.5, 178 samples were sent for deeper sequencing. 106 of these samples were processed with the mitochondrial pipeline and are included in the tree in Figure 14. An additional set of 21 ancient samples either screened or sequenced in previous projects were amplified and added to the dataset. The sample with the lowest coverage is Tar1, 5.7X. The highest value is 590.9X for Sza13. On average, the 127 samples sequenced or amplified in this chapter reached a coverage of 71.95X, with a median of 56.5X.

Mitochondrial capture using an array of RNA-baits was attempted on 17 samples from Central and Eastern Europe (n=6), the Caucasus (n=5) and South-Western Asia (n=6), which date from the Neolithic to the Iron Age. It was successful for 14, the 3 failing being from the Caucasus (Bronze Age) and Iran (Paleolithic and Neolithic). These 14 samples reached on average a coverage of 61.0X, the median being at 54.7X. The lowest coverage

achieved is 12.8X for a Neolithic sample from Poland, Ruc1, while the highest value is 187.8X for Sza12 (Bronze Age Hungary).

3.3.2. Mitochondrial phylogeny

3.3.2.1. Maximum likelihood tree

The maximum likelihood tree obtained from the 127 ancient samples processed for this chapter, plus the 54 published ones (and the 34 moderns), was collapsed in Figure 14 onto the major haplogroups found in *Bos primigenius*, *Bos taurus* and *Bos indicus*: I, E, R, K, C, P, Q, T5, T1, T2 and T3-T4. One sample, Far1 (Neolithic Iran), is located at the base of the T1-T2-T3-T4 cluster, but not within the T5 group.

The I cluster, at the base of the tree, is the indicine haplogroup. E and R on one side, and K and C on the other form two clusters, both basal to P, Q and lastly T. The detailed branches of the tree can be seen in Figure 15.

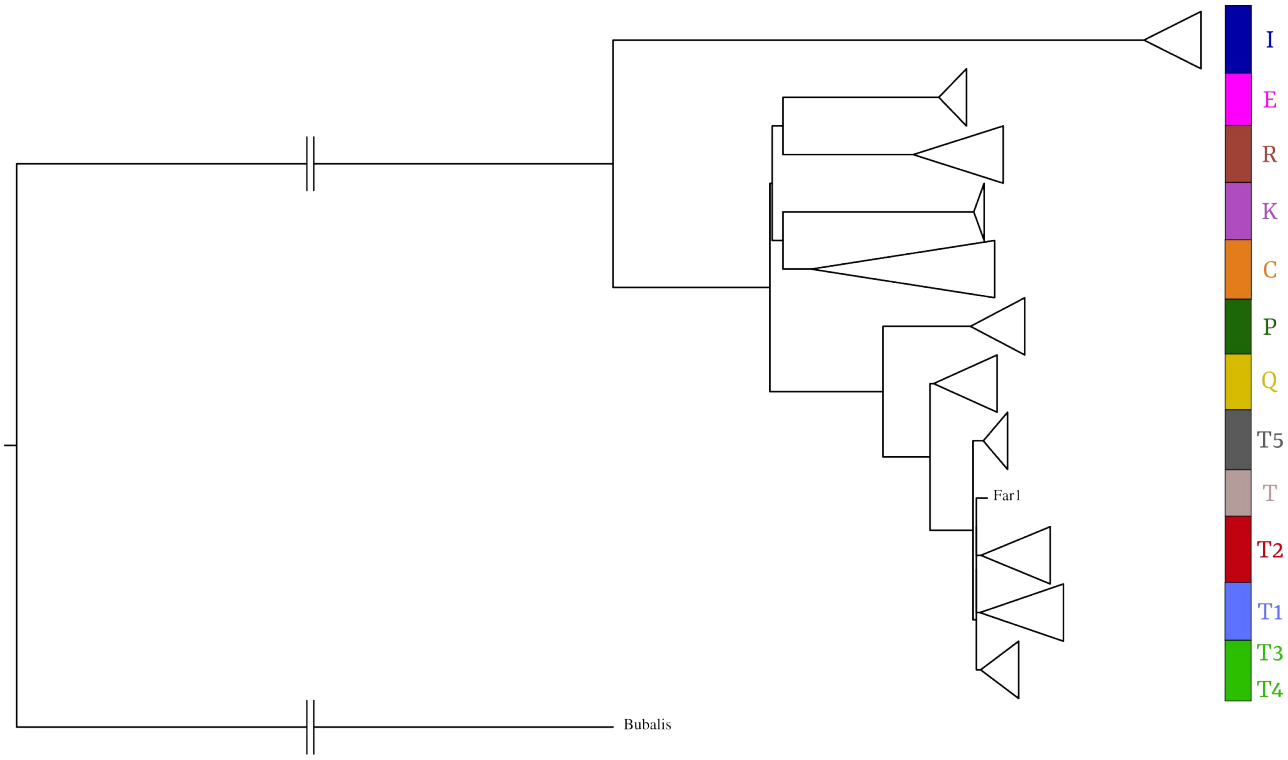

Fig. 14. Collapsed maximum likelihood tree of 215 mitochondrial genomes. The tree was computed on both wild and domesticated samples that had at least 90% of their genome covered at minimum 3X.

Due to the tree size, close ups of clusters are shown in Figure 15, allowing readable haplogroup identification for each sample. The figure shows indicines at the base of the tree, outgrouping every other cluster. Ancient domesticates from the new dataset can be found in every non-indicine haplogroup, except for K, C and T5. Only *Bos primigenius* individuals are represented by K and C haplogroups, T5 is only represented by two modern taurine samples. Interestingly, 12 individuals in P are *Bos primigenius* (described in Supplementary Table S1), but two *Bos taurus* are included in this cluster (Sha1, from Chalcolithic Ukraine, and Kok1, from Iron Age Uzbekistan). Two *Bos taurus* form the E cluster (Figure 15.A) and are individuals from the Caucasus and Turkmenistan. The Q cluster is divided in two branches: on one side are mainly wild samples, and on the other domesticates (Figure 15.B).

The number of samples in each group has been recorded in Table 10. The majority of the ancient samples fall in the T3 cluster (84 samples, Figure 15.D). T2 and Q are the next

most frequent haplogroups, with 29 and 24 samples respectively. P presents in 14 samples while E, R, K and C are represented by 2 to 5 samples each (Table 10).

Table 10. Haplogroup distribution in ancient samples. Haplogroups were determined according to the sample's position in the maximum likelihood tree.

	E	R	K	C	P	Q	T1	T2	T3-T4
Number of samples	2	2	3	5	14	24	16	29	85

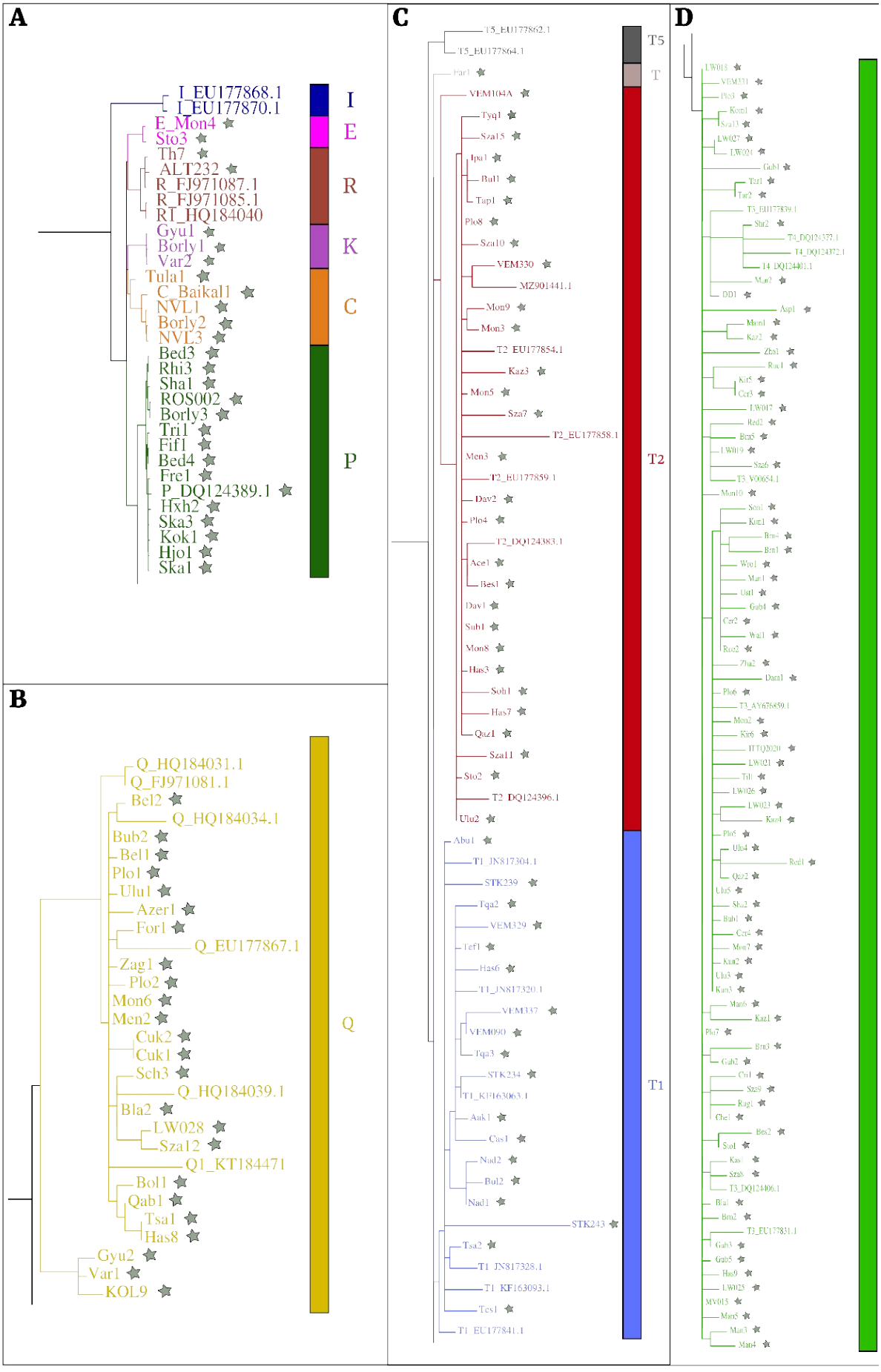
A
I_EU177868.1
I_EU177870.1
E_Mon4
Sto3
Th7
ALT232
R_FJ971087.1
R_FJ971085.1
RT_HQ184040
Gyu1
Borly1
Var2
Tula1
C_Baikal1
NVL1
Borly2
NVL3
Bed3
Rhi3
Sha1
ROS002
Borly3
Tri1
Fif1
Bed4
Fre1
P_DQ124389.1
Hxh2
Ska3
Kok1
Hjo1
Ska1
I
E
R
K
C
P
B
Q_HQ184031.1
Q_FJ971081.1
Bel2
Q_HQ184034.1
Bub2
Bel1
Plo1
Ulu1
Azer1
For1
Q_EU177867.1
Zag1
Plo2
Mon6
Men2
Cuk2
Cuk1
Sch3
Q_HQ184039.1
Bla2
LW028
Sza12
Q1_KT184471
Bol1
Qab1
Tsa1
Has8
Gyu2
Var1
KOL9
Q
C
T5_EU177862.1
T5_EU177864.1
For1
VEM104A
Tyq1
Sza15
Ipa1
Bul1
Tap1
Plo8
Sza10
VEM330
MZ901441.1
Mon9
Mon3
T2_EU177854.1
Kaz3
Mon5
Sza7
T2_EU177858.1
Men3
T2_EU177859.1
Dav2
Plo4
T2_DQ124383.1
Ace1
Bes1
Dav1
Suh1
Mon8
Has3
Soh1
Has7
Qaz1
Sza11
Sto2
T2_DQ124396.1
Ulu2
Abu1
T1_JN817304.1
STK239
Tqa2
VEM329
Tef1
Has6
T1_JN817320.1
VEM337
VEM090
Tqa3
STK234
T1_KF163063.1
Ank1
Cus1
Nad2
Bul2
Nad1
STK243
Tsa2
T1_JN817328.1
T1_KF163093.1
Tes1
T1_EU177841.1
T5
T
T2
T1
D
LW018
VEM331
Plo3
Kom1
Sza13
LW027
LW024
Gub1
Tar1
Tar2
T3_EU177839.1
Shr2
T4_DQ124377.1
T4_DQ124372.1
T4_DQ124401.1
Man2
DD1
Asp1
Mam1
Kaz2
Zha1
Ruc1
Kit5
Cer3
LW017
Red2
Brn5
LW019
Sza6
T3_V00654.1
Mon10
Soo1
Kom1
Brn4
Brn1
Wro1
Mar1
Ust1
Gub4
Cer2
Wal1
Ruc2
Zha2
Dara1
Plo6
T3_AY676859.1
Mon2
Kir6
ITTQ2020
LW021
Tri1
LW026
LW023
Kaz4
Plo5
Ulu4
Red1
Qaz2
Uli5
Sha2
Bub1
Cer4
Mon7
Kan2
Uli3
Kan3
Man6
Kaz1
Plo7
Brn3
Gub2
Cri1
Sza9
Rag1
Che1
Bes2
Sto1
Kan1
Sza8
T3_DQ124406.1
Bla1
Brn2
T3_EU177831.1
Gub3
Gub5
Has9
LW025
MV015
Man5
Man3
Man4

Fig. 15. Close-up view of each haplogroup cluster in the ML tree (Figure 14). Only clusters that contain ancient samples are represented. Modern samples are labelled with their haplogroup and NCBI accession number, ancient samples are indicated by a star-shaped symbol.

The tree, presented in Figures 14 and 15, illustrates the diversity of different haplogroups in ancient samples. This haplogroup representation seems to be linked to the sample's geographic origin. In Figure 16, the regions of Eurasia present in the ancient dataset are displayed, and range from North Western Europe to Central Asia and Africa.

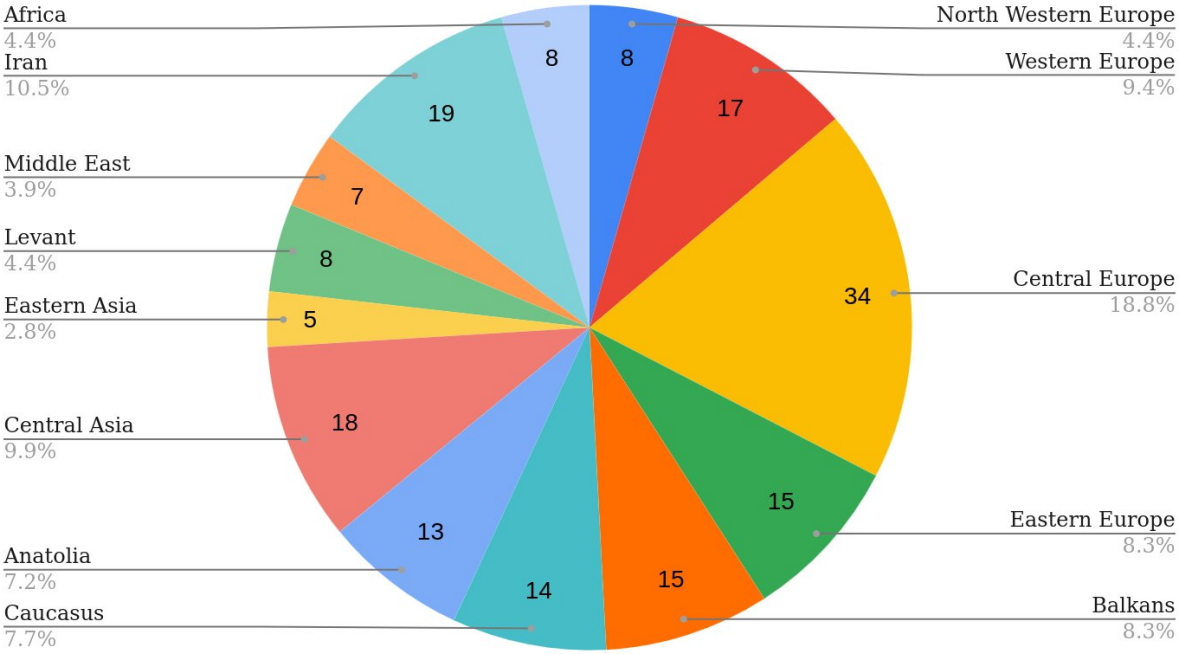

Fig. 16. Distribution of countries in the dataset of ancient samples used for the mitochondrial phylogeny.

Many samples come from Western and Central Europe. On Figure 17, we can see that these European samples are mostly T3 and P. However, T3 is not restricted to Europe, it is also found in a few samples from Western Asia: Anatolia, the Caucasus, Iran and Uzbekistan. T1 and T2 haplogroups are present in Central Europe, Anatolia, the Middle East, Iran, Turkmenistan and Africa. Numbers are higher in South-Western Asia (and Africa for T1) than in Europe though. Q is found in both Europe and Asia, from the Netherlands

to Anatolia, Iran, and Russia, while C, E and K, which are rare haplogroups in the dataset, are found in a couple of Asian countries only. C is found in the North of Kazakhstan and South of Western and Eastern Russia, E in the Caucasus and Turkmenistan, K in Armenia and Kazakhstan. R is found in an aurochs from Morocco and a domesticated taurine from Tunisia. The only T4 sample in the dataset comes from the North Caucasus.

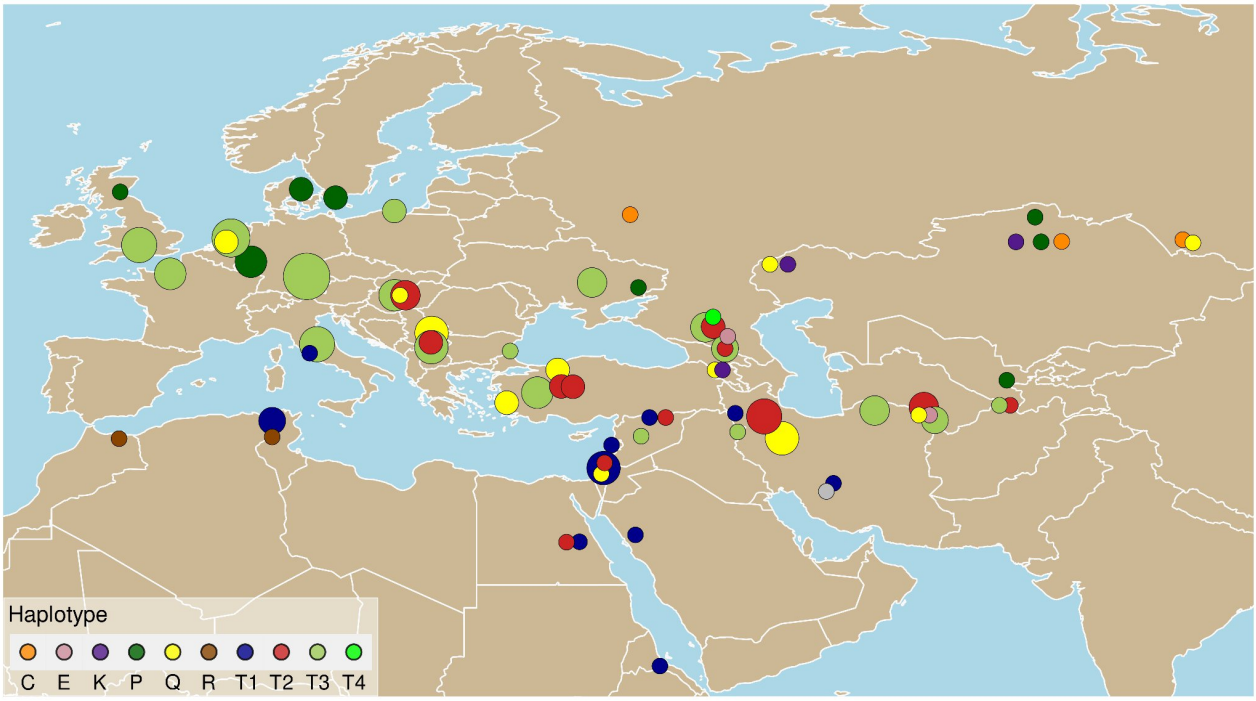

Fig. 17. Geographic distribution of haplogroups in ancient samples. Colors for each haplogroup correspond to the ones used in the ML tree. Bigger dots mean several samples in the same country share the same haplogroup. One sample in Iran is left in grey as it could not be placed in the tree further than T.

3.3.2.2. Haplogroup distribution through time

As seen earlier in Figure 17, not all haplogroups are present in every region of Eurasia and Africa. A second characteristic of this distribution is the time period. Figure 18 describes haplogroup distribution according to the age of the samples. The dataset was divided in three periods, those older than 5k years, those dating from 5 to 2.7k years and samples younger than 2.7k years.

Europe during the Neolithic is characterised by a majority of P and T3 in wild and domesticated samples respectively (with the exception of one *Bos taurus* in Ukraine which is P), with higher diversity in Serbia, which also presents T2 and Q haplogroups. In South Western Asia, variability is also high, with 6 haplogroups present: E, K, Q, T1, T2 and T3. None of these is predominant in the region. Russia and Kazakhstan are represented by wild samples with haplogroups C, K, Q and P. The only African sample is an aurochs with haplogroup R.

As time progresses to less than 5,000 years ago, T3 becomes predominant in Europe (Figure 18.B), which still harbours trace numbers of Q and T2 samples, although T2 is, as before, constrained to Central Europe. This haplogroup is more present in Anatolia, the Caucasus, the Middle East and reaches Egypt in Africa and Uzbekistan in Central Asia. T1 is another important haplogroup in the Levant and the Middle East, as well as in Egypt. Q is still present but in fewer numbers in the Middle East and Eastern Russia. T3 is also found in the Caucasus and in Iran, as before, as well as in one sample from Anatolia. One sample from the Caucasus is identified as a T4, the only example in this dataset. No samples belonging to the haplogroups C, E, K or R were found from this time period.

Lastly, later samples from Figure 18.C, the most recent period, show a strong predominance of the T3 haplogroup. Most samples are from Western and Central Europe, with a few from Eastern Europe, the Caucasus and one from Iraq. Two samples in the Caucasus and the Levant have the haplogroup T2, two others from Western Iran and the Levant are Q and T1 respectively. In Africa, 4 out of the 5 samples belong to the T1 haplogroup, the last one being R.

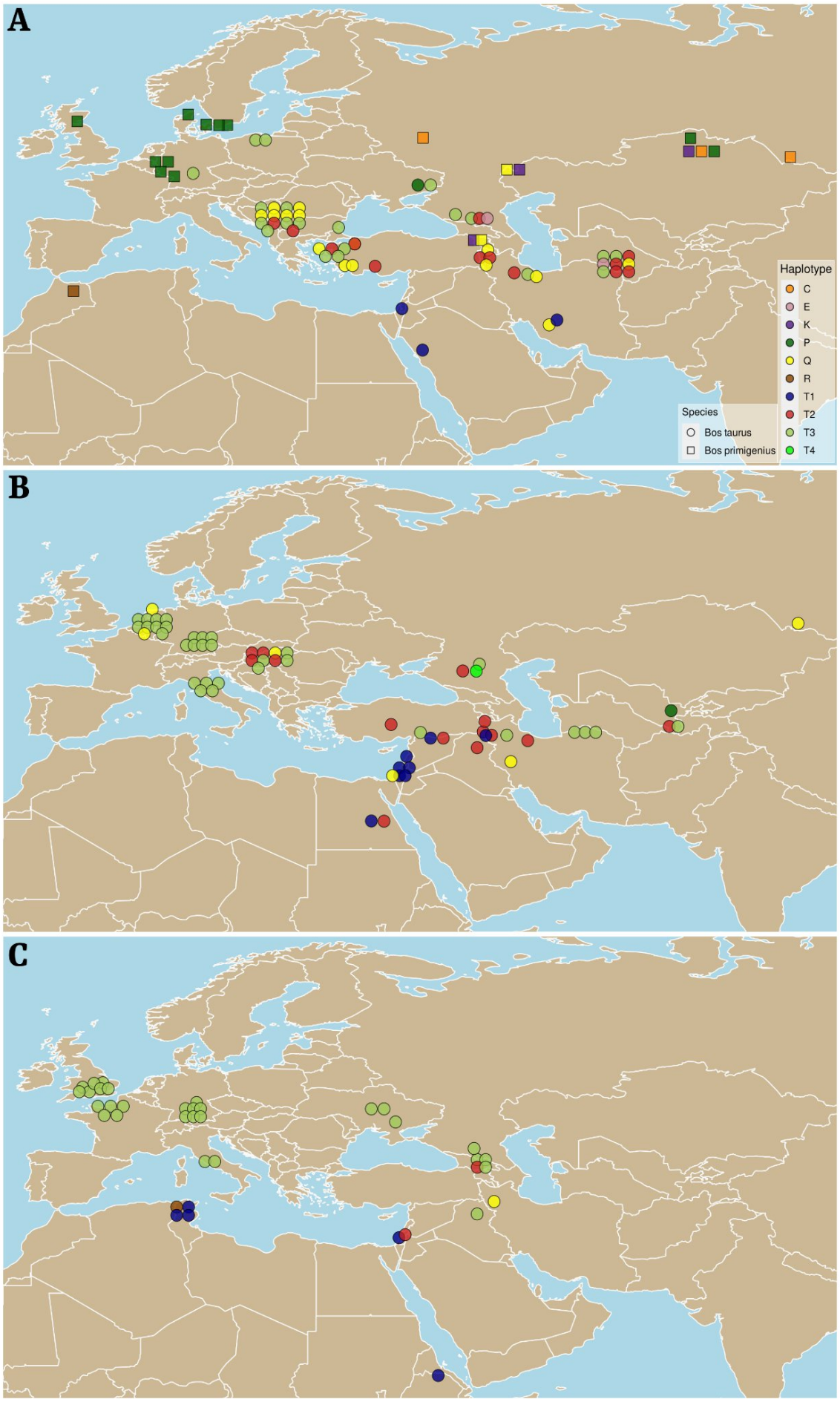
A
B
C
Haplotype
C
E
K
P
Q
R
T1
T2
T3
T4
Species
Bos taurus
Bos primigenius

Fig. 18. Haplogroup distribution according to three time periods. A) Samples older than 5,000 years. B) Samples dating from between 5,000 and 2,700 years ago. C) Samples younger than 2,700 years old.

3.3.3. A rare sample from Saudi Arabia

3.3.3.1. Capture output

This purportedly Neolithic sample from a rock shelter site (Aak1, Kennedy et al., 2023) was amplified with 7 PCRs, but one led to ambiguous results when processing the data through fastQ Screen, indicating possible contamination. This single PCR was discarded, while the 6 others were merged together, and their sequencing allowed the reconstruction of a mitogenome covered at 20.1X. In Figure 15.C, Aak1 is found located in the T1 cluster, among samples from the Levant, the Middle East, and Africa (Tunisia and Ethiopia).

A BEAST analysis was performed to confirm its location in the T1 cluster, as well as to check the archaeologically estimated age of this sample (~ 7,000 years old), since its radiocarbon dating failed.

3.3.3.2. BEAST phylogeny

The consensus tree resulting from the merge of 3 runs of BEAST is displayed in Figure 19. The software estimated the sample Aak1 to be aged (with a 95% confidence interval) between 6,756 and 7,280 years old. It places Aak1 in the T1 cluster, among modern T1 samples from Bonfiglio et al. (2012) (Appendix - Supplementary Table S3) and ancient samples from the Levant, Lebanon and Egypt.

Samples found to have haplogroup T2 stretch from the Balkans to Central Asia, including Anatolia, the Caucasus, the Levant, Iraq and Iran. T3 is spread from Western Europe to The Caucasus, and includes the Levant and Middle East. Q is present in the same areas as T2. Ancient samples that have haplogroup T1 come from Africa (Tunisia and Egypt), the Levant and Middle East. Today's domesticated cattle that have T1 are African

(Ethiopia, Egypt) and South-European (Portugal, Italy, France), with a few samples from the American continents (Mexico and Paraguay).

BEAST allowed the estimation of separation times between the main haplogroups Q, T3, T2 and finally T1. First, a separation between Q and the taurine haplogroups happened between 18,914 and 32,772 years ago, leading to Q on one side and T1-T2-T3 on the other. The basal divergence between T haplogroups represented is that between, T3 and the group T1+T2 diverged between 10,679 to 16,716 years ago, followed by the split between T1 and T2 10,285 to 15,384 years ago. These 95% posterior density ranges only barely overlap with the archaeological estimate for cattle domestication, ~10,500 BP (Vigne 2011; Zeder 2012). Coalescence estimates within the T1, T2 and T3 haplogroups (Table 11) have later ranges which have more plausible overlaps with domestic history.

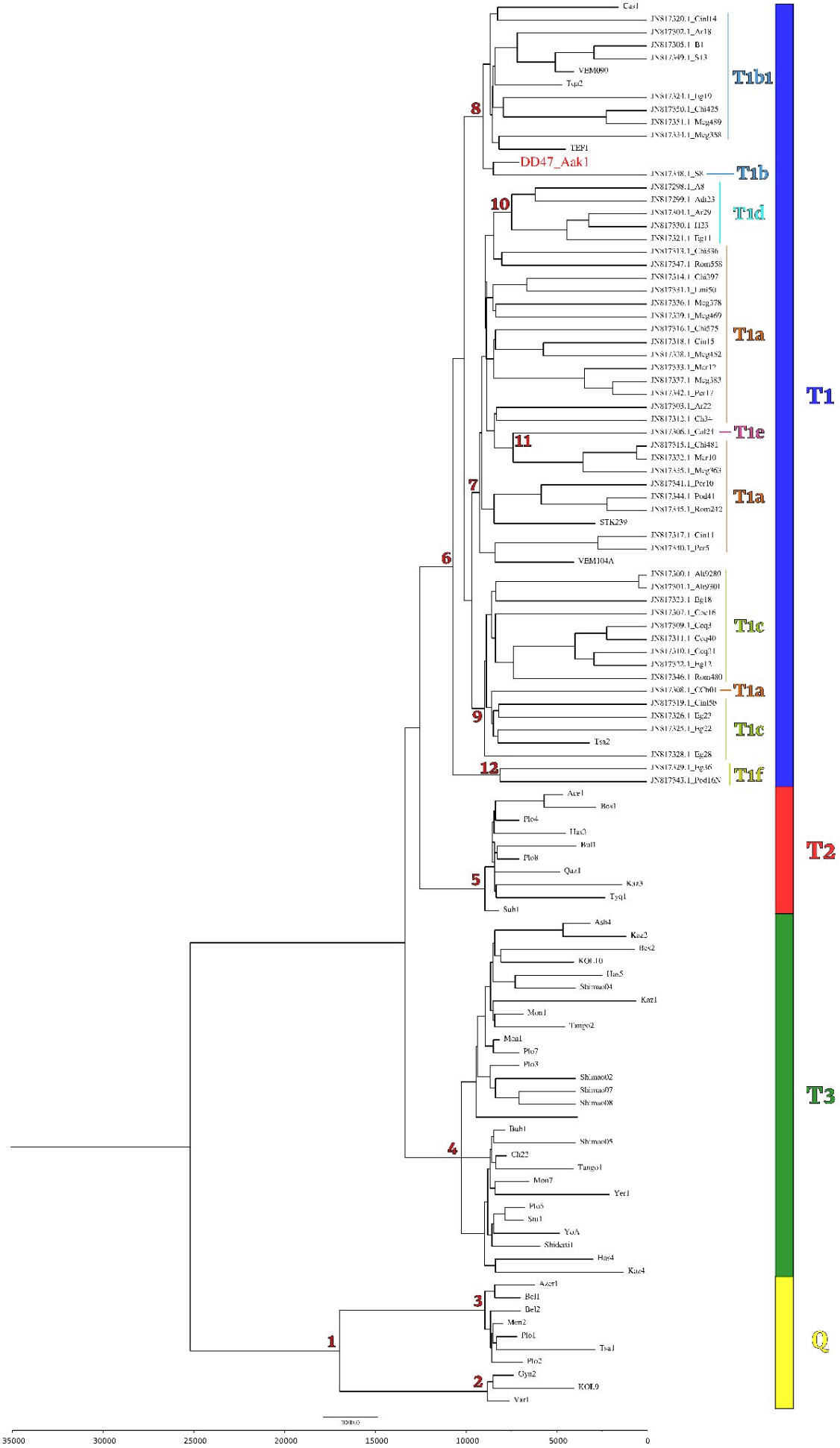
Cas1
JN817320.1_Cin114
JN817302.1_Ar18
JN817305.1_B1
JN817349.1_S13
VBM060
Tqu2
JN817324.1_Ig19
JN817350.1_Chi425
JN817351.1_Mcg489
JN817334.1_Mcg358
TEF1
DD47_Aak1
JN817348.1_S8
JN817298.1_A8
JN817299.1_Ad123
JN817304.1_Ar29
JN817330.1_I123
JN817321.1_Tg11
JN817313.1_Chi336
JN817347.1_Rom558
JN817314.1_Chi397
JN817331.1_mi50
JN817336.1_Mcg378
JN817339.1_Mcg469
JN817316.1_Chi575
JN817318.1_Cin15
JN817338.1_Mcg482
JN817333.1_Mcr17
JN817337.1_Mcg383
JN817342.1_Per17
JN817303.1_Ar22
JN817312.1_Chi3-
JN817306.1_Cal24
JN817315.1_Chi481
JN817332.1_Mcr10
JN817335.1_Mcg463
JN817341.1_Per10
JN817344.1_Pod41
JN817345.1_Rom242
STK239
JN817317.1_Cin11
JN817340.1_Per5
VBM104A
JN817300.1_Alt9280
JN817301.1_Alt9301
JN817323.1_Tg18
JN817307.1_Coc16
JN817309.1_Ceq4
JN817311.1_Ceq40
JN817310.1_Ceq31
JN817322.1_Tg13
JN817346.1_Rom480
JN817308.1_CCh01
JN817319.1_Cin15b
JN817326.1_Tg23
JN817325.1_Tg22
Tsa2
JN817328.1_Tg28
JN817329.1_Tg36
JN817343.1_Pod16N
Ace1
Bes1
Plo4
Has3
Bal1
Plo8
Qaz1
Kaz3
Tyq1
Suh1
Ash4
Kaz2
Bes2
KOL10
Has5
Shimao04
Kaz1
Mon1
Tango2
Mon1
Plo7
Plo3
Shimao02
Shimao07
Shimao08
Buh1
Shimao05
Ch22
Tango1
Mon7
Yer1
Plo5
Stu1
YoA
Shidert11
Has4
Kaz4
Azer1
Bel1
Bel2
Men2
Plo1
Tsa1
Plo2
Glyt2
KOL9
Var1
1
2
3
4
5
6
7
8
9
10
11
12
T1b1
T1b
T1d
T1a
T1e
T1a
T1c
T1a
T1c
T1f
T1
T2
T3
Q
35000
30000
25000
20000
15000
10000
5000
0

Fig. 19. BEAST tree with samples representing the haplogroups T1, T2, T3 and Q. Aak1 is indicated in red. Ancient samples were aligned with moderns (belonging to haplogroup T1, labelled with their NCBI accession number) published in Bonfiglio et al. (2012). Inside the T1 cluster, sub-haplogroups are indicated. The main divergence times are numbered on the tree and detailed in Table 11.

Table 11. Divergence times of main haplogroups (and T1 sub-haplogroups) estimated by BEAST. T1 sub-haplogroups are defined following Bonfiglio et al., 2012. Each node is labelled on the tree in Figure 19.

Haplogroup	Node label on tree	Beast divergence times (BP)
Q	1	12,482 - 22,374
Qwild	2	7,936 - 10,185
Qdom	3	8,073 - 10,170
T3	4	8,678 - 12,254
T2	5	8,165 - 10,011
T1	6	8,956 - 12,861
T1a	7	8,228 - 10,468
T1b	8	8,099 - 10,342
T1c	9	8,065 - 10,088
T1d	10	4,505 - 9,377
T1e	11	3,360 - 9,250
T1f	12	4,191 - 9,879

3.4. Discussion

3.4.1. Mitochondrial DNA is more easily extracted from ancient samples than nuclear data

The 106 mitogenomes extracted from whole genome sequencing gave at least 5.7X in mitochondrial coverage. The nuclear coverage for the least densely sequenced sample (Tar1) was 0.009X, barely enough for population genomics analyses. The mean coverages in mitochondrial vs. nuclear genomes, 72.0X vs. 0.527X (see Chapter 2 section 2.3.5 for the nuclear coverage), is markedly larger, probably reflecting higher copy number in bone cells of mtDNA. This shows that even if a sample seems too old or too degraded to hold preserved nuclear DNA, it may often still be salvaged for mitochondrial studies.

3.4.2. Haplogroups became structured over time

3.4.2.1. A reduced haplogroup diversity in younger samples

In Figure 18 one can see a visible reduction in the range of haplogroups observed in different regions occurring through time. There are a number of factors influencing this. First, the earlier samples include aurochsen with divergent mtDNA to domesticates. These reduce as the wild progenitor becomes rarer and eventually extinct. Second, there are likely founder effects in the Neolithic when cattle are first brought outside the Near East and subsequently as herds move into different regions. Third, although it is undoubtedly the case that initial herds were influenced by subsequent migrations and crossbreeding, including with aurochs, this may not have increased mtDNA diversity. This is because directional hybridisation can occur mainly (and most effectively) through males, altering the autosomal genome but leaving the maternal lineage relatively unchanged. This has been shown to be the case for aurochs introgression (Park et al., 2015, Upadhyay et al., 2017, Rossi et al., 2024) and with ancient zebu introgression in the ancient Near East (Verdugo et al., 2019). However, it is to be noted that there are lower frequency examples of divergent P and R mtDNA lineages in Europeans and Asians, a legacy of secondary maternal input (Achilli et al., 2009; Bonfiglio et al., 2010; Rossi et al., 2024). One domestic sample in this current dataset shows a P haplogroup also. Lastly, the effective population

size of mtDNA chromosomes is reduced, one quarter that of autosomal variation, leaving them more susceptible to diversity loss through genetic drift.

3.4.2.2. Specific haplogroups became representative of regions

In Europe the predominance of T3 in modern cattle has been noted before, first in Troy et al. (2001), and confirmed by many subsequent publications. This pattern is stronger in Northern Europe, with for example T1 present in Mediterranean countries possibly as a result of secondary migration from Africa but also probably including the influence of early movement (Cymbron et al., 1999; Bonfiglio et al., 2010; Colominas et al., 2015). This T3-focus might have been attributed to the spread and success of modern high production breeds - even via their influence in less productive cattle (rare breeds and village cattle), which would have occurred through their numerical predominance. However these results show that the T3 drift to numerical superiority has an earlier basis - already being strongly present in the west from 5,000 BP onwards.

In Africa modern African cattle are very strongly predominantly T1 (Bradley et al., 1996, Troy et al., 2001). This was initially taken as suggestive of a separate domestication. However, given that T1 has a coalescence range with T2 that includes the time of Near Eastern domestication (circa 10,500 BP) it is not possible to make an argument for this from mtDNA divergence. The estimated coalescence using BEAST of the different sublineages of T1 defined by Bonfiglio et al. (2012) at 8,956 - 12,861 BP suggests that all T1 may have had a single Near Eastern maternal ancestor within the domestication period ~10,500 years ago. Achilli et al. (2008) in the initial analysis of whole cattle mtDNA diversity used a calibrated mtDNA mutation rate to arrive at similar estimates. A further argument against an indigenous African origin for the T1 diversity that predominates there today comes from a single African aurochs and several early African domesticates which do not show this haplogroup, instead presenting with the highly divergent R haplogroup (Verdugo et al., 2019; Ginja et al., 2023). The additional data in this chapter also shows that T1 was present throughout the Levantine region of the ancient Near East, including close to the Upper Euphrates region of primary domestication in the Syrian Chalcolithic. This work also establishes the earliest Near Eastern occurrence of T1 in a bone from the rock shelter Al'Ula in Saudi Arabia which a Neolithic site dated to ~7,000 BP (Kennedy et al., 2023).

Here, although the bone was not directly datable using radiocarbon, its temporal provenance is confirmed using BEAST, which estimates its age at 6,756 and 7,280 years old. T1 has also been identified by PCR based ancient DNA work in Neolithic Spain (Colominas et al., 2015). It is most likely that T1 is a Near Eastern originating haplogroup that became established in the Levant before introduction to African and its establishment as the almost fixed variant there.

Table 9. Dataset of ancient samples sequenced for this chapter (n=127). The column lab work indicates other researchers involved in the drilling and/or DNA extraction (DD = Deborah Diquelou, MV = Marta Verdugo, CR = Conor Rossi, VEM = Victoria Elizabeth Mullin, CB = Catherine Butt)

Sample	Sex	Country	Site	Period	MT coverage (X)	Haplogroup	Capture	Lab work
Aak1	Female	Saudi Arabia	Al' Ula	Neolithic	20.130	T1	Yes	DD
Abu1	Male	Levant	Abu Gosh	MPPNB	7.470	T1		MV/DD
Ace1	Female	Anatolia	Acemhöyük	Bronze Age	35.330	T2		MV/DD
Asp1	Female	UK	Aspen Gardens, Stotfold	Medieval	117.750	T3		DD
Bla1	Male	Serbia	Blagotin	Neolithic	59.790	T3		MV/DD
Bla2	Male	Serbia	Blagotin	Early Neolithic	21.150	Q		MV/DD
Bol1	Female	Iran	Darre-ye Bolaghi	Neolithic	38.350	Q		CR/DD
Brn1	Female	France	Brionne	Roman Period	101.650	T3		DD
Brn2	Female	France	Brionne	Roman Period	130.550	T3		DD
Brn3	Female	France	Brionne	Roman Period	153.159	T3		DD
Brn4	Male	France	Brionne	Roman Period	77.290	T3		DD
Brn5	Female	France	Brionne	Roman Period	79.230	T3		DD
Bub1	Female	Serbia	Bubanj	Neolithic	42.760	T3		MV/DD
Bub2	Female	Serbia	Bubanj	Neolithic	25.800	Q		DD

Bul2	Female	Uzbekistan	Tilla Bulak	Late Bronze Age	84.820	T3		CR/DD
Cer2	Female	Germany	Cherry Tree Cave	Neolithic	21.940	T3		DD
Cer3	Female	Germany	Cherry Tree Cave	Late Neolithic?	78.350	T3		DD
Cer4	Male	Germany	Cherry Tree Cave	Iron Age	41.690	T3		DD
Che1	Male	Bulgaria	Chernomorets	Neolithic	162.370	T3		CR/DD
Cri1	Male	UK	Cribbs Causeway	Medieval	61.210	T3		DD
Cuk1	Male	Anatolia	Çukuriçi Höyük	Neolithic	22.140	Q	Yes	DD
Cuk2	Male	Anatolia	Çukuriçi Höyük	Neolithic	40.490	Q		DD
Dam1	Male	North Caucasus	Damba Kalaus 2	Bronze Age	35.080	T3	Yes	DD
Dav1	Female	Iran	Dava Goz	Chalcolithic	140.230	T2		DD
Dav2	Male	Iran	Dava Goz	Chalcolithic	270.930	T2		DD
DD001	Female	Hungary	Szazhalombatta-Sancalja	Bronze Age	15.140	T3	Yes	DD
Far1	Female	Iran	Fars	Neolithic	13.460	T		MV/DD
For1	Male	Iran	Tepe Forudgah	Bronze Age	61.230	Q		DD
Gub1	Female	Italy	Gubbio	Bronze Age	87.560	T3		DD
Gub2	Female	Italy	Gubbio	Bronze Age	56.580	T3		DD
Gub3	Female	Italy	Gubbio	Bronze Age	38.480	T3		DD
Gub4	Female	Italy	Gubbio	Bronze Age	66.170	T3		DD
Gub5	Male	Italy	Gubbio	Bronze Age	40.640	T3		DD
Has6	Female	Iran	Hasanlu	Iron Age	11.040	T1		DD
Has7	Female	Iran	Hasanlu	Early Bronze Age	51.540	T2		DD
Has8	Female	Iran	Hasanlu	Iron Age	77.420	Q		DD
Has9	Male	Iran	Hasanlu	Iron Age	79.180	T3		DD
Ipa1	Male	North Caucasus	Ipatovo 3	Bronze Age	10.190	T2		DD
ITTQ2020	Male	Italy	Tarquinia	Roman Period	63.830	T1	Yes	DD
Kas1	Female	Germany	Kaspauer	Bronze Age	19.130	T3		DD
Kir5	Female	Germany	Kirschbaumhöhle	Neolithic	29.690	T3		DD

Kir6	Female	Germany	Kirschbaumhöhle	Neolithic	54.760	T3		DD
Kom1	Male	North Caucasus	Komsomolec 1	Iron Age	110.740	T3		DD
Kun1	Female	Germany	Kunzig-Unternberg	Neolithic	9.090	T3		DD
Kun2	Female	Germany	Kunzig-Unternberg	Neolithic	85.660	T3		DD
Kun3	Female	Germany	Kunzig-Unternberg	Neolithic	80.370	T3		DD
LW017	Male	Netherlands	Mienakker	Chalcolithic	8.750	T3		VEM/CB/DD
LW018	Male	Netherlands	Mienakker	Chalcolithic	10.450	T3		VEM/CB/DD
LW019	Female	Netherlands	Mienakker	Chalcolithic	144.250	T3		VEM/CB/DD
LW021	Male	Netherlands	Mienakker	Chalcolithic	203.330	T3		VEM/CB/DD
LW023	Female	Netherlands	Zeewijk	Chalcolithic	125.800	T3		VEM/CB/DD
LW024	Male	Netherlands	Zeewijk	Chalcolithic	113.640	T3		VEM/CB/DD
LW025	Female	Netherlands	Zeewijk	Chalcolithic	145.830	T3		VEM/CB/DD
LW026	Male	Netherlands	Zeewijk	Chalcolithic	11.730	T3		VEM/CB/DD
LW027	Female	Netherlands	Zeewijk	Chalcolithic	7.480	T3		VEM/CB/DD
LW028	Female	Netherlands	Zeewijk	Chalcolithic	7.540	Q		VEM/CB/DD
Mam1	Female	Ukraine	Mamai-Gora	Iron Age	67.930	T3		DD
Man1	Male	Germany	Manching	Iron Age	29.010	T3		DD
Man2	Male	Germany	Manching	Iron Age	160.220	T3		DD
Man3	Female	Germany	Manching	Iron Age	52.120	T3		DD
Man4	Female	Germany	Manching	Iron Age	66.530	T3		DD
Man5	Male	Germany	Manching	Iron Age	44.480	T3		DD
Man6	Female	Germany	Manching	Iron Age	95.370	T3		DD
Men3	Male	Anatolia	Mentese	Neolithic	57.770	T2		DD
Mon10	Female	Turkmenistan	Monjukli Depe	Chalcolithic	127.740	T3		CR/DD
Mon2	Female	Turkmenistan	Monjukli Depe	Chalcolithic	10.610	T3		DD
Mon3	Female	Turkmenistan	Monjukli Depe	Chalcolithic	10.680	T2		DD
Mon4	Female	Turkmenistan	Monjukli Depe	Chalcolithic		E		CR/DD
Mon5	Female	Turkmenistan	Monjukli Depe	Chalcolithic	35.910	T2		CR/DD
Mon6	Female	Turkmenistan	Monjukli Depe	Chalcolithic	53.670	Q		CR/DD
Mon7	Female	Turkmenistan	Monjukli Depe	Chalcolithic		T3		CR/DD
Mon8	Male	Turkmenistan	Monjukli Depe	Chalcolithic	42.280	T2		CR/DD

Mon9	Male	Turkmenistan	Monjukli Depe	Chalcolithic	42.320	T2		CR/DD
MV015	Female	Anatolia	Barcin Hoyuk	Neolithic	32.560	T3	Yes	MV/DD
Nad1	Female	Iran	Naderi Tepe	Early Bronze Age	50.230	T3		DD
Nad2	Male	Iran	Naderi Tepe	Bronze Age	56.250	T3		DD
Plo6	Male	Serbia	Plocnik	Neolithic	37.910	T3		MV/DD
Qab1	Male	Iran	Tepe Qabrestan	Chalcolithic	30.940	Q	Yes	DD
Qaz2	Male	Iran	Alou	Chalcolithic	77.400	T3		DD
Red1	Male	UK	Redcliffe	Medieval	89.540	T3		DD
Red2	Male	UK	Redcliffe	Medieval	126.580	T3		DD
Ruc1	Male	Poland	Rzucewo	Neolithic	12.830	T3	Yes	DD
Ruc2	Female	Poland	Rzucewo	Neolithic	52.780	T3		DD
Rug1	Female	UK	Rugby SUE16	Medieval	231.530	T3		DD
Sch3	Male	Netherlands	Schipluiden	Neolithic	24.590	Q		MV/DD
Sha1	Male	Ukraine	Shakter	Chalcolithic	119.720	P		DD
Sha2	Male	Ukraine	Shakter	Chalcolithic	77.480	T3		DD
Shr2	Female	North Caucasus	Sharakhalsun 3	Bronze Age	40.850	T4		DD
Soh1	Male	Iran	Soha Chai Tepe	Chalcolithic	54.660	T2		DD
Son1	Female	Germany	Sonnenschacht cave	Neolithic	158.100	T3	Yes	DD
Sto1	Male	North Caucasus	Stoderovskoe 1	Bronze Age	65.000	T3		DD
Sto2	Male	North Caucasus	Stoderovskoe 1	Bronze Age	92.210	T2	Yes	DD
Sto3	Female	North Caucasus	Stoderovskoe 1	Bronze Age	126.520	E	Yes	DD
Sza10	Male	Hungary	Szazhalombatta-Sancalja	Early Bronze Age	58.070	T2		DD
Sza11	Female	Hungary	Szazhalombatta-Foldvar	Early Bronze Age	30.610	T2		DD
Sza12	Male	Hungary	Szazhalombatta-Foldvar	Bronze Age	187.790	Q	Yes	DD
Sza13	Male	Hungary	Szazhalombatta-Foldvar	Bronze Age	590.910	T3		DD

Sza15	Male	Hungary	Szazhalombatta-Foldvar	Middle Bronze Age	9.990	T2		DD
Sza6	Male	Hungary	Szazhalombatta-Foldvar	Early Bronze Age	218.330	T3		DD
Sza7	Female	Hungary	Szazhalombatta-Foldvar	Middle Bronze Age	69.230	T2		DD
Sza8	Female	Hungary	Szazhalombatta-Foldvar	Middle Bronze Age	129.360	T3		DD
Sza9	Female	Hungary	Szazhalombatta-Sancalja	Early Bronze Age	162.360	T3		DD
Tap1	Male	Iran	Tappeh Hesar Markazi	Bronze Age	108.030	T2		DD
Tar1	Male	Italy	Tarquinia	Antiquity	5.710	T3		DD
Tar2	Male	Italy	Tarquinia	Iron Age	36.680	T3	Yes	DD
Tef1	Female	Lebanon	Tell Fadous	Bronze Age	42.040	T1		CR/DD
Tes1	Male	Levant	Tel-el-Safi	Bronze Age	7.540	T1		DD
Til1	Female	Anatolia	Tilbeshar	Chalcolithic	10.480	T3		DD
Tqa2	Female	Levant	Tel es-Qashish	Bronze Age	99.020	T1		MV/DD
Tqa3	Female	Levant	Tel es-Qashish	Bronze Age	15.550	T1		MV/DD
Ulu1	Female	Anatolia	Ulucak	Neolithic	24.540	Q		DD
Ulu2	Female	Anatolia	Ulucak	Neolithic	67.390	T2		DD
Ulu3	Male	Anatolia	Ulucak	Neolithic	33.620	T3		DD
Ulu4	Male	Anatolia	Ulucak	Neolithic	70.450	T3		DD
Ulu5	Male	Anatolia	Ulucak	Neolithic	30.810	T3		DD
Ust1	Male	North Caucasus	Ust Dzheguta	Bronze Age	19.460	T3	Yes	DD
VEM090		Egypt	Thoth Hill	Bronze Age	133.620	T1		VEM/DD
VEM104A		Egypt	Deir el Bahri	Bronze Age	14.800	T2		VEM/DD
VEM329		Syria	Wakarida	Recent	80.510	T1		VEM/DD
VEM330		Syria	Tell Chuera	Bronze Age	148.240	T2		VEM/DD
VEM331		Syria	Tell Chuera	Bronze Age	58.760	T3		VEM/DD
VEM337		Ethiopia	Tell Sheikh Hassan	Chalcolithic	8.230	T1		VEM/DD
Wal1	Female	UK	Wallingford	Medieval	77.380	T3		DD
Wro1	Female	UK	Wroughton	Iron Age	111.340	T3		DD

Zag1	Male	Iran	Zagheh Abyek	Chalcolithic	85.530	Q		DD
Zha1	Male	Ukraine	Zhabotin Settlement	Iron Age	73.350	T3		DD
Zha2	Male	Ukraine	Zhabotin Settlement	Iron Age	9.600	T3		DD

4. A window into the structure and history of past populations of domesticated cattle

4.1. Introduction

For more than ten thousand years, humans have been raising animals, adding to the development of complex societies and the production of various goods such as meat, dairy, leather and wool. Larger animals are also a good source of labour, whether it be for help in farms or for transportation. This phenomenon was described as the Neolithic Revolution due to its impact on human societies. However, the importance of livestock in human communities did not decrease over time. On the contrary, today, domesticated cattle can be found all over the globe, and have reached an estimated number of 1.5 billion individuals, according to the Food and Agriculture Organization of the United Nations (Figure 20, Wint and Robinson, 2007, Gilbert et al., 2018), far from the estimated 80 females that comprised to the first domesticated cattle (Bollongino et al., 2012). This long shared history with humans led to a deeper relationship than that between hunters and their food sources. Different cultures have developed a strong connection with cattle, often relying on them for economy, or including them in religious beliefs (Alders et al., 2021); they have been proven to be good for mental health, helping in behavioural therapies or assisting patients with dementia through animal-assisted therapy (Hassink et al., 2017). This is why cattle have been widely studied, to understand their evolutionary history, and find ways to improve their management and production.

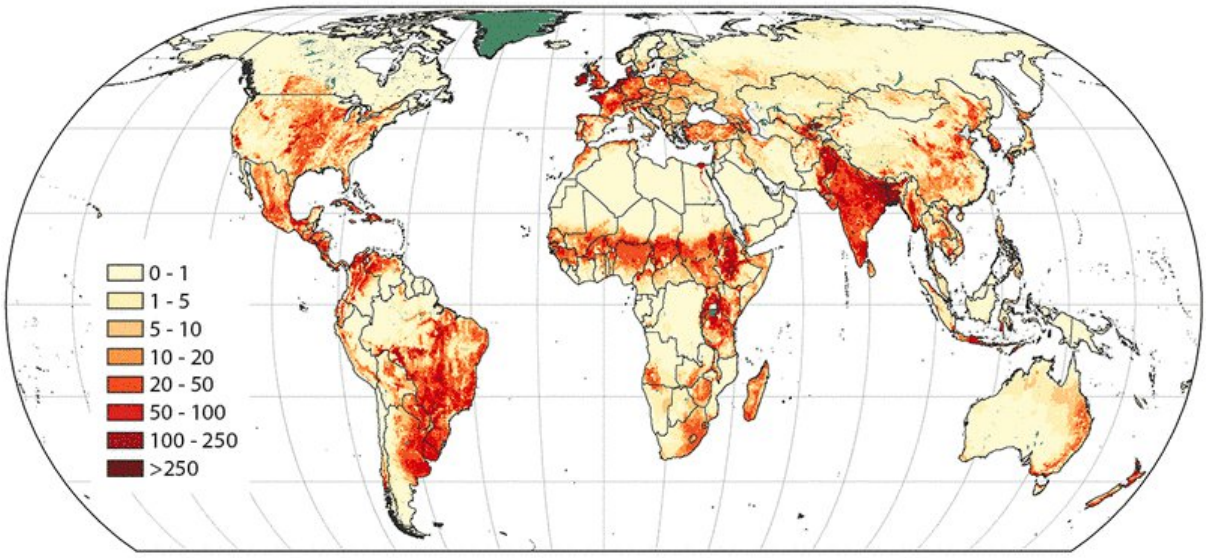

Fig. 20. Number of cattle per square kilometre in 2015. The map was taken from the FAO website and shows global cattle density using a color gradient from white (low) to dark red (high). (https://www.fao.org/livestock-systems/global-distributions/cattle/en/)

Domestication started in the Fertile Crescent, in the case of *Bos taurus*, and subsequently spread to neighbouring regions. The wild progenitor was *Bos primigenius*, a large, strong and violent animal that was present in Europe, Asia and Africa (Rossi et al., 2024). Co-existence between species in the same area has permitted admixture between them. If we add to this relationship a third participant, indicine cattle *Bos indicus* domesticated 8.5k years ago in the Indus valley, we start to see a glimpse of the complex history of taurine cattle, which is further enhanced by the involvement of human management regarding breeding and mobility.

Indeed, major human migrations helped shape modern cattle population structure. Nowadays, the main mitochondrial haplogroups are differentially represented in specific continents (see Chapter 3). Humans have spread cattle throughout the world and managed their breeding in such ways that we can see in their genomes traces of admixture. Hence, massive indicine introgression throughout the Near East was detected ~4k years ago (Verdugo et al., 2019) possibly as a way to help taurine herds to adapt to a changing more arid environment. Wild introgression can also be detected. Local aurochsen in Eurasia and Africa managed to breed with domesticated herds: although it is uncertain if this was facilitated by herders or resulted despite their efforts to keep wild cattle away.

Similarly to *Bos taurus*, *Bos primigenius* have geographical structure, helping us keep track of genetic flow from wild into domesticated animals. One particular gene-flow is from Eastern aurochsen, a wild population inhabiting North and Central Asia, and characterised by a mitochondrial haplogroup C. During the Bronze Age, a human population from the Steppe, the Yamnaya, undertook a migration towards Eastern Europe, influencing European human genomics (Haak et al., 2015).

In this chapter, we use allele frequency-based analyses to investigate gene-flow from Eastern aurochsen, as well as from European aurochsen and indicine cattle. The goal is to characterise ancient European cattle and model their structure, which we hypothesise to be a mixture of domesticated and different wild animals. To do so, we use pseudohaploid data in order to compensate for low coverage leading to the loss of allelic information. The data is then processed through different analyses adapted to the study of ancient DNA (laser PCA, D-statistics, admixture, qpADM).

4.2. Material and Methods

4.2.1. Samples

373 ancient samples encompassing published samples, previously sequenced and newly sequenced samples were analysed in this chapter via different methods. The list of samples and the analyses they are included in are informed in Supplementary Table S4.

4.2.2. Projection PCA

Procrustes projections were performed using LASER v2.04 (Wang et al., 2014 and 2015). Two reference PCAs were made using datasets of either both modern breeds of *Bos taurus* and *Bos indicus* or *Bos taurus* only, representing Europe, Asia and Africa. Ancient samples were processed with samtools v1.13 (Li et al., 2009) to produce pileup files, using the command mpileup with parameters:

```
-q 30 -Q 20
```

Projections of ancient samples on the modern reference PCA were performed for 100 runs, and the average result was plotted using RStudio v2023.03.0+386 (Posit Team, 2023).

4.2.3. D-statistics

D-statistics were computed on all chromosomes using ANGSD v0.941 (Korneliussen et al., 2014), using a water buffalo as an outgroup with parameters:

```
-doAbbababa 1 -doCounts 1 -rmTrans 1 -trim 4 -minMapQ 30 -minQ 20
```

Three tests were performed, one to detect indicine introgression in taurine cattle, another one to assess wild European introgression and the last one to assess wild Eastern introgression. A modern individual of Sahiwal breed, Sha3b, was used to represent *Bos indicus* in the first test. Wild aurochs from Germany and Eastern Russia, Bed3 (Erven et al., 2024) and Baikal1 (Rossi et al., 2024), were used for the second and third tests respectively. Results were filtered to keep samples that had at least 500 ABBA+BABA sites, a bigger value than Verdugo et al. (2019) who used 200 sites, as to be more stringent and reduce bias due to low coverage. Final results were plotted with RStudio v2023.03.0+386.

4.2.4. Pseudohaploid dataset

To overcome the issue ancient DNA poses regarding missing data, pseudohaploid genomes were generated for samples of the ancient dataset (Supplementary Table S4). To do so, a list of SNPs published in Erven et al. (2024) was used as input for ANGSD v.0.936, and called in the ancient dataset using the following parameters:

```
-doHaploCall 1 -doCounts 1 -doMajorMinor 3 -dumpCounts 1 -C 50 -
uniqueOnly 1 -remove_bads 1 -trim 5 -minMapQ 30 -minQ 25
```

The sites called were then compared with those from the modern reference dataset to make sure they are correct. Sample files were converted into tfam/tped files using ANGSD's haploToPlink command, and then converted into binary files thanks to PLINK v.1.9 (Purcell et al., 2007). Transitions were filtered out to remove the risk of DNA damage induced errors.

With PLINK, the dataset was filtered for a minimum allele frequency of 2.5% (`--maf 0.025`). The usual parameter for missingness is 20%, but this threshold resulted in very few SNPs left (501). Testing a 25% missingness led to only 10,455 SNPs kept. Finally, using a threshold of 30% (`--geno 0.3`) allowed a final SNP number of 367,817. Samples that had less than 50 thousands sites were removed and the dataset was pruned for linkage disequilibrium (`--indep-pairwise 50 5 0.5`).

4.2.4.1. Admixture

Analysis of ancestries in the pseudohaploid dataset was performed using ADMIXTURE v1.3 (Alexander et al., 2009), after testing a range of K from 3 to 8, with a thousand bootstrap replicates (`-B1000`). Admixture plots for the best K value were generated with Rstudio v2023.03.0+386.

4.2.4.2. Admixtools2

4.2.4.2.1. F4 ratio

In order to model samples as resulting from admixture between domesticated and wild aurochs, f4-ratios were computed using the admixtools2 R package (Maier et al., 2023). For each ancient sample, the ratio was calculated under the form f4(1, 2; 3, 4)/f4(1, 2; 5, 4), where:

- 1 = Tula1, an aurochs from Western Russia;
- 2 = a water buffalo as outgroup;
- 3 = samples from the pseudohaploid dataset;
- 4 = Sub1, a Neolithic Anatolian domesticates;
- 5 = Baikal1, an aurochs from Eastern Russia.

Baikal1 is used in various analyses to assess Eastern introgression, Tula1 is used here as a sister-taxa of Baikal1. Sub1 is a *Bos taurus* sample used as reference for domesticates in the calculation. All chromosomes were taken into account, instead of 22 by default, by setting the parameter `auto_only = FALSE`.

4.2.4.2.2. qpAdm

This other admixtools2' tool was used to model European samples as a mixture of domesticates, wild Europeans and wild Easterns. To do so, several samples were classified as "left populations" and "right populations" in order to test introgression from the left into a "target" population, while the right populations act as outgroups.

The left populations were set as Neolithic Anatolian domesticates (samples Men1, Men2 and Sub1), European aurochs (Bed3, Bed4, Fif1, Hjo1, Hxh2, Rhi3, Ska1, Ska3, Tri1, Zea1, Zea2) and Eastern aurochs (Baikal1, Tula1). Every sample of the ancient pseudohaploid dataset is rotated in the target population.

The command was run with parameters `auto_only = FALSE` and `allsnps = FALSE`. This allows the use of all chromosomes (instead of 22 by default) with the first parameter, the second forcing the f4 statistics to be calculated on SNPs present in every population for more accurate comparisons between them. Results were filtered to remove non-significant samples before plotting.

4.3. Results

4.3.1. Principal Component Analysis (PCA)

Two principal components analyses were generated using different sample cohorts. The first one, that includes both *Bos primigenius* and modern *Bos indicus* samples, used a reference PCA of modern samples (taurine and indicine) based on 2,841,344 markers onto

which ancient individuals were projected. The second one, focused purely on *Bos taurus* individuals, was based on 2,177,620 variable sites. Out of the 373 ancient samples considered (Supplementary Table S4), 9 samples had less than 2,000 sites after SNP calling, and were thus left out of both PCAs : Abu1, Abu2, Dan1, Gil1, Nah1, Tda1, Tmq1, Tmq2 and Tqa1. Not counting aurochs genomes, the second PCA contains 324 taurine samples.

4.3.1.1. Projection on a modern reference dataset of both taurine and indicine cattle

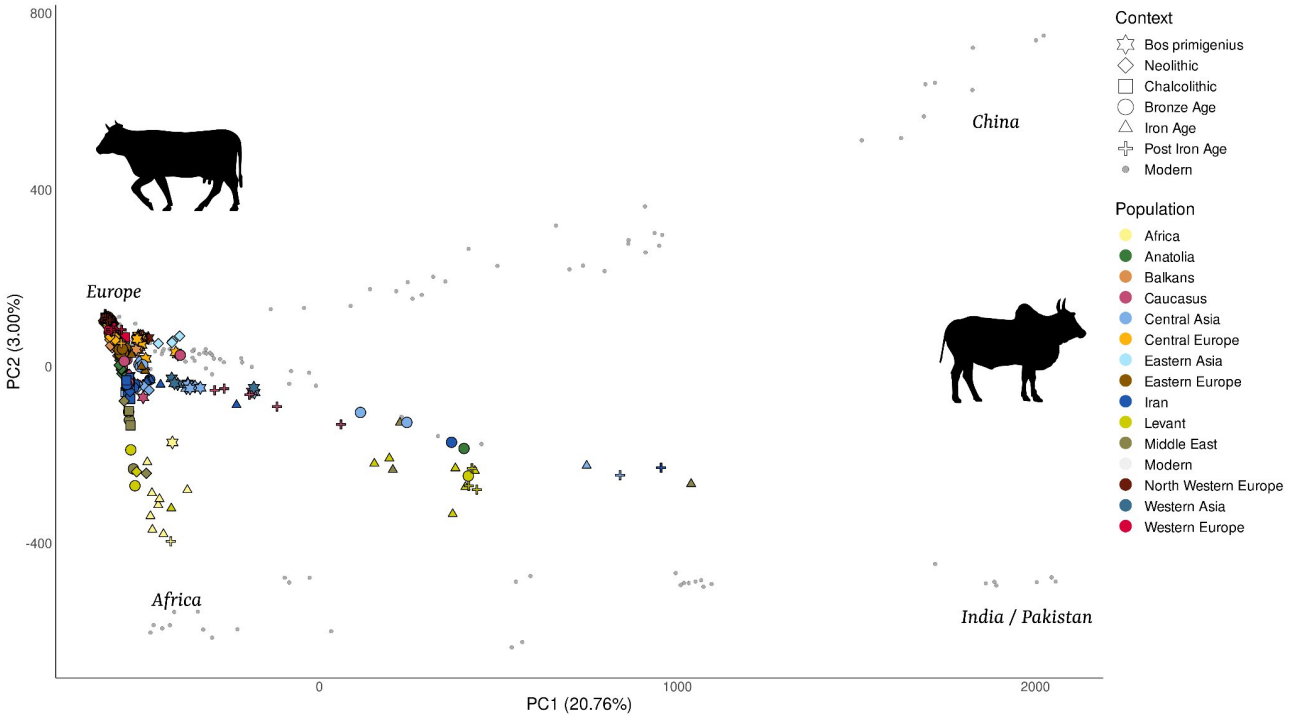

Fig. 21. Reference PCA of 171 modern taurine and indicine cattle and procrustes projection of 364 ancient samples. Modern samples are presented by smaller grey circles while ancient samples projected on the reference PCA are coloured according to geographic origin and shaped according to time context.

The modern samples used to plot the PCA in Figure 21 separate according to species and region of origin (grey symbols). On the left side of the figure are modern taurine samples, while modern indicines are located on the right. Hybrids plot in between. PC1 separates the two subspecies and PC2 separates populations according to their geographical origins; a range of regional samples from northern to southern provenance plot on PC2 from top to bottom. *Bos indicus* are thus separated in two clusters: Chinese zebu, on the top right corner of Figure 21, and Indian/Pakistan zebus on the bottom right. *Bos taurus* cattle are ranged vertically on the left, between those from Europe to those from Africa on the bottom left. PC values of the modern dataset are listed in Supplementary Table S5.

Ancient samples were projected onto the scaffold of the modern plot, and with these the same trend of geographical distribution can be observed. A close-up view of the full ancient dataset is shown on Figure 22.

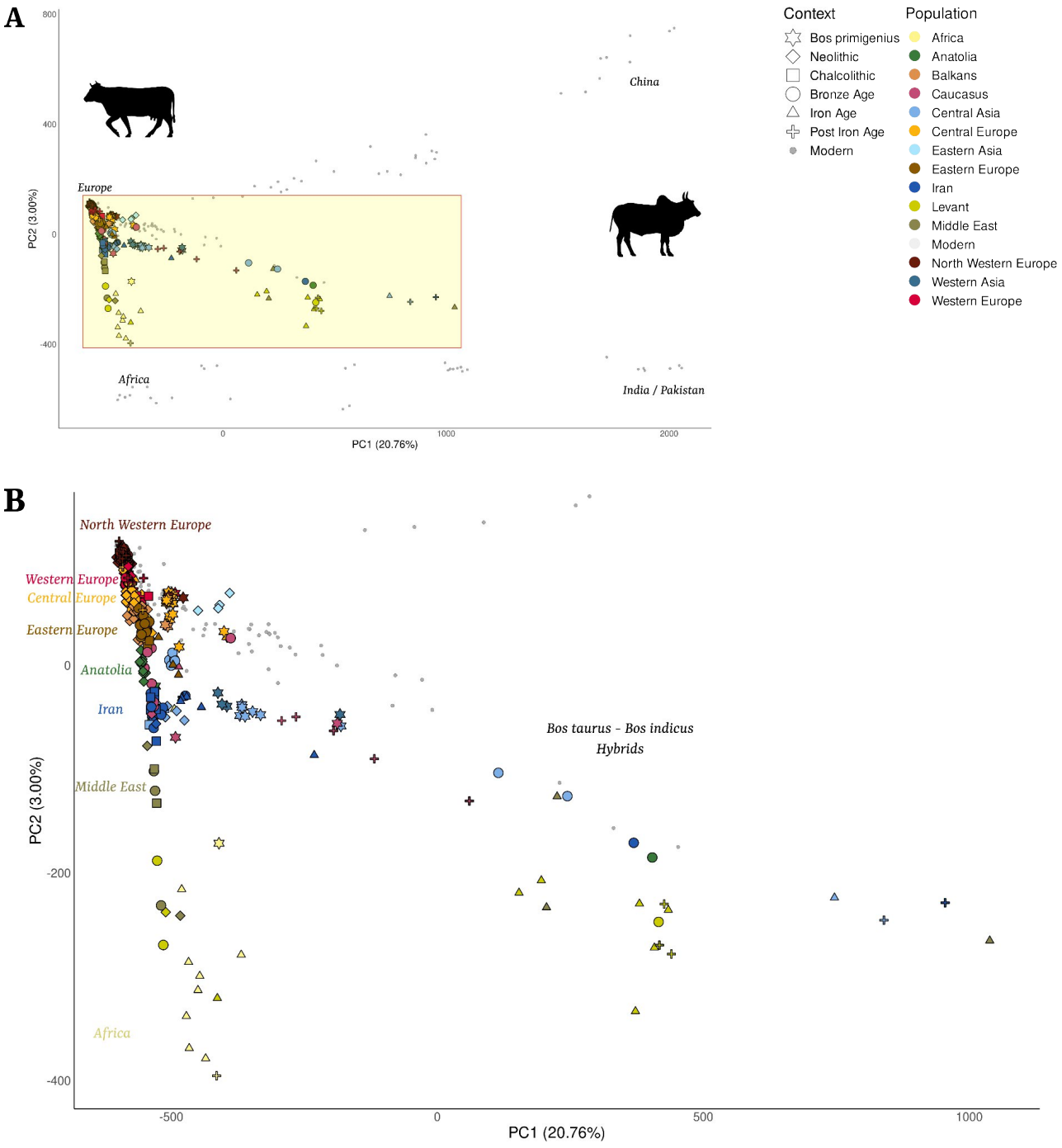

Fig. 22. Close-up view of the 364 ancient samples projected on the reference PCA. A) Figure 21 with a red rectangle surrounding the area that was zoomed in for panel B. B) Close-up view of this area showing every ancient sample on Figure 21.

The ancient samples on the PCA depicted in Figure 22 are spread along both axes. PC2 separates depending on the country of origin, as has been seen for modern samples. North-Western European samples plot at the top left corner of the PCA followed successively by those from Western Europe, Central Europe, the Balkans and Eastern Europe. These European areas, even though distinct, are clustered together on the top left corner. Next, from the Anatolian cluster, samples spread beneath in clusters from Iran, the Caucasus, the Middle East, the Levant and Africa. A subset of Central Asian, Iranian, Middle Eastern and Levantine individuals, are drawn to the right towards the indicine pole, behaving similarly to modern hybrids.

4.3.1.2. Projection on a taurine only reference dataset

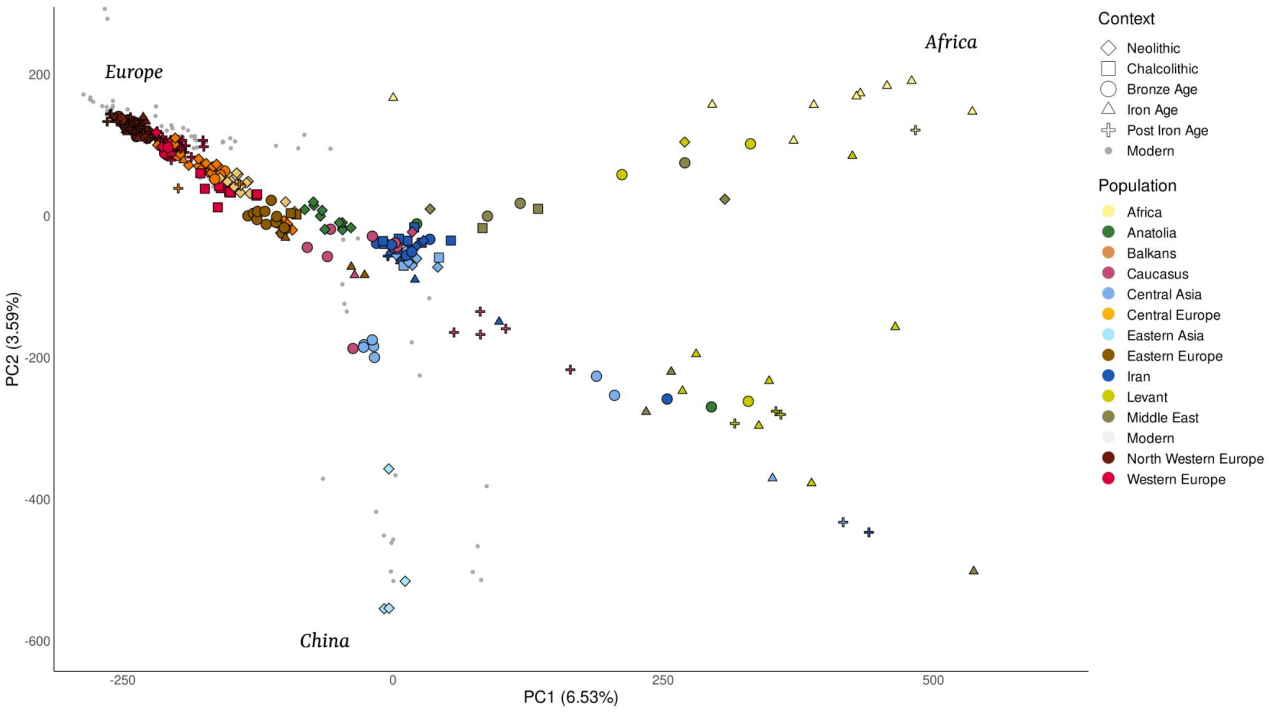

Fig. 23. PCA of 100 modern taurine samples and procrustes projection of 324 ancient samples. Modern samples are represented by smaller grey circles. Ancient samples are represented by time context via shape and coloured by their geographical origin.

To examine ancient variation within taurine cattle more closely, samples were projected on a reduced modern PCA without indicine genomes. For Figure 23, a modern dataset of *Bos taurus* samples was used to generate a PCA (see Supplementary Table S6)

and ancient samples (domesticates only) were projected onto it. The ancients plot on three axes, i.e. in the direction of modern samples from three continents: towards Europe at the top left, Africa on the top right and Asia at the bottom left. Ancient samples from these continents plot in concordance with their region.

Ancient European samples at the top left corner (Figure 23) form a cline from North-Western Europe to Eastern Europe with samples from Anatolia, Caucasus, Iran and Turkmenistan plotted centrally. On the axis towards modern Africa are found samples from Syria, the Levant and Tunisia. The axis towards Asian samples at the bottom of the figure features samples from Neolithic China, as well as a few samples from Central Asia and the Caucasus at the beginning of this trend. Several samples from the wider Near East (Levant as well as Anatolia, Iraq, Iran and Uzbekistan) plot on a separate diagonal stretching towards the bottom right. These samples are more recent than those from the same regions that plot centrally or toward Africa, and date at least from the Bronze Age. These individuals correspond to those that are drawn towards the right hand of Figure 22, presumably because of *Bos indicus* introgression.

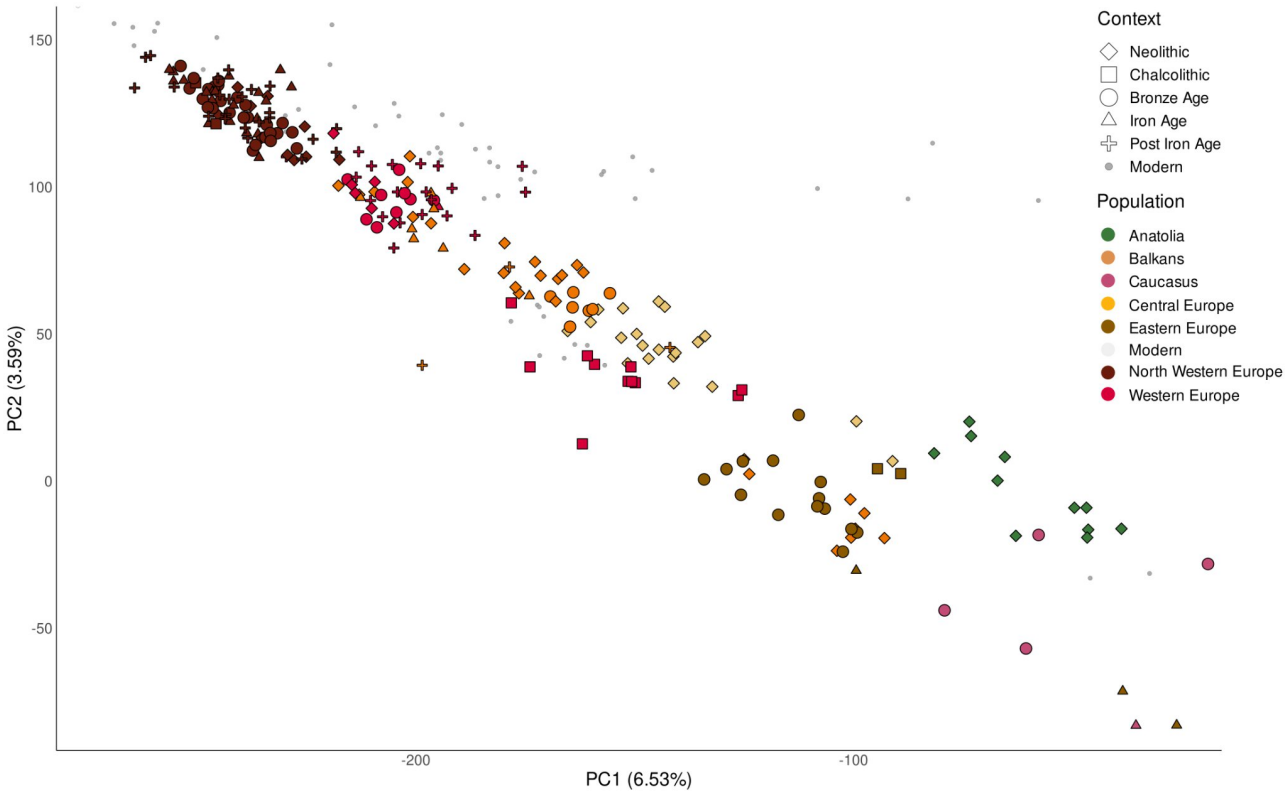

Fig. 24. Close-up view of European samples projected on Figure 23. The area of interest shows samples from North-Western Europe in the top left corner to Eastern Europe, Anatolia and the Caucasus in the bottom right corner. Colors and shapes indicate geographical origin and time period following the same code as in previous figures.

In Europe in particular (Figure 24), an interesting phenomenon can be seen among Western and Central European samples from the late Neolithic and Bronze Age. Indeed, early samples from these areas cluster on the left side of the figure. However, a shift can be seen with later samples from Netherlands and Germany. For Danish samples, we can locate Neolithic, Bronze Age and post-Iron Age samples in the second cluster starting from the top left corner (red). But Chalcolithic samples are seen away from their counterparts. They are found in the middle of PC1, closer to Bronze Age German and Neolithic Balkans samples. Indeed, German samples are mostly found in this area of the PCA, except for a group of late Neolithic samples that can be found on the right side, clustering with Eastern European samples. Samples from this time period in both Netherlands and Germany seem to be characterised by this shift towards the right side of the PCA, in direction of Eastern European (and then South-Western Asian) samples.

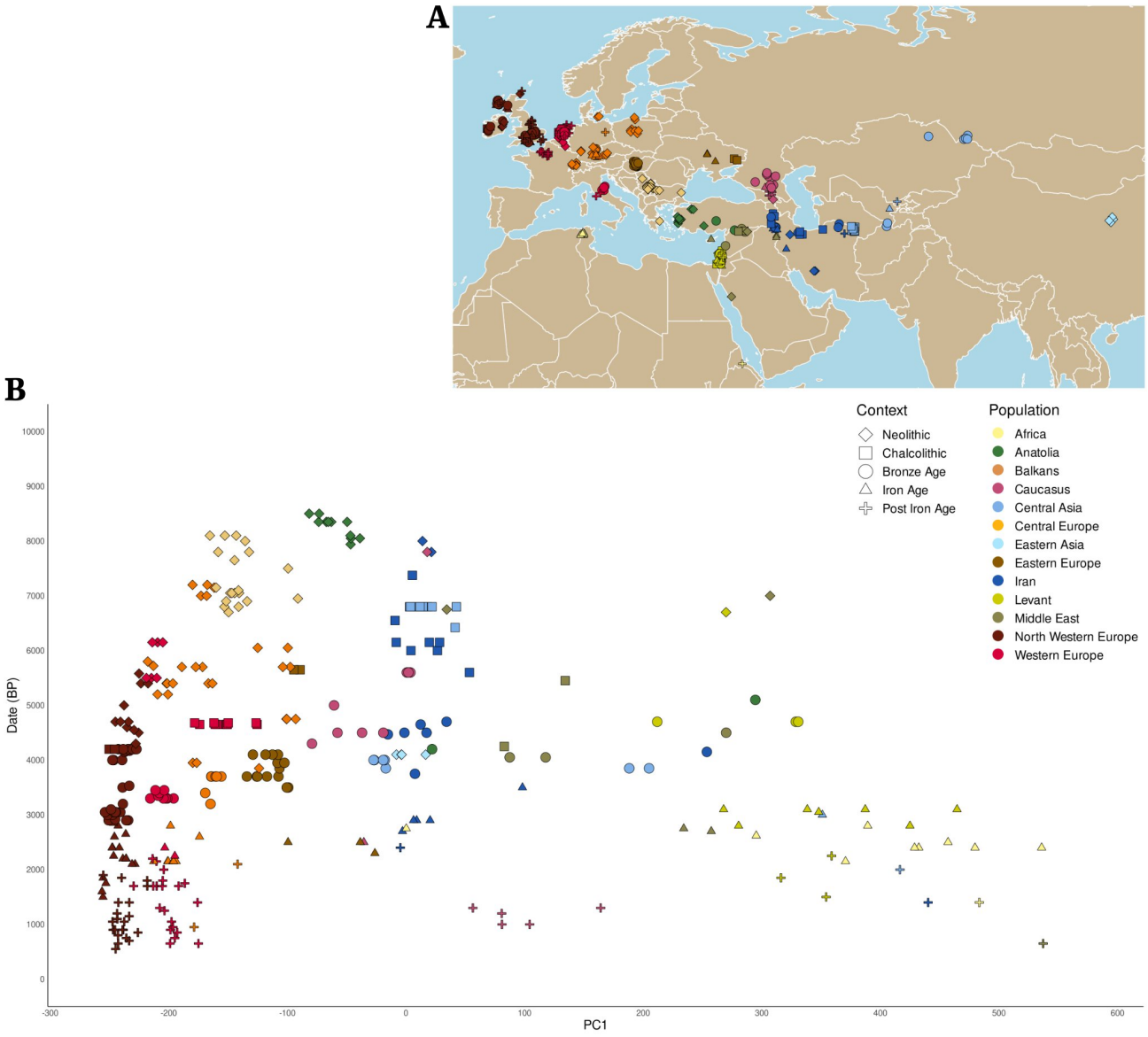

Fig. 25. Values of PC1 from Figure 23 plotted versus the age of 324 taurine samples. A) Taurine samples are plotted on a map of Eurasia and north of Africa, with colors and shapes representing population and age respectively. B) Taurine samples are plotted by age versus PC1 values. When the exact age of a sample was not known, a mean site age was determined according to archaeological information. The bottom of the graph is the present and older samples stretch toward the top.

In order to see the interaction of time with PC values from Figure 23, the results were plotted versus age of sample. Values from PC1 were chosen as this component gathers the largest fraction of variability between samples (6.5% against 3.6% for PC2). On

Figure 25, the age of the samples is represented on the vertical axis and PC1 values on the horizontal axis.

As above, PC1 separates genomes by geography (with overlap): European samples are located on the left of the graph (first North-Western, then Western Europe, Central Europe, the Balkans and Eastern Europe). Greater PC values are then found in Anatolia and the Caucasus, and then in Iran, Central and Eastern Asia. Finally, samples from the Middle East, the Levant and Africa are plotted on the right side of Figure 25. An interesting question is whether PC1 location is sustained in those geographical regions with multiple samplings through time.

PC values evolve differently depending on the region. As seen on Figure 25, North Western European samples are all aligned on a vertical line from the Neolithic to post-Iron Age. Almost the same trend can be seen in Western Europe, but with the exception of a group of Chalcolithic samples, all from the Netherlands, associated with the Corded Ware culture. In Central Europe, a linear trend similar to the West is also present but values are more spread, especially during the middle-late Neolithic.

In Anatolia, Neolithic samples from Cukurici, Ulucak, Mentese and Suberde cluster together but greater variation becomes apparent in Bronze Age samples. In the Caucasus where samples date from the Neolithic to Medieval times, there seems to have been a substantial shift after the Iron Age.

4.3.2. D-statistics

4.3.2.1. Indicine introgression in taurine samples

Samples were assessed for significant *Bos indicus* introgression using the D test - *D(outgroup, test sample; Sub1, Sha3b)*. This test (also called the ABBA-BABA test) estimates the excess of allele sharing between populations, which suggests a gene flow. The outgroup, a Water Buffalo, is used to determine the ancestral (A) or derived (B) state of the SNPs shared between the samples and Sub1 or Sha3b. Sub1 is a Neolithic Anatolian genome sampled prior to zebu domestication and Sha3b is a modern North Indian Sahiwal *Bos indicus* animal - both have been sequenced to high coverage (Verdugo et al., 2019). A

significantly positive D-value indicates excess allele sharing between the test sample and the indicine population (Sha3b) relative to the taurine population (Sub1), consistent with *Bos indicus* introgression.

The dataset of 337 ancient taurine cattle was filtered to only keep samples that have at least 500 ABBA+BABA sites. This led to the exclusion of 14 samples. Test scores are plotted in Figure 26 and a detailed table is present in the Appendix (Supplementary Table S7).

Samples stretching from North Western Europe up to the Balkans all show no trace of *indicus* except for one sample, Cer1, that has a non significant D value (number of sites = 831). Significant results first appear when we reach Eastern Europe in the form of one Iron Age sample from Ukraine (Mam1, D = 0.029, number of sites = 62,927).

It is in Anatolia and further East that D-statistics significantly indicate indicine introgression in a substantial number of samples, especially from the Bronze Age onwards. In Africa, only one sample is significant for introgression, a post-Iron Age sample from Ethiopia, VEM337. Indicine admixture is not detected in Neolithic samples from China (Eastern Asia group in Figure 26) and the highly negative values in these probably reflect an input from eastern Aurochs.

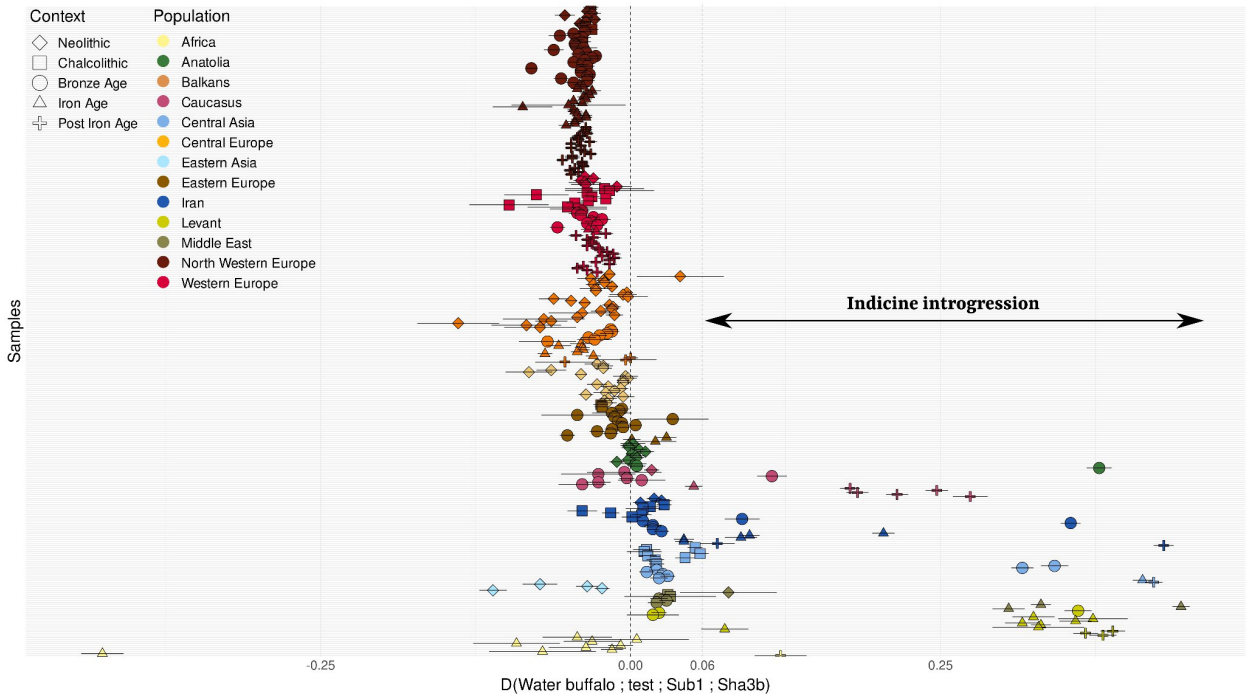

Fig. 26. Indicine introgression assessed via D-statistics in 323 ancient taurine samples. A positive D value means introgression is detected in the sample ancestry. Standard errors are represented by a bar on both the left and right side of each sample. Samples were grouped by region and ordered by time within each region, from oldest to youngest from top to bottom, as also shown by shapes denoting archaeological era. A vertical line at x=0.06 indicates the value used to exclude samples from later analyses due to indicine introgression.

4.3.2.2. Hybridization with wild *Bos*

Samples that did not present any trace of indicine introgression (Figure 26 and Supplementary Table S7; a cutoff of D=0.06 was used, which concurs with ADMIXTURE results, below) were tested for wild introgression. Two D-statistics were used assessing introgression from separate regional variants of wild aurochs: European represented by Bed3 and Eastern Asian represented by Baikal1, both, using the Neolithic Anatolian genome as an unadmixed reference, were calculated and co-plotted in Figure 27. Only samples that had at least 500 informative sites were kept, thus removing 12 samples : Aak1, Abu1, Abu2, Azer1, Dan1, Gil1, Mon1, Nah1, Tda1, Tmq1, Tmq2 and Tqa1. Figure

27.A shows the full figure, while the panel B of that same figure is a zoom in the main cluster, where all but 2 samples fall.

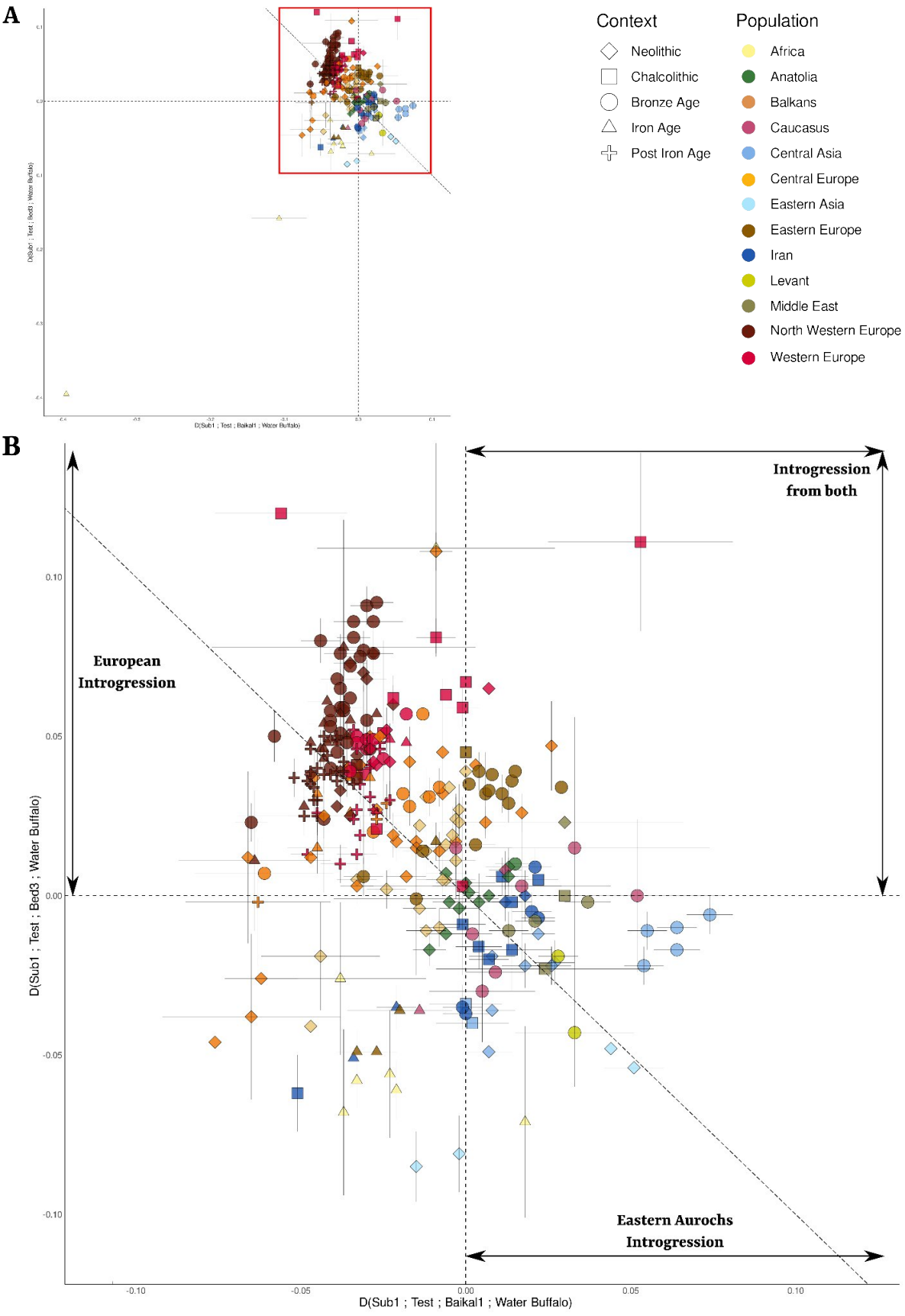

Fig. 27. Comparison of D-statistics assessing introgression from 2 aurochs in 287 ancient taurine samples. A) Original plot showing the entire dataset. B) A close-up view of 285 samples forming the "main" cluster. Samples with at least 500 ABBA+BABA sites for both Bed3 and Baikal1 computations were plotted, with standard errors. They are represented by color and shape for geographic and time information, respectively. A positive D value means an introgression was detected. The horizontal axis assesses Baikal1 (= Eastern aurochs) introgression while the vertical axis assesses Bed3 (= European aurochs) introgression. A positive signal from both types of aurochs will lead the sample to be located on the top right corner of the figure.

The two samples not included in the zoom in Figure 27.B are from Iron Age Tunisia. They appear in the lower left corner of Figure 27.A, due to negative D values indicating an absence of introgression from aurochs similar to either Bed3 (Mesolithic Germany) on the vertical axis or Baikal1 (Mesolithic Central Asia) on the horizontal axis. These individuals probably had a separate African wild input, resulting in these extreme values. The bulk of the sampled genomes fall in two quarters of the graph (Figure 27.B). First, almost all Europeans plot in the top left hand side, indicating significant Bed3 (European aurochs) introgression but not Eastern aurochs input. Within these there is a trend from west to east in the strength of this signal. Second, a group of mainly Asian (Eastern, Central and South-Western) genomes falls in the bottom right quarter indicating Baikal1 (Eastern aurochs) introgression but not wild European admixture. In the top right quarter, a group including European and Near Eastern genomes may have introgression from both. More specifically, a group of Bronze Age Hungarian and Polish samples, in addition to late Neolithic Polish and German, show Eastern introgression. One sample, LW017 (Chalcolithic Netherlands), is seen standing out of the cluster, showing significant introgression from both aurochsen. Another sample from the same site, LW028, is shifted leftward along the horizontal, indicating European but no Eastern introgression.

4.3.2.3. Autosomes vs. X

Rossi et al. (2024) used a D statistic finding that there was more wild allele sharing in autosomes relative to autosomes to assert that aurochs introgression into domestic

cattle was male-mediated. Study of the Y chromosome directly is difficult because i) it is carried by only 50% of the population, ii) it is harder to retrieve from ancient samples than mtDNA, which is present in several copies per cell, and iii) modern breeding practices tend to reduce paternal diversity (one bull is used to produce many offspring).

Consequently, for each sample with at least 500 informative autosomal sites and 100 for the X chromosome two separate D statistics of the form *D(Sub1, Test; Bed3, Water buffalo)* were calculated and plotted together in Figure 28. The horizontal axis measures introgression via autosomes while the vertical axis measures introgression via the X chromosome. In the figure, the bulk of samples with European aurochs introgression (positive values) fall below the 1:1 line indicating that admixture was more extensive on autosomes than on the X. Given that wild males transmit only half as many X chromosomes as autosomes to offspring, this is consistent with a bull-focused introgression process.

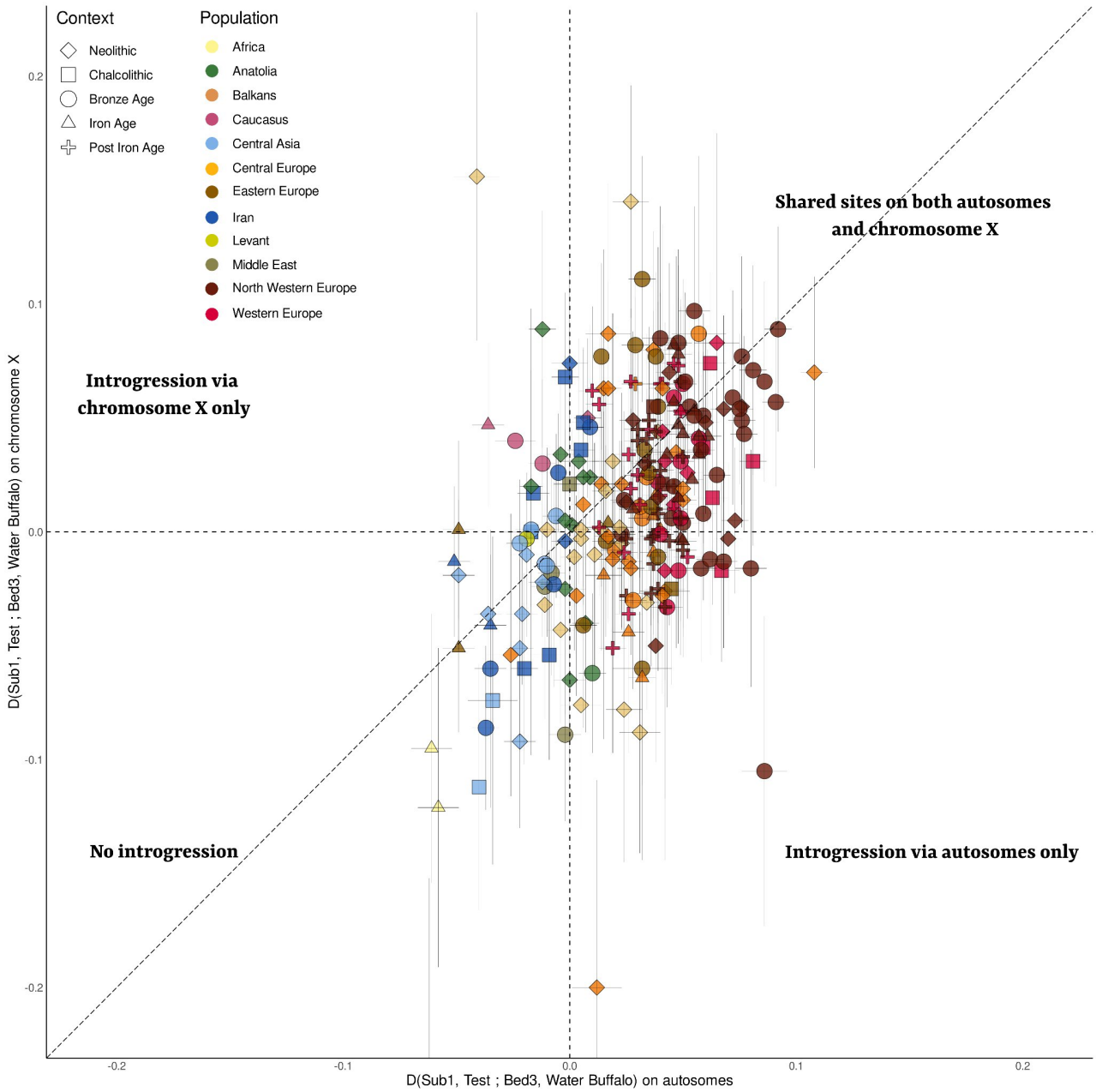

Fig. 28. D-statistics values for European aurochs introgression on X chromosomes only compared to autosomes. Eastern Asian samples were removed due to not showing introgression from European aurochs on Figure 27. Samples were selected based on the combined number of ABBA+BABA sites for autosomes (>500) and X chromosomes (>100). Colors and shapes represent geographic origin and time context respectively. D-statistics were computed using ANGSD and plotted via R.

When this process is repeated with the replacement of Bed3 with the high coverage Eastern aurochs Baikal1 as reference, a somewhat different picture emerges (Figure 29).

Here, earlier East Asian cattle with East Asian aurochs introgression show less X chromosome admixture, consistent with male introgression. However later Central Asian and Near Eastern animals show somewhat higher X introgression. As these animals likely also have European aurochs influence, this may be a confounding influence in this test.

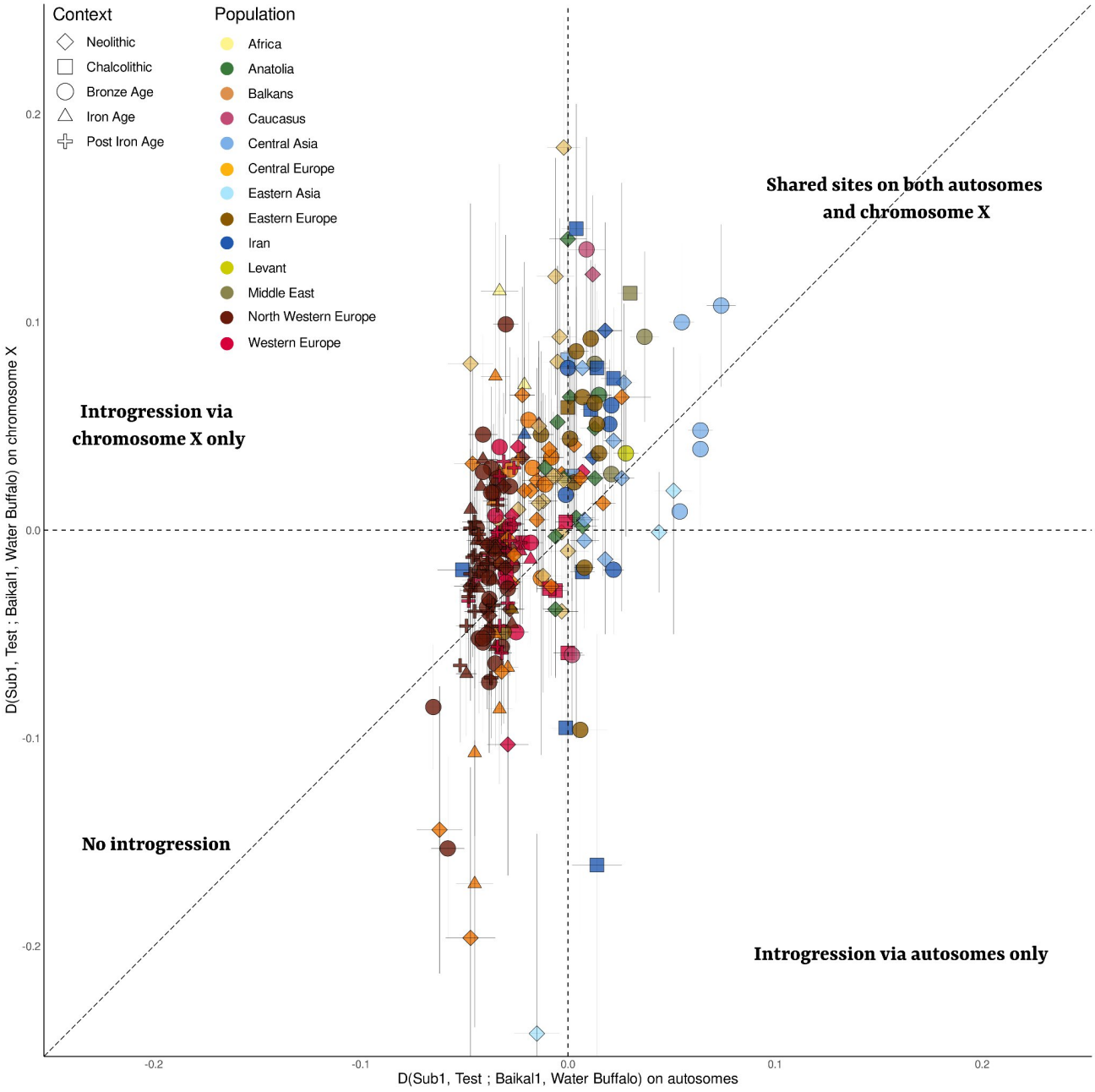

Fig. 29. D-statistics values for Eastern aurochs introgression on X chromosomes only compared to autosomes. Samples were selected based on the combined number of ABBA+BABA sites for autosomes (>500) and X chromosomes (>100). Colors and shapes

represent geographic origin and time context respectively. D-statistics were computed using ANGSD and plotted via R.

4.3.3. Ancestries assessment with Admixture

Model-based estimation of ancestral components was performed using ADMIXTURE with the dataset of ancient and modern individuals of *Bos taurus*, *Bos indicus* and *Bos primigenius*, for K values (number of components present) of 1 to 8. The cross validation (CV) errors computed showed a best fitting model at K=6 (Table 12 below).

Number of ancestries	CV error
1	0.4449
2	0.3551
3	0.3513
4	0.3508
5	0.3515
6	0.3494
7	0.3501
8	0.3517

Table 12. CV errors calculated by ADMIXTURE for different values of K (1 to 8).

This model separates populations according to species and geography, following 6 ancestries that can be identified as follows:

- *Bos indicus* of China;
- *Bos indicus* of India and Pakistan;
- *Bos taurus* of Europe;
- *Bos primigenius* of Europe;
- *Bos taurus* from the Levant and Africa;
- *Bos primigenius* from Northern Asia.

Results are detailed in Supplementary Table S8, and are plotted for ancient genomes in Figure 30 below. Blue colors represent indicine ancestry; Chinese indicine was

found to be a component in the full dataset of ancient and modern samples but no ancient samples showed signs of this ancestry.

Taurine cattle are composed mostly of a European and a Middle-Eastern ancestral component (respectively green and red in Figure 30). Both can be observed throughout the Middle East and Europe, although in different proportions, except for the British Isles where the Middle-Eastern component is not observed until the Iron Age. In Europe, local aurochs played a role in the ancestry of domesticated cattle, as seen by the portion of brown representing ancestry from that population. Eastern aurochs show a different component to European aurochs. In older samples, in Figure 30.A, the Eastern component, in yellow, is only present in Northern and South-Western Asia, and in one Polish sample. The Late Neolithic and Bronze Age in Europe (Figure 30.B) brings this Eastern component into more Central European samples in Germany and Hungary. This time period also sees the arrival of the indicine component, in high proportions, in South-Western Asia (seen in blue on Figure 30).

More recent samples in Europe seem to homogenize (Figure 30.C). The Eastern component almost disappears, except in one German sample, and the European aurochs and Levantine taurine aspects (respectively in orange and red) decrease, in favour of a European taurine component. Domesticated cattle outside of Europe are still a mixture of different taurine, indicine and wild components. In Africa, the two Iron Age samples from Tunisia show a much bigger portion of the Africa/Levantine cattle ancestral component and a smaller European one. No indicine or either European or Asian aurochs ancestry is detected in them.

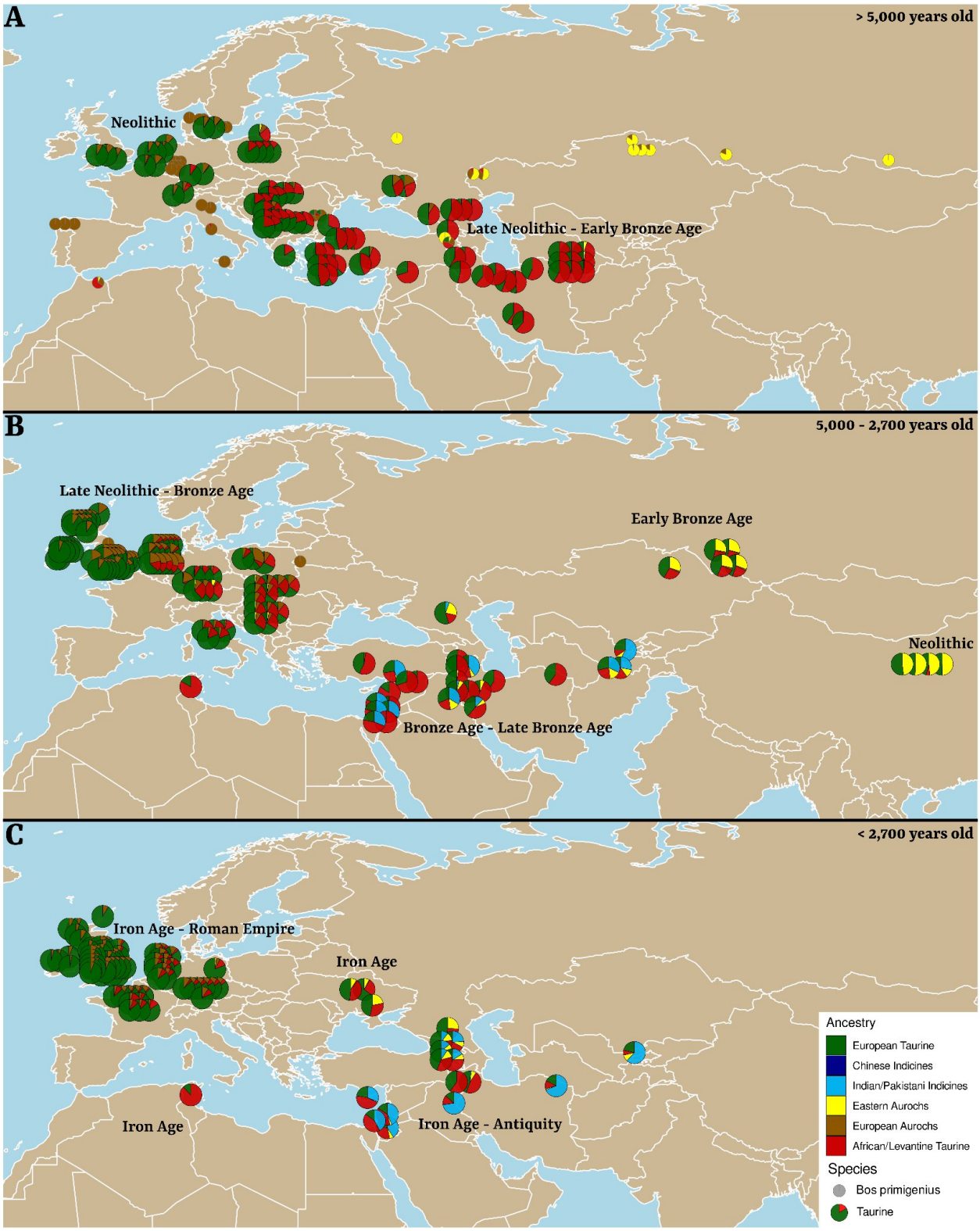

Fig. 30. ADMIXTURE results with K = 6 on domesticates and wild samples of Eurasia. Samples were selected based on the number of sites retained for analysis (50k) and divided in the three panels based on their age. Aurochs are shown in smaller circles with grey outline, while *Bos taurus* are shown in larger circles with black outline. A) Samples older

than 5,000 years; B) Samples aged between 2,700 and 5,000 years; C) Samples younger than 2,700 years.

A "traditional" admixture plot is shown in Figure 31 below. It includes modern samples that helped form the components in ancient samples. Indeed, as seen in Figure 31.C, indicine samples from modern China and the Indian subcontinent are represented as two uniform groups, colored in shades of blue. The modern “Africa Taurine” group is mainly composed of a red component, also majoritarily present in ancient Levantine taurine samples. European and Eastern aurochsen appear as distinct groups as well in Figure 31.B, which allowed for the identification of the Eastern and European wild components (yellow and brown).

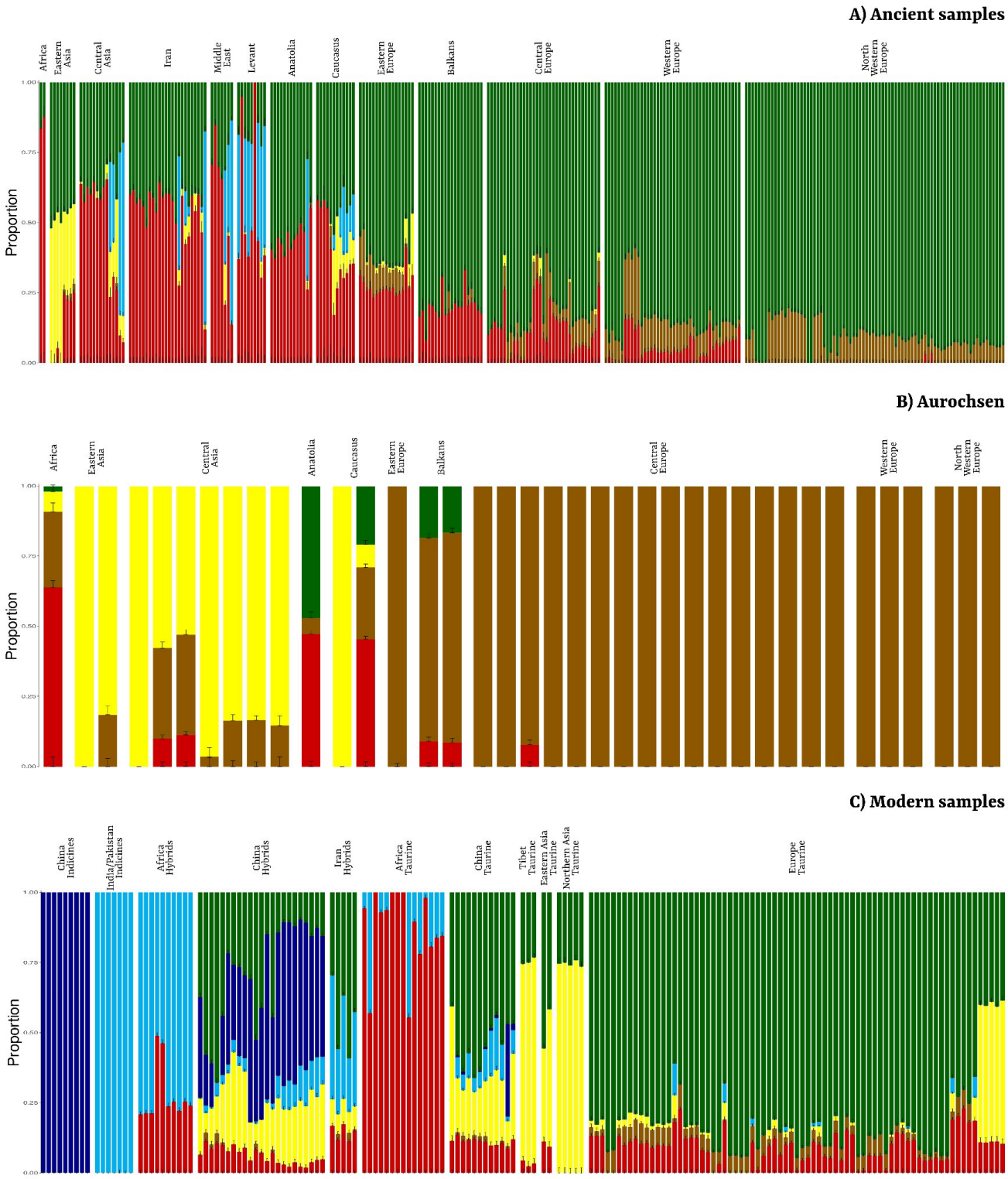

Fig. 31. ADMIXTURE results with K = 6, presented in barplot format and including modern individuals. Samples were selected based on the number of sites retained for analysis (50k) and divided in three groups based on their species and age. A = ancient domesticates, B = *Bos primigenius* individuals, C = modern samples (*Bos taurus*, *Bos indicus*, hybrids).

4.3.4. Eastern aurochs proportions estimation via f4-ratios

The introgression from eastern aurochs detectable using D-statistics was estimated for each genome by computing an f4-ratio. The ratio used was as follows:

$$\frac{f_4(Tula1, Outgroup\ ;\ Test, Sub1)}{f_4(Tula1, Outgroup\ ;\ Baikal1, Sub1)}$$

Tula1 and Baikal1 are both eastern aurochs (as seen in Figure 30), Sub1 is a Neolithic Anatolian domesticate. Results are plotted as pie charts representing two ancestries (domesticates and wilds) in Figure 32, employing only samples that have at least 0.01X coverage. Samples were divided among three maps, according to age. Oldest samples, on Figure 32.A) are more than 5,000 years old, while Figure 32.B depicts samples aged from 5 to 2,700 years, the most recent samples being represented on the last panel of Figure 32.

Pie charts summarise the amount of each ancestry, from Eastern wild aurochs (represented by Baikal1, an aurochs from Eastern Russia) Vs domesticated cattle (represented by Sub1, a Neolithic Anatolian domesticate). A slight influence of wild ancestry can be seen in early South-Western Asia, as seen via the presence of yellow in almost every individual of Anatolia, Iran and Turkmenistan in Figure 32.A. Domesticated cattle from the Caucasus all have traces of Eastern ancestry as well. In Europe, the amount of wild Eastern ancestry decreases from East to West, and is not detected in Germany, Denmark, or Great-Britain. It is detected in Ukraine, Poland, Hungary, Switzerland, Belgium and the Netherlands (with percentages ranging from 0.1 to 0.4 % in Belgium and the Netherlands to 10% in Serbia and 18% in Switzerland).

In the second panel (Figure 32.B, 5,000-2,700 years ago), we can observe the progression of wild Eastern ancestry represented by the increase of the yellow proportion in the pie charts. It is now detected in a significant amount in Germany and the Netherlands (with percentages ranging from 0.3 to 25%), and even in one sample from England (though the result for this sample is not significant). Domesticated cattle from Poland and Hungary are still showing traces of Eastern ancestry (0.1 to 5% in Poland, and 0.01 to 13% in Hungary), which is also found in Italian cattle (0.7% of ancestry). The two Caucasian

samples have a greater amount of wild ancestry than their older counterparts (8-14% vs 1.5-6.5% in earlier samples). Other South-Western Asian samples still present traces of wild ancestry. Significant results are found in Russia and China.

In the most recent samples, the wild ancestry decreases greatly and is barely seen in Europe; only in Germany and Great-Britain. The few individuals representing Eastern Europe, the Caucasus and Iran do not present any sign of Eastern wild ancestry. It is however significantly present in Iron Age cattle from Tunisia (with the highest amount of wild ancestry being 22%). This last result gives some indication of the limitations of this approach. It is unlikely that the Eastern aurochs ancestry present in Baikal1 had a specific input to these African cattle. Rather, some of our signals may be due to local aurochs' introgressions, from populations that show some affinity to Baikal1 compared with the domesticated Anatolian Neolithic reference in the calculation.

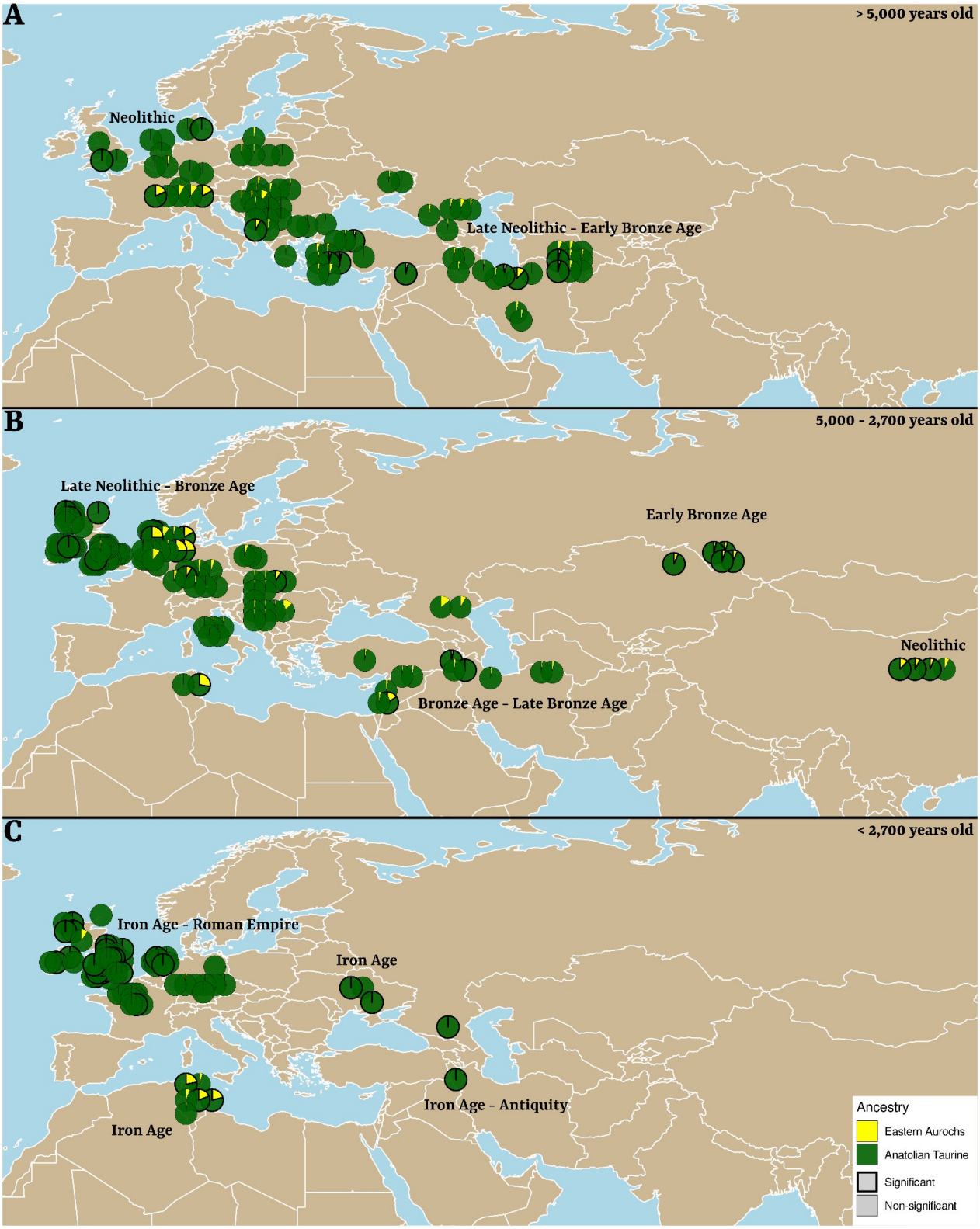

Fig. 32. F4-ratios depicting domesticate and wild ancestry in Eurasian domesticates. The wild ancestry was assessed using an aurochs representative from Eastern Russia, while a Neolithic Anatolian sample is used to represent taurine cattle. Samples were divided in three groups based on age, and represented by pie charts depicting two ancestries:

domesticated in green, wild Eastern Asian in yellow. A) Samples older than 5,000 years; B) Samples aged between 2,700 and 5,000 years; C) Samples younger than 2,700 years. Significant results (Z-score > 3) are shown in pie charts with a thicker outline.

To investigate the validity of f4-ratio results, values of Eastern introgression detected by the test were plotted against values estimated via ADMIXTURE (Figure 33). A correspondence is apparent but is not a simple linear one.

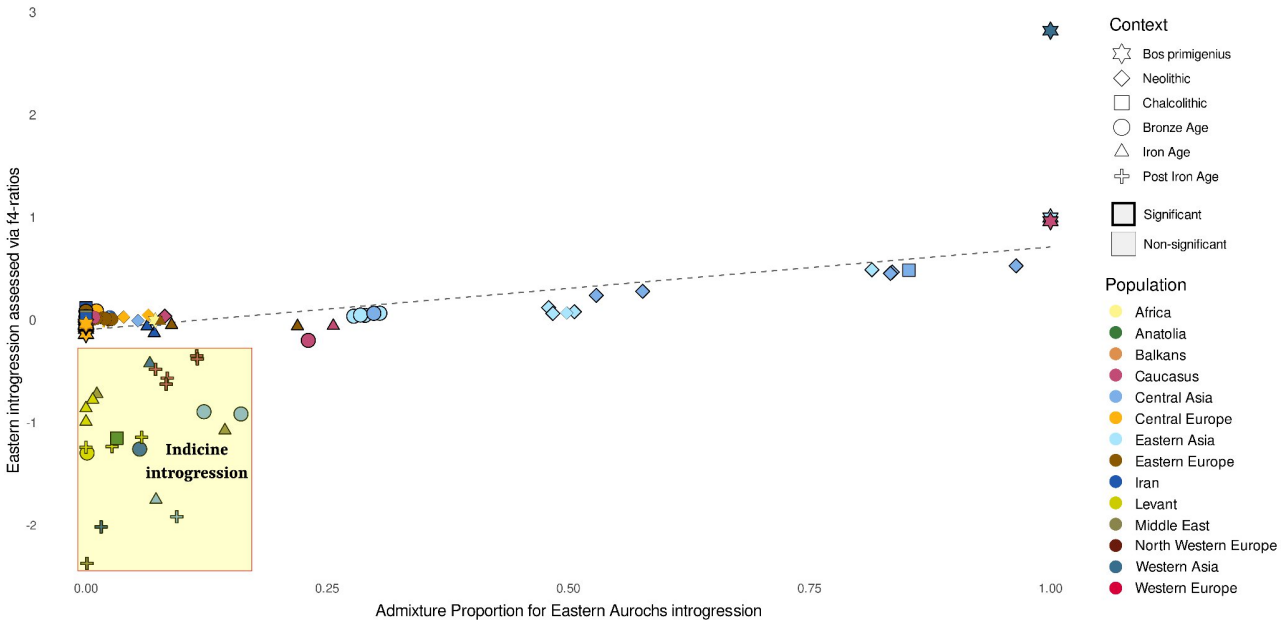

Fig. 33. F4-ratios against ADMIXTURE values for Eastern introgression. A) Samples are represented by dots shaped according to age and colored according to geographic origin. Significant f4-ratios values are indicated with a thicker outline. In the bottom left corner showed indicine introgression in ADMIXTURE (>0.09). B) Samples showing indicine introgression and the sample in the top right corner in A) were removed to show a better view of the rest of the dataset.

In Figure 33, values found from f4-ratios and admixture seem to follow the same trend, with exception of a group of samples on the bottom left of the figure for whom f4-ratio computations are the lowest of the dataset. This group is positive for indicine introgression, as seen with D-statistics results (Figure 26). On the top right can be found one aurochs acting as an outlier with the highest f4-ratio value, as it is itself an aurochs sharing the same haplogroup as Baikal1, who is used as reference for Eastern aurochs.

4.3.5. Ancestry modelling with qpAdm

F4-ratios use a constrained set of ancestries in their calculation, qpAdm seeks to avoid this by considering the effects of multiple ancestral populations within each estimate - with rejection of models for each individual where other ancestral groups have a confounding contribution. In order to assess wild Eastern introgression in Europe, a model was initially tried testing a triple origin:

- Neolithic Anatolia (which should be the closest to the original domesticated population), defined by the three samples Sub1, Men1 and Men2;
- European aurochs (since admixture and D-statistics showed introgression from local populations) represented by Bed3, Bed4, Fif1, Hjo1, Hxh2, Rhi3, Ska1, Ska3, Tri1, Zea1 and Zea2; and
- Eastern aurochs (individuals that have haplogroup C): Baikal1 and Tula1.

A second model tests if instead European samples were not influenced by an Eastern population, and would rather look like an admixture of Neolithic Anatolian taurine cattle and wild European aurochsen only.

Another possibility was modelled as European samples showing no European aurochsen component at all, but rather being the result of admixture between taurine cattle and Eastern Aurochsen only. Lastly, qpAdm was used to test a scenario where European cattle were a mixture of four populations: Neolithic Anatolia, European Aurochs and Eastern Aurochs, as earlier, and an additional one made of Central Asian Aurochs. This last population includes aurochsen harbouring a different haplogroup than C. However, these last two models proved unfitting to any samples.

In Figure 34 the same division as in previous figures was kept, with an upper map representing samples older than 5,000 years, a second for samples between 5 and 2.7k years and lastly the more recent samples, less than 2,700 years old. The triple ancestry model (right side of the figure) gave no acceptable fit for any samples from the first group (the oldest), while it does fit for six animals in the second group (between ~2,800 and ~4,750 years old). Taurine individuals showing ancestry from both wild European and wild Asian populations are from Hungary (Sza1, 2.0%), Germany (Kir4 and Son1, 2.7 and 7.4%),

Ireland (New4, 2.2%), Scotland (Cla1, 2.2%) and Tunisia (STK239, 6.4%). Cla1 and STK239 are the youngest samples of this group showing Eastern ancestry (Bronze Age Scotland and Iron Age Tunisia), Kir4 and New4 are the oldest (Neolithic Germany and Ireland).

The more recent (post 2,700 yrs ago) group is the best adapted to this model, with 32 samples fitting, from several North-Western and Western European countries (Ireland, UK, France, Netherlands, Germany), and Tunisia. These samples are evidence of Eastern aurochs' introgression into later European domestic cattle, with contributions ranging from 1.5% to 3.7% (while the sample from Tunisia, STK230, has 12.3% Eastern ancestry, a figure that may be confounded by the absence of an African aurochs in the reference data). They also show the westward progression of Eastern ancestry, which is found in many samples from France and the Netherlands on the continent, and UK and Ireland across the Channel.

A dual origin model was also tested using European aurochs and Neolithic Anatolian ancestors. This fitted a larger number of genomes in each time period (Figure 34, left hand side plots), with roughly a quarter of each sample's ancestry being influenced by wild European animals. This dual model is well present in North-Western Europe as well as on the Western part of the mainland. Eastern Europe, as well as the Balkans, are poorly defined by either of the models, especially from 5k years ago onwards. Older samples from Poland and Serbia (pre-5k years old) are correctly characterised by the dual European - Anatolian model. However, only two out of the fourteen samples from Hungary find a place in these models, and neither of the five Ukrainian samples nor the one Greek fit any model.

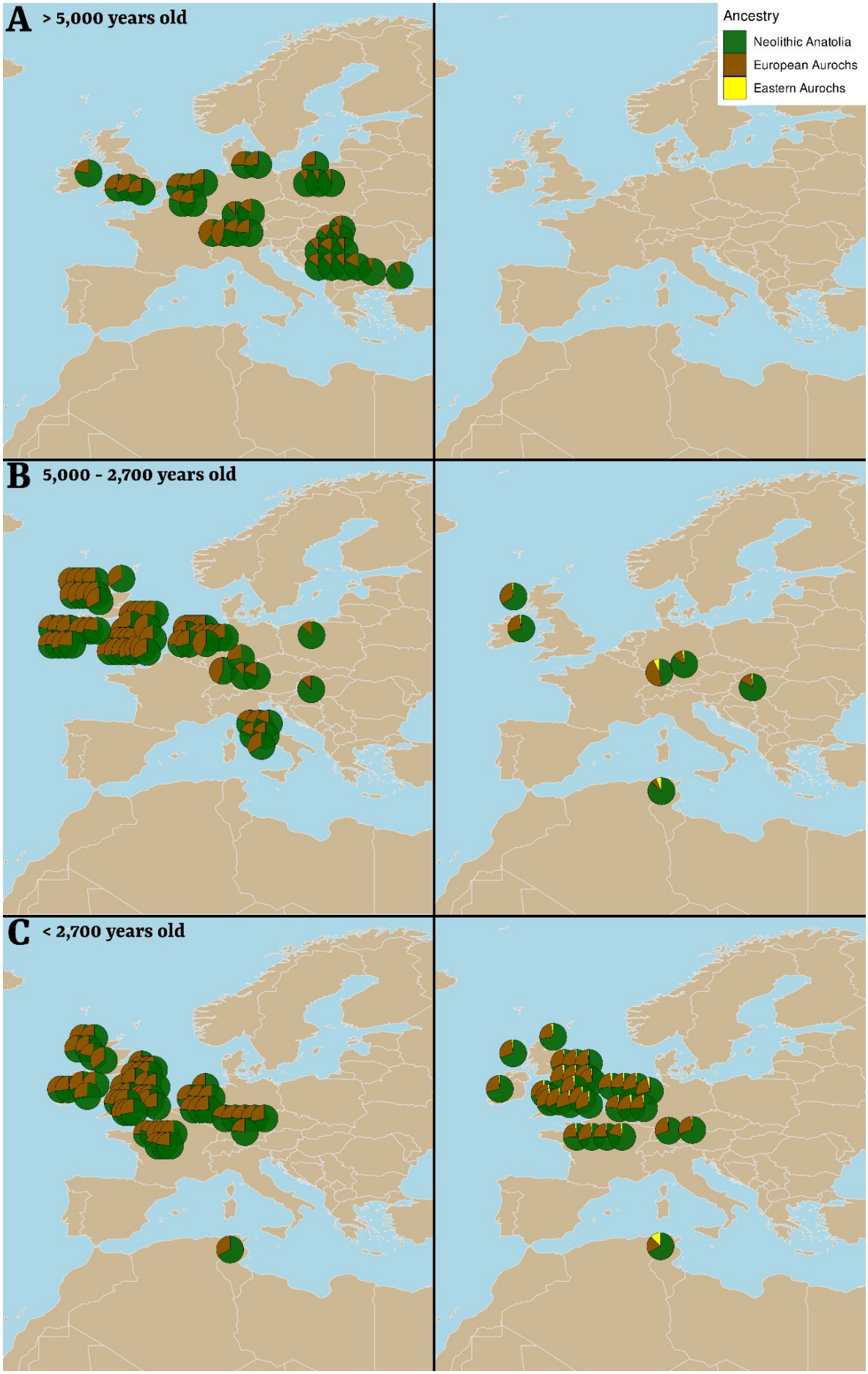
A > 5,000 years old
B 5,000 - 2,700 years old
C < 2,700 years old
Ancestry
Neolithic Anatolia
European Aurochs
Eastern Aurochs

Fig. 34. Ancestry models tested with qpAdm. Samples were tested for a model combining a domesticate, wild European and/or wild Eastern ancestries. The left side is the dual model testing domesticate + European aurochs ancestry only, right side is the model testing for the three ancestries. Plotted samples are the ones significantly fitting the models. A) Samples older than 5,000 years. B) Samples dating between 5,000 and 2,700 years ago. C) Samples younger than 2,700 years old.

4.4. Discussion

4.4.1. A variety of ancestries in ancient cattle

Ancient cattle do not form a singular group that spread to different continents from a simple single Near Eastern origin population. On the contrary, samples cluster according to geography, as seen in PCAs (Figures 23 and 24), but also according to time, as seen on Figure 25. Further analyses focusing on dissection of ancestry show several components playing a role in the genetic make-up of ancient cattle. D-statistics and F4-ratios highlight the involvement of introgressing wild populations (*Bos primigenius)* and genetics from a second, separate domestication process further east in the Indus valley (*Bos indicus*). Admixture analysis reveals 6 ancestral components, involving 3 taxa (*Bos taurus*, *Bos primigenius* and *Bos indicus*), all indicating a complex history shaped by migrations and local encounters.

4.4.1.1. The origin of domestication in the Near-East

The Fertile Crescent is known to be the centre of cattle domestication (Hongo et al., 2009, Bollongino et al., 2012), so it is not surprising to find a component using ADMIXTURE that is found maximally in the Near-East (the red cluster on Figure 30). From Anatolia to Turkmenistan, and including the Levant and the Caucasus, samples show a majority of this cluster in the earliest period featured in this study (more than 5,000 years ago - Figure 30.A). Its importance then decreases over time and space, throughout Asia (first in Central Asia and then becoming barely noticeable in China) and Europe (Figure 30.A, B and C). Samples from Western and North-Western Europe are the only ones not showing this ancestral component. These are dominated by a different European domestic

component, one driven in the ADMIXTURE analysis by the inclusion of modern domestic European breeds, within which it predominates (Figure 31.C). This ancestry is likely a description of the differentiation due to drift of these breeds during their formation in migrations from the Near East.

4.4.1.2. A European domesticates component

The decrease of the Levantine/African cluster from Eastern to Western Europe is replaced by the increase in a green cluster, labelled as “European taurine” on Figure 30. This component is mostly found in Europe, with the largest proportions located in North-Western Europe (where the Levantine/African component is absent). However, it is interesting to note that this European component is still part of Near-Eastern samples of all time periods in the admixture analysis. Also, modelling European samples with qpAdm (Figure 34) showed an accepted model when using Neolithic Anatolian samples as proxy for early domesticates. This component, while being a clear European signal, hence shows links with the Near-East, as a remnant of early domesticates from the Fertile Crescent. As time went by, genetic drift likely shaped European cattle into their own population, while keeping the signal of a Near-Eastern origin, a result already seen in Scheu et al. (2015).

4.4.1.3. Introgression from local aurochsen in Europe

Domesticated cattle were not strictly taurine in Europe. A third major component in European cattle is the one identified on Figures 32 and 33 as “European aurochs” (in brown). *Bos primigenius* from the continent (denoted as small circles in Figure 30) are defined by this ancestry (also identified in Figure 31.B). Together, with D-statistics testing for this introgression (Figure 27), and accepted fits from qpAdm models assuming the same (Figure 34), this is proof that local wild populations admixed with local domesticates, whether passively or as a means for breeders to improve their herds. Such findings were already described in several papers, via mitochondrial studies suggesting a female-driven introgression implied by the discovery of divergent haplogroups (Achilli et al., 2008; Schibler et al., 2014), but also alternately via whole-genome sequencing that found wild introgression but no trace of the wild haplogroup P which is characteristic of wild European oxen, suggesting a male-driven introgression (Park et al., 2015; Upadhyay et al., 2017).

Indeed, direct observation of introgression via Y chromosome is challenging, as modern breeding practices have reduced paternal diversity, and nuclear DNA is more difficult to retrieve than mtDNA. It was still attempted in a study by Götherström et al. (2005), who suggested male introgression in Europe, but was later contradicted by Bollongino et al. (2008b). However, a signal of an eastern aurochs male-driven introgression was highlighted in Rossi et al. (2024), confirming results we found with D-statistics that compared autosomes and X chromosome alleles (Figure 28). Indeed, the profile seen in allele-sharing seems biased towards the autosomes rather than the chromosome X, suggesting a male mediated introgression in Europe.

4.4.1.4. Eastern aurochsen ancestry can also be found in domesticates of Eurasia

Wild introgression did not only involve European aurochsen. A signal from an Eastern aurochs was found in our dataset, even outside of its geographical area (Eastern Russia). In Figure 27 D-statistics detected this introgression, in domesticated cattle from Russia, as expected, but also in other countries, mostly in South-West Asia (Turkmenistan, Iran, Levant, Anatolia and the Caucasus), and more surprisingly in Eastern Europe, despite the distance between the source and the target. This signal is also found via f4-ratio computations from the Bronze Age in Central Europe and even in two Scottish samples. Additionally, admixture analysis confirms the Eastern component in several European samples, first in a Neolithic Polish sample, but it is in the Bronze Age that this signal moves westwards to Hungary and Germany. This signal can be observed on the PCA as well (figure 24), where Central and Western European domesticates are found away from their respective clusters during the late Neolithic and Bronze Age, a behaviour not seen in later samples. This likely coincides with the spread of the Yamnaya, a people who migrated from the Asian Steppes towards Eastern Europe around the Bronze Age, and was then followed by the Corded Ware culture in Central Europe (Haak et al., 2015, Juras et al., 2018). Every positive introgression detected by one or more analyses is informed in Tables 13, 14 and 15, each focusing on one of the three time periods used to represent samples according to age (older than 5,000 years, dating between 5,000 and 2,700 years, and younger than 2,700 years).

A curious result can be seen in our dataset, concerning four samples from Twann, Switzerland. These pre-date the Corded Ware in Central Europe, but f4-ratios seem to contradict this. Indeed, an abnormally large amount of the Eastern component was found in these samples (Figure 32), which is at odds with the hypothesis that this ancestry only arrived with the Yamnaya in Eastern Europe and the following Corded Ware culture in Central Europe. However, it is noticeable that, firstly, these samples are very low coverage (0.01X) and results are not significant for two of them, and secondly, other analyses (D-statistics, qpAdm) do not show this Eastern component, but rather a typical Neolithic profile for Europe.

Thus, apart from the f4-ratios for these 4 samples, the various analyses concur on the timing for the Eastern input in Europe, during the late Neolithic and Bronze Age, and agree on a strong decrease, if not a total disappearance in some places, in later time periods. With the exception of Ukraine which shows similar proportions to the Caucasus in the admixture analysis (but not the f4-ratios).

A similar scenario happens in the Middle East where an Eastern aurochs component is detected via admixture, D-statistics and f4-ratios from the Neolithic in Turkmenistan and the Bronze Age in Iran and the Caucasus. This could come from introgression with a local population of aurochsen similar to the one found in Russia, or perhaps Eastern component progressed through Central Asia (as it is detected in Chalcolithic Turkmenistan) into South-Western Asia with migrating stock as it did in Europe.

Table 13. Comparative table of analyses testing for Eastern introgression into European samples older than 5,000 years.

Lab name	Country	Coverage	Position shift on Laser PCA?	D-statistics	Admixture	qpAdm	F4-ratio
Bel1	Serbia	0.090					Y
Bel2	Serbia	0.130					Y
Bis1	England	3.020					

Bis2	England	2.740					
Bla1	Serbia	0.300					Y
Bla2	Serbia	0.060					Y
Bla3	Serbia	0.020					Y
Bub1	Serbia	0.800					
Bub2	Serbia	0.110					Y
Cer2	Germany	0.090					
Che1	Bulgaria	0.480	Y				Y
Den1	Denmark	3.550					
Den2	Denmark	3.930					
Dzh1	Bulgaria	9.680					
Dzh2	Bulgaria	4.100	Y				
Hxh1	Germany	8.390					
LARK02	UK	0.140					
Lud1	Poland	0.510					Y
Lud2	Poland	0.350					Y
Lud3	Poland	1.800					
Lud4	Poland	0.190					Y
MDR1	Netherlands	0.370					Y
Plo1	Serbia	0.130					Y
Plo2	Serbia	0.480					Y
Plo3	Serbia	2.630					Y
Plo4	Serbia	2.860					
Plo5	Serbia	0.080					Y
Plo6	Serbia	0.460					Y
Plo7	Serbia	0.110					Y
Plo8	Serbia	0.200					Y
Ruc1	Poland	0.008	Y				
Ruc2	Poland	0.490	Y	Y	Y		Y
Sar38	Greece	0.480					Y
Sch1	Netherlands	1.710					

Sch2	Netherlands	1.270					Y
Sch3	Netherlands	0.040					
Sha1	Ukraine	0.370					Y
Sha2	Ukraine	0.440					Y
Stu1	Serbia	0.150					Y
Sut1	Switzerland	1.870					
TAW1	Belgium	0.050					Y
TAW2	Belgium	0.040		Y			Y
TWAN01	Switzerland	0.010					Y
TWAN02	Switzerland	0.040					
TWAN03	Switzerland	0.010					Y
TWAN04	Switzerland	0.010					Y
TWAN05	Switzerland	0.010					Y

Table 14. Comparative table of analyses testing for Eastern introgression into European samples dating between 2,700 and 5,000 years.

Lab name	Country	Coverage	Position shift on Laser PCA?	D-statistics	Admixture	qpAdm	F4-ratio
Bisk1	Poland	2.430	Y	Y			Y
Cer1	Germany	0.005	Y				
Cer3	Germany	0.520	Y		Y		Y
Cla1	Scotland	15.250				Y	
Cla2	Scotland	1.860					
Cla3	Scotland	0.150					
Cla4	Scotland	0.320					
Cla5	Scotland	3.330					
Cla6	Scotland	0.360					
Cla7	Scotland	0.290					
Cla8	Scotland	3.100					
Cla9	Scotland	2.920					

Dur1	England	2.440					
Dur2	England	2.140					
Dur3	England	0.380					
ENK1	Netherlands	1.910					
ENK2	Netherlands	0.280					Y
ENK3	Netherlands	2.670					
ENK4	Netherlands	2.140					
ENK5	Netherlands	0.270					
GQ1	England	0.070					Y
Gub1	Italy	1.050					Y
Gub2	Italy	0.550					Y
Gub3	Italy	0.420					Y
Gub4	Italy	0.440					Y
Gub5	Italy	0.490					Y
HF1	N. Ireland	0.170					
Irt1	England	0.900					
Irt2	England	0.510					
Irt3	England	1.020					
Irt4	England	0.890					
Irt5	England	0.250					
Irt6	England	1.290					
Irt7	England	0.660			Y		
Irt8	England	1.370					
Kas1	Germany	0.040			Y		Y
Kie1	Poland	0.040					Y
Kie2	Poland	0.970					Y
Kir4	Germany	14.880	Y	Y	Y	Y	
Kir5	Germany	0.060	Y				Y
Kir6	Germany	0.030	Y	Y	Y		Y
Kun1	Germany	0.050					Y
Kun2	Germany	1.000					Y

Kun3	Germany	1.000					Y
LARK01	UK	0.700					
Lec1	England	1.770					
Lec4	England	2.460					
LW017	Netherlands	0.010	Y	Y			Y
LW018	Netherlands	0.010	Y				Y
LW019	Netherlands	0.170	Y				Y
LW020	Netherlands	0.010	Y				Y
LW021	Netherlands	0.410	Y				Y
LW023	Netherlands	0.750	Y				Y
LW024	Netherlands	0.570	Y				
LW025	Netherlands	0.440	Y				Y
LW026	Netherlands	0.010	Y				Y
LW027	Netherlands	0.010	Y				Y
LW028	Netherlands	0.010	Y				Y
Ness5	Scotland	2.730					
New1	Ireland	16.380					
New3	Ireland	5.200					
New4	Ireland	4.230				Y	
New5	Ireland	4.700					
Par1	Ireland	1.710					
Par2	Ireland	2.020					
Pot2	England	0.440					
Pot3	England	1.600					
Pot5	England	1.720					
Pot6	England	0.350					Y
Rou1	Ireland	2.930					
Rou2	Ireland	2.980					
Rou3	Ireland	1.300					
Rou4	Ireland	0.980					
Son1	Germany	0.010				Y	Y

STB1	Netherlands	1.650					
STB2	Netherlands	1.690					
STB3	Netherlands	1.280					Y
Stra1	Scotland	1.250			Y		
Sza1	Hungary	1.540		Y	Y	Y	Y
Sza10	Hungary	0.470		Y	Y		Y
Sza11	Hungary	0.030		Y			Y
Sza12	Hungary	0.010		Y			Y
Sza13	Hungary	0.690		Y	Y		Y
Sza15	Hungary	0.007					
Sza2	Hungary	1.630		Y	Y		Y
Sza3	Hungary	2.240		Y	Y		Y
Sza4	Hungary	1.780		Y	Y		Y
Sza5	Hungary	1.850		Y	Y		Y
Sza6	Hungary	0.480		Y	Y		Y
Sza7	Hungary	0.470			Y		Y
Sza8	Hungary	1.020		Y	Y		Y
Sza9	Hungary	0.430			Y		Y
Viz1	Netherlands	0.710					

Table 15. Comparative table of analyses testing for Eastern introgression into recent (less than 2,700 years old) European samples.

Lab name	Country	Coverage	Position shift on Laser PCA?	D-statistics	Admixture	qpAdm	F4-ratio
Asp1	UK	1.850				Y	
Bal1	Ireland	2.640					
Bal2	Ireland	2.030			Y	Y	
Bor1	Scotland	2.680				Y	
Bri1	Netherlands	7.380					

Brn1	France	1.000					
Brn2	France	1.180					
Brn3	France	2.790				Y	
Brn4	France	1.120					
Brn5	France	1.430					
Bun1	Netherlands	1.760					
Bun2	Netherlands	2.380					
Cer4	Germany	0.150					
Cha1	England	2.170				Y	
Cri1	UK	0.690				Y	
Da1	England	7.280					
Da2	England	0.260					
Da3	England	2.510					
Da4	England	4.830					
Da5	England	5.440				Y	
Da6	England	3.450					
Dro1	Netherlands	2.840				Y	
Dub1	Ireland	2.180					
Dyr1	Germany	2.310			Y		
Els1	Netherlands	2.520					
Fir1	Netherlands	3.150				Y	
Fis1	Scotland	2.930					
Fis2	Scotland	2.690					
Had1	England	1.460					
Hou1	Netherlands	2.560				Y	
Hou2	Netherlands	2.630				Y	
Kil1	Scotland	1.930					
LISM01	Scotland	0.000					
LISM03	Scotland	0.010					Y
Lym1	England	4.467					
Lym2	England	2.970				Y	

Mam1	Ukraine	0.410	Y		Y		
Man1	Germany	0.100					
Man2	Germany	0.440					Y
Man3	Germany	1.060				Y	
Man4	Germany	1.040				Y	Y
Man5	Germany	0.520					
Man6	Germany	1.050					
New2	Ireland	3.930					
Otb1	Netherlands	2.580				Y	
Poi1	France	3.900				Y	
Poi2	France	1.390				Y	
Poi3	France	0.830					
Poi4	France	1.980					
Poi5	France	0.380				Y	
Pot1	England	2.550					
Pot4	England	2.640					
Red1	UK	1.680					
Red2	UK	3.310				Y	
Rug1	UK	9.140				Y	
Snu1	Scotland	3.420				Y	
Ste1	Netherlands	0.950					
Tar1	Italy	0.009					
Tar2	Italy	0.009					
Viz2	Netherlands	5.750				Y	
Wal1	UK	0.940				Y	
Win1	Netherlands	18.690				Y	
Wro1	UK	1.630					
Yor1	England	7.850				Y	
Yor11	England	0.980					
Yor12	England	4.480					
Yor2	England	3.910				Y	

Yor3	England	2.500					
Yor4	England	3.140				Y	
Yor5	England	2.960				Y	
Yor6	England	6.810				Y	
Yor7	England	6.630				Y	
Zha1	Ukraine	0.390	Y		Y		
Zha2	Ukraine	0.020			Y		

4.4.1.5. The indicine arrival in South-Western Asia

As seen with the first PCA (Figures 21 and 22), taurine and indicine cattle plot on opposite sides of the graph. The two taxa can thus be differentiated, and hybrids identified. This is the case for several individuals who plot in the middle of the PCA, revealing indicine introgression in Bronze Age and later samples in the Levant, Iran, Anatolia, the Caucasus and Central Asia. D-statistics (Figure 26) and admixture (Figure 30) also highlight this introgression in the Near-East from the Bronze Age, and specifically from *Bos indicus* originating from the Indian subcontinent, and not from China.

This was already found in Verdugo et al. (2019), with an arrival of indicine hybrids from ~4,000 years ago detected in Central Asia, Iran and the Levant, but only from the post-Iron Age in the Caucasus. In the latter we find it in older samples, one dating from the 3rd millennium BCE, and one from the Iron Age. We hence highlight the indicine progression westwards earlier than previously described, even suggesting an indicine signal from D-statistics in Eastern Europe, though not significantly, in 3 samples (one Bronze Age and two Iron Age). The two Iron Age samples that are not significant are plotting away from European samples on the PCA, towards the hybrids' area.

4.4.2. African cattle ancestry

Verdugo et al. (2019) used D-statistics to establish that aurochs introgression had occurred to differentiate later Levantine cattle from Neolithic Turkish herds. The best proxy for this source was an African aurochs from pre-Neolithic Morocco, probably indicating some ancestral concordance between this and wild stock in the eastern

Mediterranean. In ADMIXTURE analysis, an African genomic component defined by modern data is also the largest one involved in the two samples from Tunisia dating from the Iron Age (Figures 30, 31) These samples also have unusually extreme D-statistics values (Figure 27), consistent with ancestry varying from that in the majority of the ancient data set. However, this ADMIXTURE component also features in Near Eastern individuals according with Verdugo's suggestion of common ancestry in local wild strains that contributed to both Levantine and later to African herds.

5. Ancient cattle populations in the Middle East

5.1. Introduction

The Fertile Crescent saw the emergence of animal domestication during the Neolithic. *Bos taurus* was domesticated around 10.5 kya from a small population of *Bos primigenius* (Bollongino et al., 2012; Rossi et al., 2024), and from then on spread to other parts of the world. But little is known about the genetic profile of ancient cattle in Iran and close areas, where DNA is difficult to extract due to the arid environment. Past research has relied on mitochondrial DNA to study the region (Bollongino et al., 2012; Scheu et al., 2015), with only limited data available for whole genomes (Verdugo et al., 2019).

Within this project, new Near Eastern samples, mainly from Iran and Syria, were sequenced and previous ones were improved (see Chapter 2). In this chapter, these samples will be discussed in more detail than in Chapter 4, in order to focus on ancient populations from the Middle East including beyond the Fertile Crescent. One particular analysis (qpWave) will focus on early samples from the Early Bronze Age, Chalcolithic and Neolithic, that pre-date indicine introgression in South-Western Asia, to avoid bias arising from this admixture. Using PCA, D-statistics, admixture and qpWave modelling, we show that early *Bos taurus* of the Middle East were not a uniform population, despite the closeness to the domestication centre of origin.

5.2. Material and methods

5.2.1. Samples

5.2.1.1. The Middle East and Central Asia region

Samples described in Chapter 4 were filtered based on their geographic origin. For the purpose of this chapter, the following regions were targeted for analyses: Anatolia, Caucasus, Cyprus, Levant, Lebanon, Saudi Arabia, Syria, Iraq, Iran and Turkmenistan. They date from the Neolithic to the Middle Age and represent a total of 97 individuals.

5.2.1.2. Early populations

A total of 41 samples were used to study the early times in the region. They represent a period from the Neolithic to the Early Bronze Age, and come from various sites across 6 regions of the Middle East:

- Anatolia = 13 samples (including one *Bos primigenius*) from 6 sites;
- Caucasus = 3 samples (2 from the Northern Caucasus and one from the South);
- Iran = 13 samples from 9 sites;
- Turkmenistan = 7 samples, 1 site;
- Syria = 3 samples from 2 sites;
- Lebanon = 1 sample;
- Levant = 1 sample.

Samples with insufficient coverage were filtered out; in order to obtain a good representation of the region's genetic diversity, a minimum of 0.1X coverage was decided. However, to be able to represent various sites, a coverage as low as 0.08X was accepted for 3 samples: Sto3 (0.08X, Bronze Age Northern Caucasus), Cuk1 (0.09X, Neolithic Western Anatolia) and Ulu5 (0.09X, Neolithic Western Anatolia). The highest coverage is 13.85X for Sub1 (Neolithic Anatolia), but most of the samples have a coverage between 0.1 and 2.9X. The final list of samples is described in Table 16 below. All individuals but one are of *Bos taurus* species. The only exception is Ch22, an aurochs from Chalcolithic Central Anatolia, kept in the test due to its close temporal and geographical location to other domesticates.

Table 16. Samples from the Middle East dating at least from the Early Bronze Age. Only samples with a coverage > 0.08X were kept.

Name	Region	Site	Context	Dates	Coverage
Ace1	Anatolia	Acemhöyük	Bronze Age	2500 BC	0.310
Bol1	Iran	Darre-ye Bolaghi	Neolithic	7000-5000 BC	0.690
Ch22*	Anatolia	Çatalhöyük	Chalcolithic	7750-7550 cal. BP	0.310
Cuk1	Anatolia	Çukuriçi Höyük	Neolithic	6000-7000 BC	0.090

Cuk2	Anatolia	Çukuriçi Höyük	Neolithic	6000-7000 BC	0.440
Dav1	Iran	Dava Goz	Chalcolithic	5000-4500 cal. BC	1.000
Dav2	Iran	Dava Goz	Chalcolithic	4200-3800 cal. BC	1.730
Far1	Iran	Fars	Neolithic	6000 -5500 BC	0.110
Has3	Iran	Hasanlu	Bronze Age	2470 cal. BC	1.500
Men1	Anatolia	Mentese	Neolithic	6400-5700 cal. BC	2.920
Men2	Anatolia	Mentese	Neolithic	7998-7843 cal. BP	1.850
Men3	Anatolia	Mentese	Neolithic	6400-5700 cal. BC	0.330
Mon10	Turkmenistan	Monjukli Depe	Chalcolithic	Ca. 5100-4500 cal. BC	0.740
Mon4	Turkmenistan	Monjukli Depe	Chalcolithic	4492-4349 cal. BC	0.490
Mon5	Turkmenistan	Monjukli Depe	Chalcolithic	Ca. 5100-4500 cal. BC	0.310
Mon6	Turkmenistan	Monjukli Depe	Chalcolithic	Ca. 5100-4500 cal. BC	0.220
Mon7	Turkmenistan	Monjukli Depe	Chalcolithic	Ca. 5100-4500 cal. BC	0.440
Mon8	Turkmenistan	Monjukli Depe	Chalcolithic	Ca. 5100-4500 cal. BC	1.270
Mon9	Turkmenistan	Monjukli Depe	Chalcolithic	Ca. 5100-4500 cal. BC	0.460
Nad1	Iran	Naderi Tepe	Bronze Age	2800-2500 BC	0.300
Nad2	Iran	Naderi Tepe	Bronze Age	3000-2000 BC	0.310
Qaz1	Iran	Qazvin	Bronze Age	2882-2636 cal. BC	1.650
Qaz2	Iran	Qazvin	Chalcolithic	3700-3400 cal. BC	0.260
Sac3	Iran	Tapeh-Sang-e-Caxmaq	Late Neolithic	5373 cal. BC	0.190
Shu1	Caucasus	Shulaveris Gora	Neolithic	5800 cal. BC	1.530
Soh1	Iran	Soha Chai Tepe	Chalcolithic	4200-3900 BC	0.490
Sto1	Caucasus	Stoderovskoe 1	Bronze Age	3900/3800-3500/3400 cal. BCE	0.450
Sto3	Caucasus	Stoderovskoe 1	Bronze Age	3900/3800-3500/3400 cal. BCE	0.080

Sub1	Anatolia	Suberde	Neolithic	6221-6024 cal. BC	13.850
Tap1	Iran	Tappeh Hesar Markazi	Bronze Age	3000-2000 BC	0.940
Tef1	Lebanon	Tell Fadous	Bronze Age	3000-2000 BC	0.290
Tqa2	Levant	Tel es-Qashish	Bronze Age	3150-2200 BC	0.560
Ulu1	Anatolia	Ulucak	Neolithic	7000-5700 BC	0.190
Ulu2	Anatolia	Ulucak	Neolithic	7000-5700 BC	0.460
Ulu3	Anatolia	Ulucak	Neolithic	7000-5700 BC	0.550
Ulu4	Anatolia	Ulucak	Neolithic	7000-5700 BC	0.500
Ulu5	Anatolia	Ulucak	Neolithic	7000-5700 BC	0.090
VEM329	Syria	Tell Chuera	Early Bronze Age	2100-2000 BC	0.470
VEM330	Syria	Tell Chuera	Early Bronze Age	2100-2000 BC	0.370
VEM331	Syria	Tell Sheikh Hassan	Chalcolithic	3300-3600 BC	0.280
Zag1	Iran	Zagheh Abyek	Chalcolithic	4800-4300 BC	0.210

5.2.2. Projection PCA

Using the pipeline as described in Chapter 4, ancient samples were projected on a reference dataset of modern individuals (Figure 21). For the purpose of this chapter, a close-up view on the region of interest was made using R, allowing the observation of samples of Iran and the surrounding areas.

5.2.3. D-statistics

To assess indicine introgression, D-statistics were computed from bam files with ANGSD v0.941 (Korneliussen et al., 2014), using parameters as follows:

```
-doAbbababa 1 -doCounts 1 -rmTrans 1 -trim 4 -minMapQ 30 -minQ 20
```

A modern water buffalo was used to define ancestral alleles. The individual Sub1, from Neolithic Anatolia, was used as reference for early domesticates, while the modern

individual Sha3b (Sahiwal breed) was used as reference for indicine cattle. The number of ABBA and BABA sites between these two references and the sample were counted and the final D-statistics computed for individuals with at least 500 ABBA+BABA sites, as described in Chapter 4. Results were plotted with RStudio v2023.03.0+386.

To plot D-statistics by time, an average date was used for the time period based on archeological information.

5.2.4. Admixture

The admixture analysis run in Chapter 4 is re-used here but modified to zoom in the area targeted in this chapter (the Middle East; from Anatolia to Turkmenistan, Caucasus to Saudi Arabia).

5.2.5. qpWave

With the R wrapper Admixtools2 (Maier et al., 2023), qpWave was run on the early samples described in Table 16. Using various aurochsen (from Europe, Caucasus and Central Asia), zebus and water buffalos as "right" populations, the software processed each combination of samples against each other. The "right" populations are:

- Modern Water Buffaloes (8 individuals, Dutta et al., 2020);
- Modern *Bos indicus* from India/Pakistan (7 individuals, Chen et al., 2018 and Koufariotis et al., 2018);
- European aurochsen: Bed3, Bed4, Fif1, Hjo1, Hxh2, Rhi3, Ska1, Ska3, Tri1, Zea1, Zea2;
- Rhine aurochsen: Rhi1 and Rhi2;
- Eastern aurochsen: Baikal1 and Tula1;
- Gyu1 and Gyu2 individually.

Ancient samples used in the "right" populations are included in the Supplementary Table S4.

A p-value above 0.05 (H_0 is not rejected) indicates a clade may be formed including the targets to the exclusion of the outgroup or "right" populations. A lower p-value would

indicate an admixture event from a right population into the target population, breaking their clade.

A symmetric heatmap of the p-values was compiled with Rstudio v2023.03.0+386 and the package “pheatmap” to show the results in the Early Middle East, allowing assessment of the presence or not of an admixture event from another species into the area in the Neolithic/Early Bronze Age. Colors in the figure indicate the p-value, with blue for p <0.01 (H_0 is rejected), shades of white-yellow for 0.01 ≤ p ≤ 0.05 and orange-red for values >0.05. The lower threshold 0.01 instead of the commonly used 0.05 is shown on the figure to allow interpretation of results otherwise hidden by a stricter limit.

In order to observe clusters within samples, p-values were converted into a distance matrix using a negative base-10 logarithmic transformation $d_{ij} = -\log_{10}(p_{ij})$, where d is the distance between samples i and j, and p is the p-value for the pair (i; j). The resulting distance matrix was used to perform hierarchical clustering on R. The heatmap was plotted using original p-values for visualisation, and transformed values for ordering.

5.3. Results

5.3.1. Principal Component Analysis (PCA)

The projection of ancient samples onto a reference plot of modern samples is shown in Figure 35. Taurine samples are on the left of the figure and indicine on the right. Interesting samples are labelled in black, while blue labels indicate samples identified in Chapter 3 that have the haplogroup T1.

Ancient Iranian samples from the Neolithic to the Bronze Age are localised in this close-up. More recent samples plot on the right of the PCA, towards the indicine side. This cluster of ancient Iranians individuals also encompasses samples from the same periods but from several other areas, such as Anatolia, the Caucasus and Central Asia (more specifically Turkmenistan). Several Syrian (labelled as Middle East in the figure: VEM329, VEM330 and VEM331) samples are located near this cluster, but lower on the y-axis, oriented towards ancient and modern African samples. Next on this axis are found a Bronze Age sample from Lebanon, Tef1, and a Neolithic sample from Saudi Arabia, Aak1, clustering

with Bronze Age and Neolithic Levant and Iron Age Tunisia. Younger Levantine samples are seen on Figure 35 in the middle of PC1, half-way between taurine and indicine samples.

The series of samples from the Middle East to Africa gathers most of the individuals characterised by the haplogroup T1, found in Bronze Age Tell Chuera (Syria, VEM329), Tel es-Qashish (Levant, Tqa2/Tqa3), Neolithic Tel Hreiz (Levant, Thr1), Al'Ula (Saudi Arabia, Aak1), Iron Age Tel es-Safi (Levant, Tsa3), Althiburos (Tunisia, STK samples) and Post-Iron Age Wakarida (Ethiopia, VEM337). The other 5 are younger Levantine samples plotting in the middle of the figure, who come from the sites of Tel Miqne-Ekron (Tmq4), Tel Masos (Mas1), Tel-el-Safi (Tes1), Horvat Castra (Cas1) and Tel es-Safi (Tsa2).

In South-West Asia, samples plot in the middle left side of Figure 35. But the central cluster does not gather the entirety of ancient samples from this area. Indeed, Anatolian and Caucasian samples are spread from the Eastern and Central Europe cluster to the ancient Iranian group. Earlier samples from Anatolia are closer to European samples, while the aurochs Ch22 and a Bronze Age individual (Ace1) are plotting in the Iranian cluster.

Two groups can be seen for the Caucasus; a first one near European and early Anatolian samples, and a second in the early Iranian - later Anatolia. The first group includes three Bronze Age samples from the Northern Caucasus, the second one consists of three samples from the site of Stoderovskoe in the Northern Caucasus and one sample from Shulaveris Gora in the Southern Caucasus. A difference between the Bronze Age samples can be noted thanks to the archeological information associated with Dam1 and Ipa1, the two samples located at the top of the figure, that are younger than other Bronze Age samples from the North Caucasus located in the center of Figure 35. The sample located on the left, close to early Anatolian, would be slightly older than the two at the top, but not as old as the four in the center of the figure.

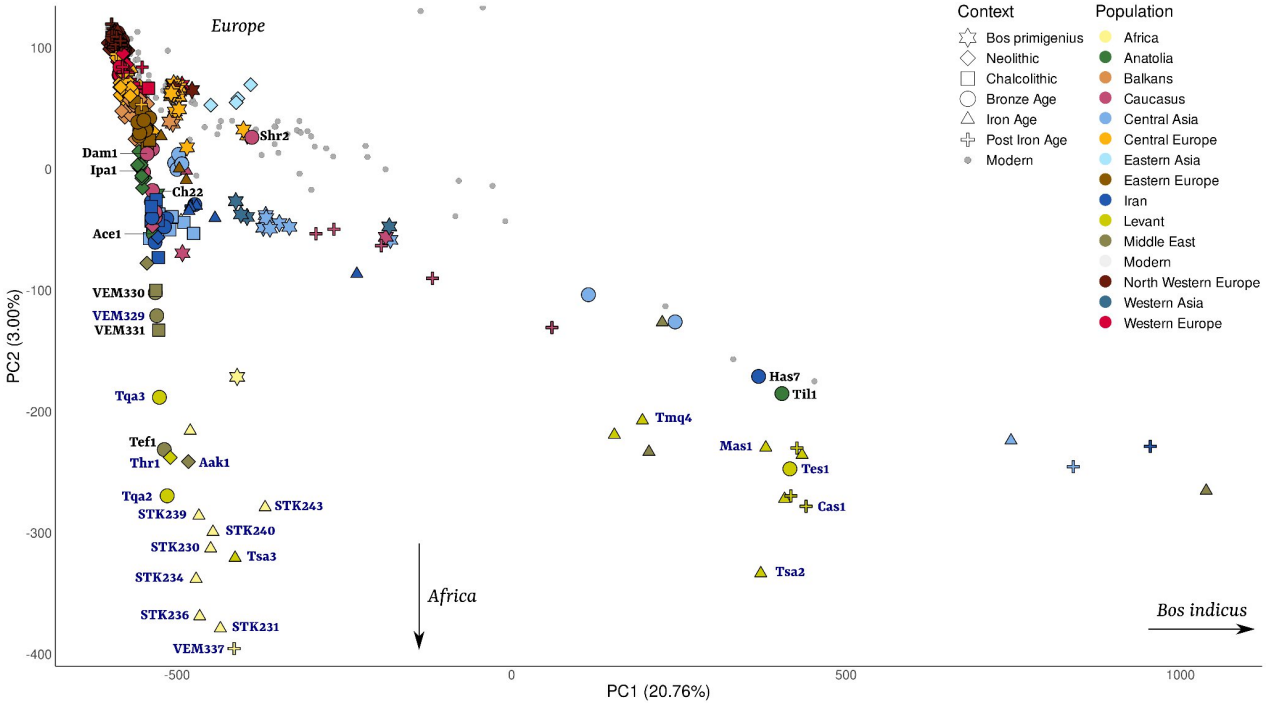

Fig. 35. Procrustes projection of ancient *Bos* samples on a modern reference PCA, focused on ancient samples. The PCA from Chapter 4 Figure 22.B) was adapted to indicate samples with mtDNA haplogroup T1 (blue sample labels), while black labels indicate samples of interest that are not T1. Modern African samples are located outside of this figure, lower on PC2. Modern *Bos indicus* samples are on the right side of PC1.

5.3.2. D-Statistics for indicine introgression assessment

Indicine introgression was assessed via the test *D(outgroup, test sample; Sub1, Sha3b)* in the 97 ancient samples from South-Western Asia, and results were filtered to keep individuals where nABBA+nBABA ≥ 500, leading to 83 remaining samples. Results were plotted in Figure 36 (values are listed in Supplementary Table S9). Samples are ordered by region (Anatolia, Caucasus, Iran, Central Asia, Middle East and Levant) and time (Neolithic, Chalcolithic, Bronze Age, Iron Age and Post Iron Age) from top to bottom. Values are simply plotted versus archaeological time in Figure 37. Significant values are indicated with thicker edges for samples that have |Z-score| > 3. An average date was used for samples with indirect dates, based on archeological information.

Samples that date from the Neolithic and Chalcolithic periods tend to have a D-statistic close to 0, indicating no zebu introgression. The only exception is a sample from

Syria that has a value of 0.079, but with only 667 ABBA+BABA sites and a non-significant Z-score (Figure 37).

After the Chalcolithic, samples have higher D values, especially Iron Age and later time periods. The Bronze Age however shows a range of D values from negative to high positive ones, though most of them are showing the same pattern as Neolithic and Chalcolithic samples, with a few exceptions. A first group is seen with three samples from Anatolia, Iran and the Levant plotting on the right side of Figure 36, with significant high D values (Til1, Has7 and Tes1). A second group of Bronze Age samples is made of Shr2 from the Caucasus and Has1 from Iran, plotting at D = 9.388 and 6.440, and appear significant on Figure 37.

Indicine introgression is clearly detected from the Iron Age in South-Western Asia, with every sample tested resulting in high D-statistics (at least 0.043) and significant Z-scores. However, among the younger samples with positive D-statistics values, three groups can be identified (horizontally on Figure 36 and vertically on Figure 37). First, Iron Age and post-Iron Age samples from Iran have lower values than other samples from the same periods, between 0.043 and 0.096. Two Bronze Age and one Iron Age samples from the Caucasus and Iran are found in this group. The second group has medium values, and is characterised by Medieval samples from the Caucasus, and one Iron Age sample from Western Iran is plotting in the same area. In the last group, on the right side of the plot, are found the highest D-values, in samples from the Levant and Syria-Lebanon, as well as two Bronze Age from Anatolia and Iran, and a more recent Iranian sample.

These three groups are easily identified in Figure 37, and all samples within them have significant positive values for indicine introgression. The oldest samples are Tes1 and Til1, though for Til1 the age was approximated following site chronology established by archeologists (Berthon and Mashkour, 2008), and an average date was used to plot the sample.

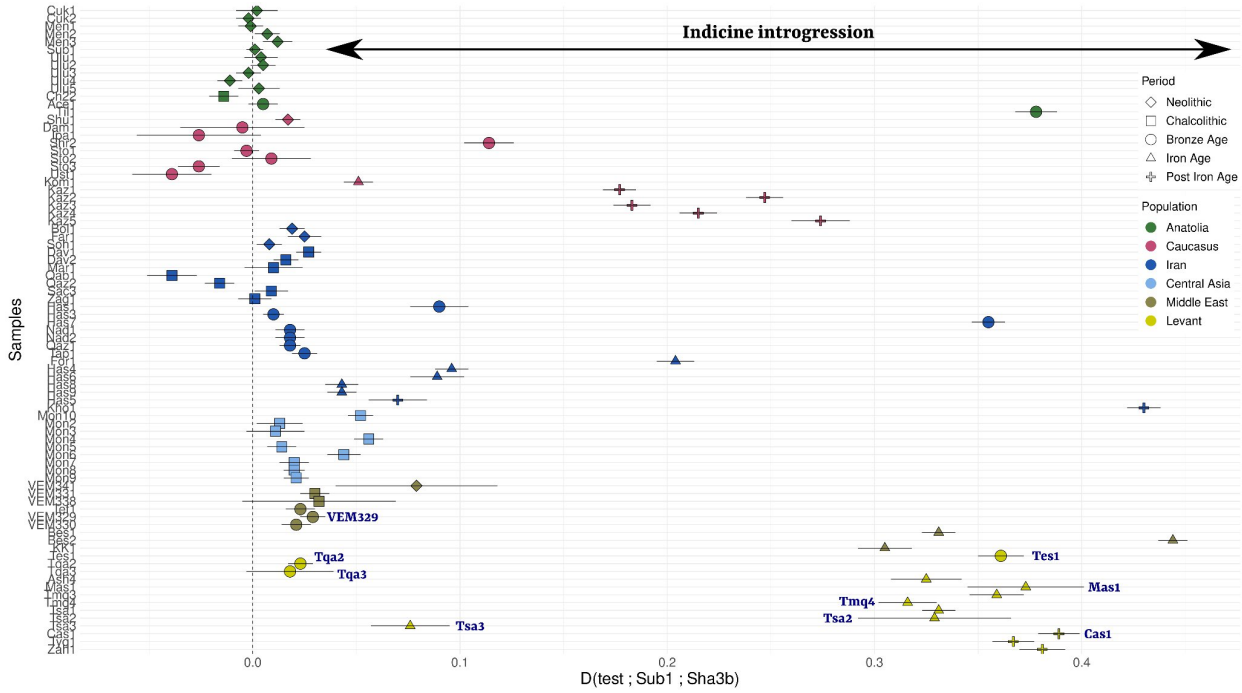

Fig. 36. Indicine introgression assessed via D-statistics in 83 ancient samples from the Middle East. A positive D value means introgression is detected in the sample's ancestry. Standard errors are represented by a bar on both the left and right sides of each sample. Samples were grouped by region and ordered by time within each region, from oldest to youngest from top to bottom, as also shown by shapes denoting archaeological era. T1 samples are labelled in blue next to their D value.

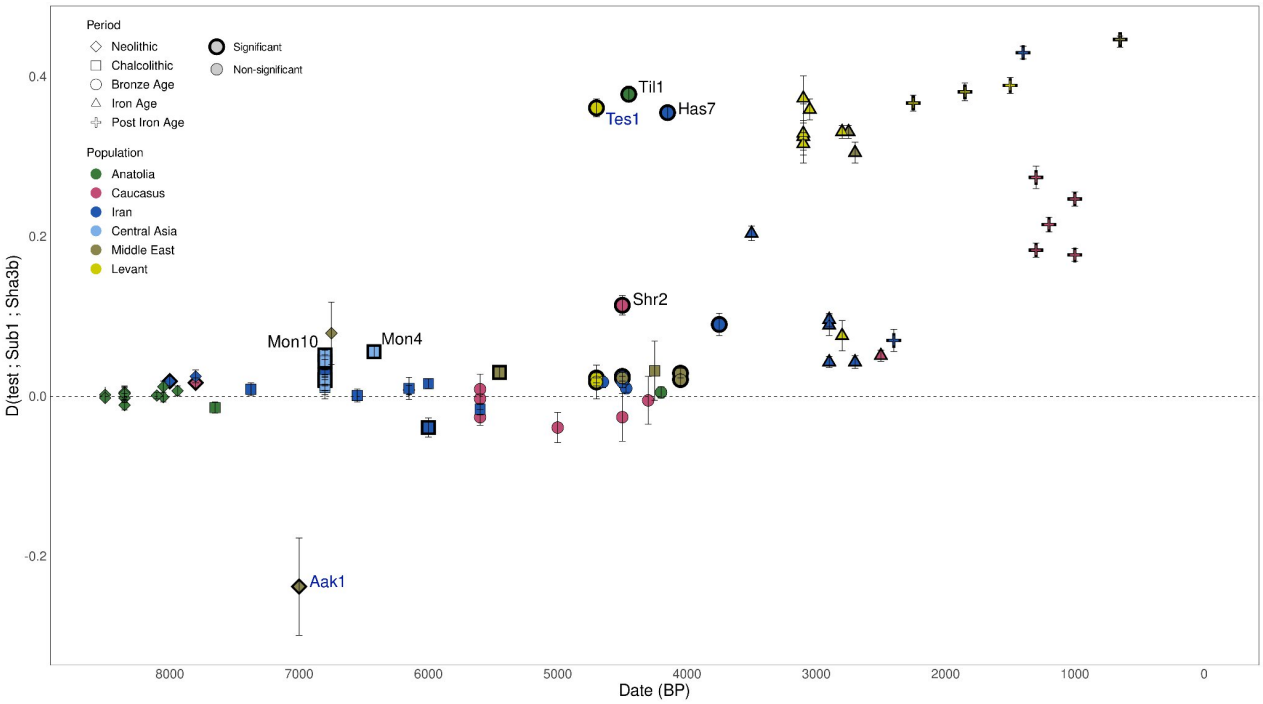

Fig. 37. Indicine introgression through time assessed via D-statistics in 83 ancient samples from the Middle East. A positive D value means introgression is detected in the sample ancestry. The sample Aak1 was added, even though it has a low number of ABBA+BABA sites, as it represents a valuable individual both geographically (the only sample from Saudi Arabia) and temporarily (Neolithic). Samples of interest are labelled, with T1 in blue.

5.3.3. Admixture analysis in South-Western Asia

The admixture analysis shown in Chapter 4 (Figure 30) was modified to only show the Middle East in Figure 38. The three panels A, B and C depict different time periods from oldest in A to most recent in C. As described in Chapter 4 section 4.3.3 “Ancestries assessment with Admixture”, 6 components were identified in the analysis, belonging to wild, taurine and indicine clusters from Europe and Asia.

When we look at earlier samples (Figure 38.A), South-Western Asia is characterised by a predominance of the taurine component labelled as “African/Levantine Taurine”, representing at least 50% of the ancestry in every sample of the region. This component is fixed in modern African *Bos taurus*. A component typical of modern European cattle gives the rest of the ancestry composition. A minority of samples, found in Turkmenistan, present a minority of Eastern aurochs component (yellow).

In the 5 kya to 2.7 kya time period (Figure 38.B), the Eastern aurochs component becomes more present, reaching Iran, the Caucasus, Anatolia, and is found as a very small portion of one Levantine sample. Also, the Indicine component appears for the first time in the region. This component, identified in modern *Bos indicus* from the Indian subcontinent in Figure 31 (Chapter 4), is found in most samples present in panel B, in Iran, the Levant, Anatolia and the Caucasus.

In the most recent samples, from the Iron Age and later in Figure 38.C), the indicine component takes a bigger portion of ancestry in samples where it is present, in Iran, Iraq, the Levant, Cyprus and the Caucasus. The Eastern aurochs component is still present, mostly in the Caucasus, but can be found in the Levant, where it was barely noticed before. A small portion of this Eastern component is also present in the sample from Cyprus.

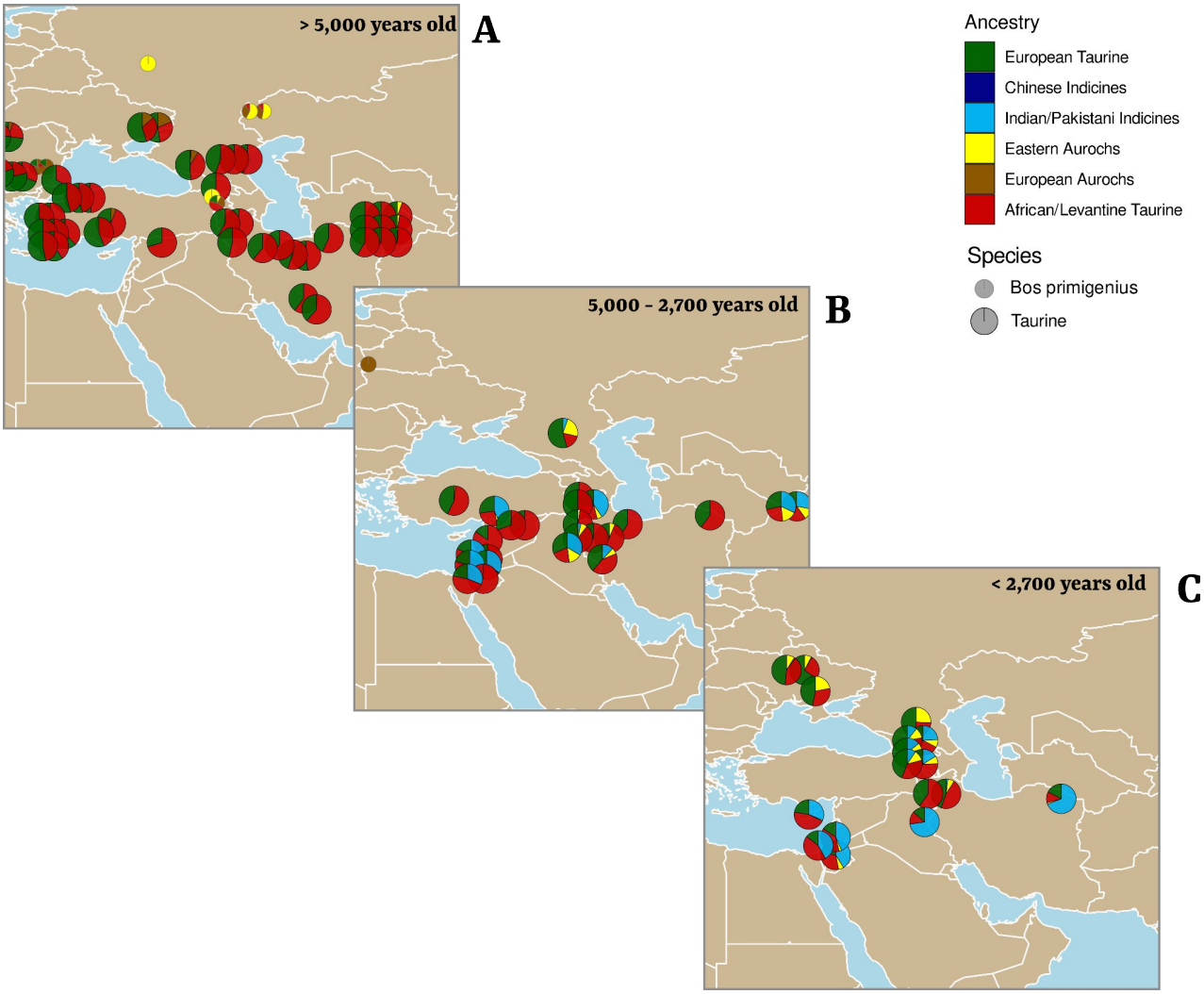

Fig. 38. ADMIXTURE results with K = 6 on domesticates and wild samples of Eurasia, centered in results from South-Western Asia. A) Samples older than 5,000 years; B) Samples aged between 2,700 and 5,000 years; C) Samples younger than 2,700 years.

5.3.4. Early Middle East structure estimated via qpWave

A qpWave analysis was designed to test whether earlier domesticated samples in the data set formed clades with each other that were undisturbed by secondary introgression from a range of candidate exotic ancestries. To perform this analysis, ancient samples were filtered according to specific criteria. They must:

- ❖ have at least a 0.08X coverage;
- ❖ date at least from the Early Bronze Age (Table 16);

- ❖ have no indicine introgression detected via D-statistics and Admixture analysis.

Using the Figures 36 and 37 for D-statistics-detection of indicine introgression, samples fitting the above criteria were then filtered out based on their D values. The lowest D-value seen in any Iron Age sample (Has8 and Has9, D=0.043) was used as a limit to choose which samples to exclude from the qpWave analysis, as this time period is seen to behave differently on the PCA (Figure 35 and 39), pointing towards hybridization. However, four exceptions were made for the samples Mon4, Mon6, Mon10 and VEM341 due to their age (Chalcolithic for Monjuklu and Neolithic for Syria), as well as the absence of indicine component detected with the Admixture analysis (Figure 38) for Monjuklu samples. VEM341 could not be included in the Admixture due to its low number of sites in the pseudohaploid dataset.

The remaining samples are labelled on the PCA on Figure 39, which is adapted from Figure 35, and plotted on a map in Figure 40. They include 13 Anatolian samples (12 domesticates and 1 wild individual), 3 from the Caucasus, 13 from Iran, 7 from Turkmenistan (all from the same site of Monjuklu), 3 from Syria, 1 from Lebanon and 1 from the Levant. On the PCA, early Anatolian samples plot on the left side, close to Eastern European samples and a group from the Caucasus. One Bronze Age Anatolian, along with an aurochs, are clustering with Neolithic to Bronze Age Iran, and another group from the Caucasus. Samples from Monjuklu, Turkmenistan are spread from the Iranian cluster in the center to the right side of the plot, towards indicine hybrids.

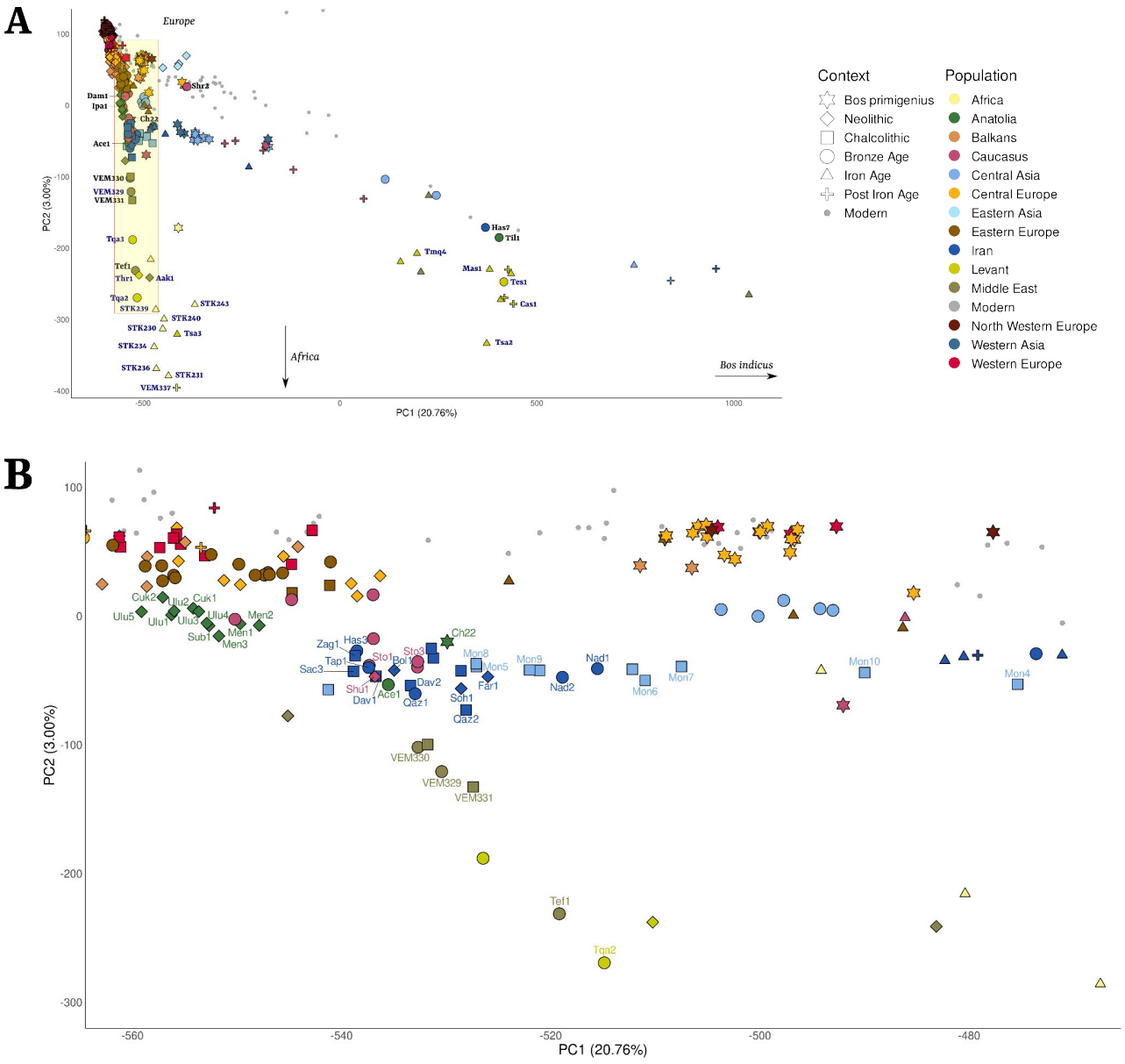

Fig. 39. Procrustes projection of ancient *Bos* samples on a modern reference PCA. A) Figure 35 is the original PCA with a highlighted area to show the close-up view in the bottom. B) Close-up view of the area highlighted on the top figure. Samples included in the qpWave analysis are labelled and colored according to their geographic origin.

Because they represent the earliest samples from the Middle East, and plot in several distinct regions on the PCA (Figure 39), they were analysed with qpWave. In this test, we hypothesize that no populations from the outgroup are involved in the ancestry of two individuals tested simultaneously. If the test turns out to be not significant (P-value > 0.05), this is consistent with the two samples tested forming a clade relative to the

outgroups. The six outgroups included were representatives of modern water buffalo and *Bos indicus*, European Holocene aurochsen and pre-LGM aurochsen, East Asian aurochsen and two Armenian aurochs, one pre- and one post-Holocene. An African aurochs (Verdugo et al., 2019) could not be included because of low coverage but it was hoped that the combined diversity of other included outgroups could aid in detection of influx of this ancestry. The qpWave results are conveyed in Figure 40 by a heatmap matrix showing the test's P-value for each pairwise combination of the 41 samples, with a scale from blue to yellow and red indicating the significance of the test. Blue indicates significant values (P-value < 0.01), implying that the samples need at least one additional admixture event from one outgroup population to explain their relationship, while orange and red indicate non-significance, consistent with the two studied samples forming a cluster. A lower threshold for significant P-value is indicated in shades of white to yellow, when the test leads to a P-value between 0.01 and 0.05.

The tested samples from the Early Middle East are ordered by location in the symmetric heatmap in Figure 40, with Anatolian samples on the top and left, followed by the Caucasus, Iran, Turkmenistan, Syria, Lebanon, the Levant and finally the only aurochs tested as target, Ch22, identified by its large size in a Neolithic domestic context in Anatolia (Verdugo et al., 2019). Geographic position is also indicated in more detail by the map below the heatmap. There is within-country diversity in sampling with sites in Western and Central Anatolia, the North and South Caucasus, and Western, Central, Eastern and Southern Iran.

Two main clusters of cladal pairwise (i.e. non-significant) values can be identified in Figure 40. The first is located on the top left corner of the heatmap, encompassing Anatolia, the Caucasus and most of Iran. Only the two Eastern samples Nad1 and Nad2 are not part of this cluster, but show a very high P-value when tested with Monjuklu samples from Turkmenistan, and high values with Syria, Lebanon and the Levant. A second cluster is located on the bottom right of the heatmap, with individuals from Syria, Lebanon, Israel and Turkmenistan (but with the exceptions of Mon4 and Mon10). Nad1 and Nad2 seem to be included in this cluster, but intermediate as they also present high P-values with other samples from Iran.

Outside of the two main clusters, some samples show higher values with geographically close individuals. The four samples from Lebanon and Syria, which are quite central on the map among the dataset, present high P-values with Iran, the Caucasus and Anatolia, and have lowest values with Western Anatolian samples. Tqa2 (from Israel, Tel es-Qashish) is one of the samples with low P-values in most of the tests. It has higher values with only a few samples from Lebanon and Syria, which are the closest to Tqa2 geographically, and Sto3 from the Caucasus, Nad2 and Mon5/Mon6/Mon9. Two Monjukli samples, Mon4 and Mon10, seem to require admixture events to explain their p relationships with other samples, it is notable that these are the furthest rightward on the PCA, suggesting that this may be because of a degree of indicine ancestry.

The aurochs Ch22, which clusters with ancient domesticates on Figure 39, shows some affinity with several Anatolian samples, the Caucasus (via Sto3) and some samples from Western and Central Iran.

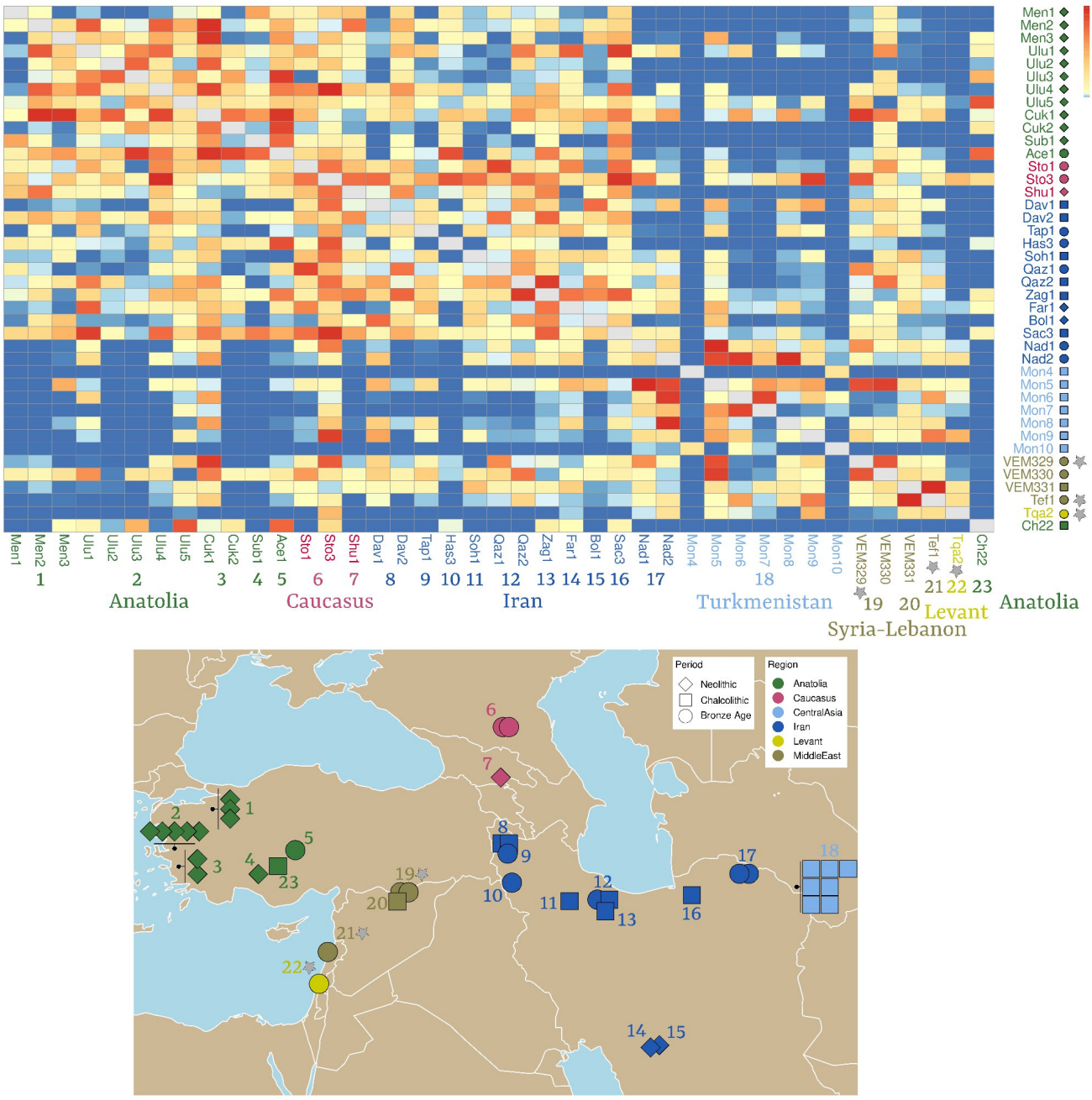

Fig. 40. Heatmap of p-values calculated with qpWave on samples from the Early Middle East. Only samples dating from Neolithic to Early Bronze Age, with a coverage ≥ 0.08X, were selected for this analysis. Ch22 is the only Bos primigenius tested and was placed at the extremity of the heatmap. The map associated shows geographic origin of the samples, with numbers indicating the archeological site. Samples that have haplogroup T1 are indicated with a grey star-shaped symbol.

The heatmap from Figure 40 was rearranged on Figure 41 after performing a hierarchical clustering of the samples based on their p-values (after transformation). This led to the appearance of 4 clusters, identified on Figure 41 by numbers on both axes. The main cluster observed previously on Figure 40 forms now the clusters 1 and 2 on Figure 41 with Anatolian, Caucasian and most Iranian samples. Individuals from Syria, Lebanon, Israel and 5 out of the 7 samples from Turkmenistan are indeed clustering together, as previously, with Nad1 and Nad2, while the remaining 2 Monjuklu samples (from Turkmenistan) form the fourth cluster.

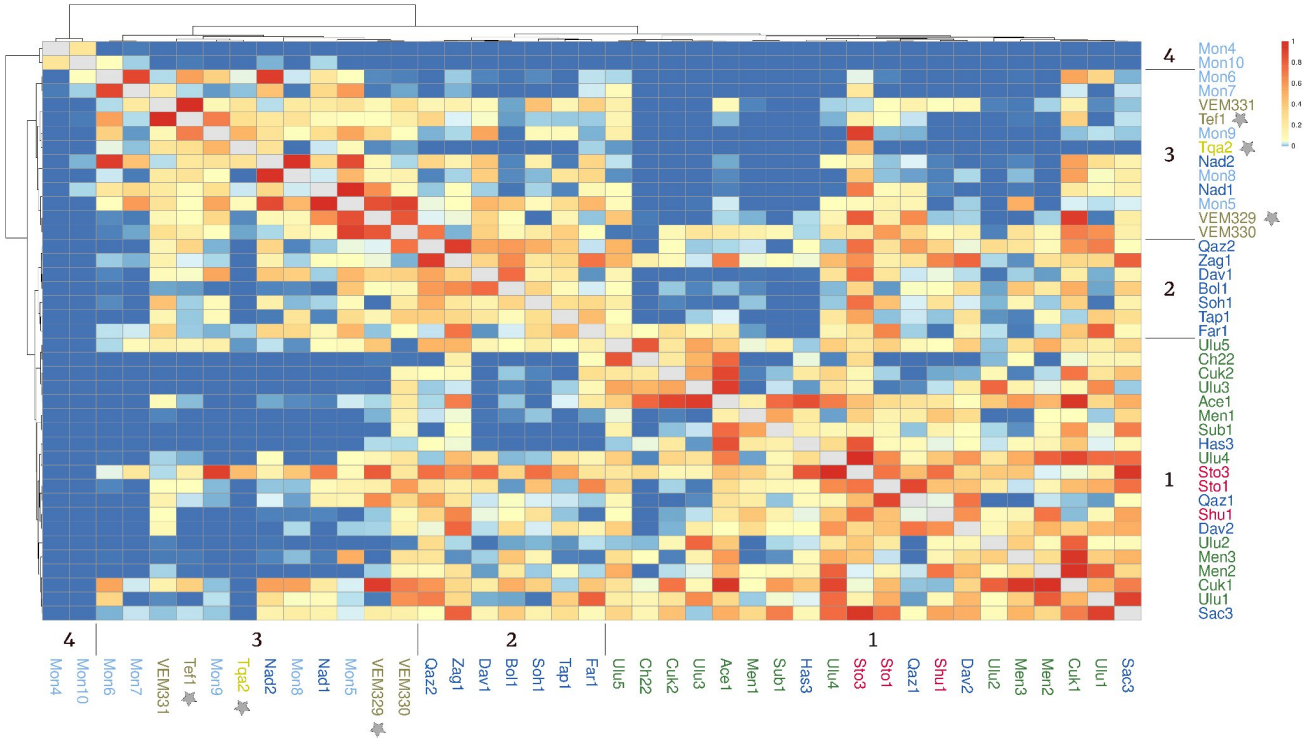

Fig. 41. Heatmap of p-values calculated with qpWave on samples from the Early Middle East, ordered after hierarchical clustering (indicated by numbers on both axes). Labels are colored according to the geographical origin of the samples, corresponding to the map from Figure 40. Star-shaped symbols indicate T1 samples.

5.4. Discussion

5.4.1. Early indicine introgression in South-Western Asia

Verdugo et al. (2019) in their survey of 67 ancient Near Eastern cattle genomes discovered massive and widespread indicine introgression throughout the region which

was observable from ~4,000 years BP. The addition of new data from the region allows the testing of this finding, in particular resolution of the time at which migration from the east occurs. Figure 35 shows a subset of Near Eastern cattle plotting in intermediate PCA positions between *Bos taurus* and *Bos indicus*, and ADMIXTURE analysis also indicates the presence of indicine ancestry within the region (Figure 38), especially from the Iron Age. Figure 39 plots the timing of evidence for zebu admixture and the general pattern holds, even though two samples seem a little older than this time threshold. These are Til1 and Tes1, from Anatolia and the Levant respectively, and are not directly dated. Tes1, a T1 sample, is from Tel-el-Safi which is an Early Bronze Age site in the Levant, occupied between 2,850 to 2,600 BC (Arnold et al., 2018). Til1 is a sample from Tilbeshar, which dates to the transition between Chalcolithic and Bronze Ages for Anatolia (3,300-1,600 BC, Berthon and Mashkour, 2008).

A few other samples showed introgression on Figure 39, earlier than for Til1 and Tes1, but the validity of this result is questionable. Indeed, three Monjuklu samples, Mon4, Mon6 and Mon10, would represent earlier traces of indicine introgression, as well as the sample from Tell Seker VEM341 (Neolithic Syria). But indicine admixture was not detected in Monjuklu (Figure 38) during the Chalcolithic, and could not be verified in Tell Seker, as VEM341 did not yield enough DNA for the test. Furthermore, indicine introgression in Neolithic or Chalcolithic sites in South Western Asia has never been described. Hence, only Til1 and Tes1 can be presented here as early hybrids. Another interesting case remains within Sharakhalsun, a site in the North Caucasus, with a lower D-value seen in Shr2 than in Til1 or Tes1 (Figure 39), and admixture detected in Figure 38. Other samples from the same region and time period do not show indicine traces; it is possible that Shr2 is slightly more recent than its neighbours from Stoderovskoe.

The work in this thesis has added 11 samples showing admixture from later sites and 38 samples without admixture in earlier sites. There does seem to be a sharp increase in admixture through cattle migration from the East just before 4,000 years ago, a process that may be linked to the well known 4.2 ka climate episode which gave a period of increased aridity detectable in several regions across South-Western Asia and which had

far reaching consequences, including empire collapses in Mesopotamia and Egypt (Weiss et al., 1993; Matthews 2002; Wiener 2014).

There are several points to be made about this admixture. First, the lack of zebu influence prior to 4,000 years ago might be seen as surprising given the proximity of the homeland of *Bos indicus* in the Indus valley, that they had been domesticated there since 8,000 BP and that there is archaeological evidence for contact (Matthews 2002). This could be explained by the geography between the Indus Valley and Iran, which is defined by an arc of mountains from Southern Iran to Afghanistan, a challenge to overcome. It is also possible that the introgression was present earlier but to a more reduced degree, which would explain the results in Monjuklu (but not in Syria, the distance is too big to see zebu introgression there and nowhere else), where admixture is not detected in Figure 38 but suggested with D-statistics (Figure 36 and 39). Second, *Bos indicus* cattle are found in the tropical regions of the world today (Loftus et al., 1994b; Hansen 2004; Patel and Meadow, 2017) and are known to have higher adaptation to aridity than *Bos taurus* animals, giving a plausible reason for the sharp and substantial influx. In Figure 38, many later animals have one third or more indicine ancestral components. Third, as suggested by Verdugo et al. (2019) the admixture may have been male-mediated. In total, across new and published samples, one out of the 118 samples shows an indicine haplogroup (the previously published Kho1, from Islamic Iran). This indicine haplogroup hence happens later, and other samples have predominantly domesticated haplogroups (T1-T2-T3-T4, for 83 out of the 118 total samples).

This introgression seems to be common in the Middle East from the Iron Age, but we are also able to detect it via D-statistics computations in Eastern Europe, from the Bronze and Iron Ages (Chapter 4, Figure 26). If admixture resulted from a deliberate cross breeding of new stock into existing herds this would likely have been via choice of zebu bulls by ancient Near Eastern breeders - a process which would have enabled fast and far reaching admixture. Deliberate modern introgression of zebu animals has mirrored this pattern, for example in tropical West Africa (Troy et al., 2001; MacHugh et al., 1997; Hanotte et al., 2002). Lastly, other evidence supports influx such as depictions and osteological evidence for *Bos indicus* (Matthews 2002; Grigson 2015), genomic evidence

for human movement from the East (Haber et al., 2017; Lazaridis et al., 2016) and the appearance of other Asian animal remains such as camels, buffalo and elephants (Çakırlar and Ikram 2016; Perry-Gal et al., 2015).

5.4.2. Aurochs introgression in early domesticates

Using D-statistics, Park et al. (2015), Verdugo et al. (2019) and Rossi et al. (2024) have indicated that, after initial domestication in the upper Euphrates region in eastern Anatolia, secondary admixtures with local aurochs populations may have taken place in Europe, the Levant, Africa and Central Asia. This can be seen in Europe via the presence of typical European aurochs haplogroup P in one domesticated sample from Chalcolithic Ukraine (Sha1), other hybrids having domesticated haplogroup (T) due to a male-driven introgression. Rather than using tests that identify specific introgressing genomes, this chapter used qpWave to test for such secondary inputs simultaneously using a diverse array of outgroups. This should have greater power to identify admixtures, even where the aurochs population having input has not yet been sequenced or is a mixture of sampled strains.

On Figure 40 was identified a group of similar samples from Neolithic Western Anatolia to Neolithic Central Iran, and encompassing Bronze Age samples from Iran and the Caucasus. This may reflect the extent of initial spread of unadmixed *Bos taurus*. An alternative is that the aurochs local to these regions was too similar to early herds to register in our qpWave test. However this is unlikely because (i) we do include a Holocene Armenian aurochs, Gyu2 (5,371-5,216 cal. BCE), as an outgroup which should have some affinities with the wild animals local to the Caucasus and Iranian regions; (ii) we pick up some evidence of non-core ancestry in Syrian samples which are situated very close to the core domestication region, indicating that local aurochsen can be sufficiently divergent to give a signal. It is interesting that Ch22, an aurochs bone sample from the Central Anatolian Neolithic context of Çatalhöyük which was identified as such by its large size, shows both some indication of divergence from the core cluster and some affinity with a subset of Anatolian cattle. This intermediate signal may indicate that this animal was itself an aurochs-domesticate hybrid, a theory supported by its haplogroup T3 (Chapter 3 Figure 19).

There is another cladal cluster of early cattle which includes Chalcolithic and Bronze Age genomes from central Asia (Turkmenistan), Israel, Lebanon and Syria. This cluster shares some affinity with the more geographically adjacent Iranian samples (Nad1 and Nad2). The Lebanon/Syrian portion of the cluster also shares some weaker affinity with a wider Iranian/Caucasus subset, also reflecting geographical proximity. Verdugo (2019) previously identified that local admixture from an aurochs with some affinity to a Mesolithic African aurochs genome was detectable in the Israeli Bronze Age. These results show that secondary aurochs admixture extended much closer in time and place to the core domestication region - up to Bronze Age Syria and Lebanon (~3,000 to 2,000 BC). Their cladality with Turkmenistan samples is interesting. This may reflect shared heritage from the same admixture or that admixtures involved similar aurochs populations. Given the geographically intermediate samples not clustering with all of them, the later option seems more liquelly.

5.4.3. T1 mtDNA haplogroup distribution and early introgression in the Middle East

In the present day, the T1 haplogroup is almost fixed in Africa (Bonfiglio et al., 2012, Olivieri et al., 2015), and was found in ancient cattle from the Levant (Verdugo et al., 2019). 19 ancient samples with haplogroup T1 have been identified on the PCA (Figure 35) and come from the Levant, Syria, Saudi Arabia, Northern and Eastern Africa. Ancient Middle Eastern samples plot on the left side and a cline of mainly T1 samples forms from them through ancient Africa and towards modern day Africa. Later Levantine samples are plotting on the right side of the figure, due to indicine introgression.

Using qpWave (Figures 40 and 41), Syrians and Levantine samples form a cluster independent of Anatolia, the Caucasus and Iran. Verdugo had hypothesized that the Levantine samples resulted from admixture with a local aurochs population similar to a low coverage African aurochs. This would explain why on qpWave they form a distinct cluster, and plot differently on the PCA, in the direction of modern day Africa. Despite the predominant representation of T1 in Africa, we see its early appearance in the Fertile Crescent, as high as Syria (in the sample VEM329), and as early as 7,000 years ago in Saudi

Arabia (with Aak1 being estimated by BEAST in Chapter 3 to be between 6,756 and 7,280 years old).

The site of Al'Ula in Saudi Arabia, where Aak1 was discovered, is a radiocarbon dated Neolithic site (Kennedy et al., 2023). It represents the earliest example of a T1 cattle, and is hence a candidate for the population that harbors the founder of this haplogroup. Originating outside of Africa, T1 would have spread in the Fertile Crescent, reaching Syria and the Levant, before movement toward Africa. The clustering of ancient samples from the Fertile Crescent (the Levant, Lebanon and Saudi Arabia on the PCA) with African samples from Iron Age Tunisia and Antiquity Ethiopia would suggest that African cattle descend from a small population of taurine cattle similar to ancient samples found in the Fertile Crescent. If this small population was characterised by the haplogroup T1, and no subsequent major migrations happened, then this would explain the predominance of T1 in today's Africa. A similar scenario was suggested regarding the modern breed Nguni in South Africa, which was found to show little diversity in the tested samples (Horsburgh et al., 2013), indicating a cattle migration from North to South, involving genetic bottlenecks.

The precise origin of T1 is unclear though, as it can be found in several places Bonfiglio et al., 2012). However, divergence times estimated by BEAST (Chapter 3, Figure 19 and Table 11) overlap for taurine domesticated haplogroups, indicating a similar timing for their origins. They could have appeared in the same Near Eastern population, and migrations associated with bottlenecks would have favored the spread of T1 in Africa, in a similar way to what is seen with haplogroup T3 in Europe.

Two aspects of the Levantine local aurochs introgression are interesting and mirror the introgression processes as cattle moved west (European aurochs input) and east (east Asian Aurochs input). First, it seems to be early as the Neolithic Saudi Arabian genome sample has a similar PCA position to other older individuals from the Levant and is Neolithic. Second, it seems to not have had an influence on mtDNA. The T1 haplogroup has been almost fixed in the movement of cattle south from the centre of origin but it is unlikely to have come from Levantine aurochs. This is because it is closely related to other *Bos taurus* predominant mtDNA haplogroups (Figure 19) (Achilli et al., 2008; Bonfiglio et

al., 2012; Rossi et al., 2024) with a coalescence around the time of domestication. Therefore it likely came from the same maternal lineage pool.

This points to a male-mediated introgression of wild genomic ancestry, as is the case elsewhere. Perhaps insemination by aurochs bulls may have been a deliberate management strategy used by early herders (Peters et al., 2012) either for stock improvement or because of the difficulty of keeping sexually mature bulls in captivity. Alternatively, it may be simply that the input of wild aurochs was something that early herders could not prevent that therefore occurred with some frequency when wild and domestic herds occupied the same regions. It has been noted however, that rare divergent mtDNA lineages have been detected in *Bos taurus*, indicating some maternal lineage introgression (Bonfiglio et al., 2010; Ginja et al., 2023; Rossi et al., 2025).

The trending of these early introgressed Levant herds towards modern African cattle and their affinity with ancient Tunisian cattle (Ginja et al., 2023) is interesting. The possibility of local Aurochs input into African cattle has been discussed for decades (Wendorf & Schild 1994; Bradley et al., 1996; Grigson 1991; Gifford-Gonzalez & Hanotte 2011; Marshall & Hildebrand 2002). Modern African cattle show distinctive phenotypes including resistance to tropical infections and genomic divergence (Decker et al., 2014; MacHugh et al., 1997). The PCA pattern suggests that the process that made animals from that continent distinctive started in the southern Middle East. Ginja et al. (2023) showed that Iron Age (2,800-2,000 BP) genomes from Althiburos, Tunisia (the STK-prefixed samples in Figure 35) had this genome type as well as T1 mtDNA, supporting a foundation from the Levant. However, it seems likely that the process of introgression continued in Africa. These animals also segregated a diverse mtDNA R lineage which is shared by the Moroccan aurochs and which has not been found in the Levant to date.

6. Conclusions

6.1. Ancient DNA and paleogenomics

Since the advent of NGS technology, researchers are not restricted to mitochondrial studies anymore. We can now leverage sequencing technologies and NGS data to detect introgression otherwise absent from mitochondrial phylogenies, and difficult to observe via Y chromosome studies. Indeed, in any species, Y chromosomes are present in only 50% of the population; it is therefore easy to lose paternal lineages through genetic drift. In the case of cattle herds, breeding practices tend to reduce paternal diversity, increasing the loss of data.

In this project, both maternal and nuclear genomes were used to study ancient cattle, and introgression through autosomes and/or the X chromosome was estimated, to mitigate the difficulty encountered regarding the limited availability of Y chromosome genomes. Admixture events involving wild and indicine populations are observed, and characterisation of different taurine populations can be made thanks to the additional samples sequenced during the project.

6.2. Screening and deeper sequencing

6.2.1. New samples

As in any project, results are limited by dataset availability and impacted by sampling bias. Data uncovered by Verdugo et al. (2019) was improved and new samples were screened in Chapter 2 to help refine results and hypotheses. 148 new samples were processed and evaluated for aDNA preservation, leading to the addition of 109 new samples sequenced to a minimum coverage of 0.01X. They represent different countries and time periods from Europe and the Middle East, in an attempt to minimise sampling bias.

6.2.2. DNA preservation

Teeth and petrous bones were favoured for DNA extraction, as they were shown to preserve aDNA better than other bone types (Gamba et al., 2014; Pinhasi et al., 2015; Parker et al., 2020), a result confirmed by the endogenous content observed in the dataset. Sample age and origin must be taken into account as well when extracting DNA. As expected, extended time periods accelerate DNA degradation (Lindahl 1993; Allentoft et al., 2012), as evidenced by the endogenous content differences between Neolithic/Bronze Age samples and more recent ones from the Iron Age and later periods. Environmental conditions can be equally impactful on DNA preservation (Bollongino et al., 2008; Pinhasi et al., 2015). None of the samples screened from Pakistan yielded recoverable DNA, and Saudi Arabia proved to be a very challenging environment, with only one sample sequenced at a very low coverage (but with good mitochondrial coverage thanks to mitochondrial capture).

6.2.3. Future work

Additional sequencing could be performed to improve these new samples, as it was done for 11 samples published in Verdugo et al. (2019), following the use of a modified extraction protocol involving a bleach wash (Boessenkool et al., 2017; Dabney and Meyer, 2019; Mattiangeli et al., 2023). However, older samples, or those found in arid environments, might require a protocol that excludes the bleach step, which could otherwise destroy any surviving DNA. Another possibility for the future is the imputation of genomes already available at at least 0.25X (Erven et al., 2024).

6.3. The mitochondrial haplogroup distribution in ancient cattle populations

6.3.1. Haplogroup distribution

Chapter 3 presents the mitogenomes of 106 individuals with a minimum coverage of 5.7X (72X on average), allowing mitochondrial phylogeny analyses via maximum

likelihood and BEAST methods. This led to the identification of several taurine and wild haplogroups in Eurasia and North Africa.

As reported in previous studies (Troy et al., 2001), the 3 taurine haplogroups can be found in the Fertile Crescent in ancient samples. T3 was mainly found in European individuals, T2 in the Middle East and T1, though less frequent, was only found in samples from the Middle East and North Africa. This distribution is consistent with modern observations (Bradley et al., 1996; Troy et al., 2001; Teasdale and Bradley, 2012).

6.3.2. Hybridization with aurochsen

Other haplogroups were identified in early cattle from the dataset presented in this thesis. In Eastern Europe, a sample carries the wild haplogroup P, while E is found in the Caucasus (Early Bronze Age) and Turkmenistan (Chalcolithic). Despite the predominance of T1 in Africa, some Iron Age cattle from Tunisia have haplogroup R, previously found in an aurochs from Epipaleolithic Morocco (Verdugo et al., 2019; Ginja et al., 2023). This shows first that, with time, haplogroup distribution became closer to what is seen today, with a reduction of haplogroup diversity and an increasing predominance of taurine haplogroups (T1-T2-T3), and secondly, that hybridisation occurred between wild and domesticated populations. However, D-statistics showed that this crossbreeding could have been mediated by wild males, leaving the maternal lineage unchanged, as suggested in several studies (Park et al., 2015; Upadhyay et al., 2017; Rossi et al., 2024).

6.3.3. New data on haplogroup T1

The last part of Chapter 3 concerned a very interesting sample dating from the Neolithic in Saudi Arabia. Attempting radiocarbon dating was unsuccessful, probably due to the age of the sample and the arid environmental conditions of Saudi Arabia. A BEAST phylogeny placed this individual in the T1 haplogroup, and estimated its age at approximately 7 kya. This supports an origin of T1 in the Middle East, rather than in Africa, and helps estimate overlapping timelines for the emergence of the 3 taurine haplogroups, indicating a similar geographic origin, and confirming previous findings on the appearance of T1 by Achilli et al. (2008).

6.4. Admixture events highlighted by whole-genome data

Using whole-genome sequencing performed on new and improved samples, Chapter 4 investigated population structure in Europe through allele frequency analyses.

6.4.1. Population structure

PCAs separated taurine and indicine individuals, and showed the presence of hybrids in the dataset, as evidenced by Verdugo et al. (2019). Admixture analyses estimated the involvement of several ancestral populations in Eurasia, including several taurine, wild and indicine ancestries. F4-ratios and D-statistics also revealed admixture events in Europe. Influence from the Middle East was detected, supporting a Middle Eastern origin of domestication, as described by Scheu et al. (2015), with decreasing influence westwards, likely resulting from genetic drift and the formation of a true European taurine population over time.

6.4.2. Local introgression

As domesticated cattle populated Europe, introgression from European aurochsen can be observed via D-statistics, and modelled through qpAdm, confirming hypotheses based on mitochondrial results in this study and in other papers (Achilli et al., 2008; Schibler et al., 2014; Park et al., 2015).

6.4.3. Eastern migrations: the Yamnaya

European aurochsen are not the only wild population detected in European domesticated cattle. Yamnaya migrations and the subsequent Corded Ware culture during the late Neolithic and Bronze Age in Europe, and their impact on human population genetics (Haak et al., 2015; Lazaridis et al., 2022) led to suppose the arrival of an Eastern component in cattle as well, as this nomadic population was pastoralist and travelled with domesticated cattle (Wilkin et al., 2021; Scot et al., 2022). This signal was first observed in a PCA, and was confirmed by several analyses (D-statistics, f4-ratios and qpAdm) that showed the emergence of an Eastern component beginning in the Late Neolithic in Europe, before a decline in later time periods, after gene flow from the East ceased.

6.5. Early populations in the Middle East

Chapter 5 was a continuation of the previous one, but focused more specifically on the Middle East, rather than combining these results together with the European populations. Indeed, analyses performed on Middle Eastern samples led to many results that needed to be presented separately for clarity. This region remains poorly described in aDNA studies, due to environmental conditions affecting DNA preservation. Fortunately, a batch of very good samples from Iran were screened and sequenced during this project, shedding new light on the structure of cattle populations.

6.5.1. The timing of indicine introgression

Verdugo et al. (2019) had already detected indicine introgression and estimated its onset at around 4,000 BP. PCA and D-statistics showed this component in the Middle East, beginning around the same period. 11 new samples showing this introgression confirmed previous observations, which coincide with a climatic change that probably motivated the search for better-adapted breeds. Newly processed samples found in sites dated to this period show an indicine component and might even push the timing back a few centuries, if radiocarbon dating is performed in the future. This admixture event was also detected in the Caucasus in a Bronze Age sample, but its age is also uncertain.

6.5.2. Early populations of the Middle East

To characterise populations in this region, early samples predating indicine introgression, as confirmed by D-statistics, were used to model domesticated cattle.This was made possible by the sequencing of several well-preserved Iranian samples during this project, and the use of previously published samples (Verdugo et al., 2019). They were analysed with qpWave and revealed 2 main clusters, one of which can be explained by the geographic proximity of specific samples, while the other involves distant individuals but not intermediate ones. This might suggest the presence of unknown wild populations that admixed with domesticated animals during the eastern progression of domestication, as seen in Europe.

The fact that several of these early samples had haplogroup T1 is another indicator of this haplogroup's origin in the Near East rather than in Africa, even though it is fixed in modern African cattle. It also confirms the hypotheses presented in Chapter 3, where the detection of wild ancestries alongside mostly taurine mitochondrial haplogroups suggested a predominantly male-mediated introgression. The same scenario may have occurred in Africa where, as domesticated cattle spread, local introgression took place, as evidenced by the presence of wild haplogroup R in Iron Age Tunisian cattle (Ginja et al., 2023).

6.6. Conclusion and future directions

This thesis presents screening results for 148 new samples and 32 previously processed ones. It led to the addition of 109 new genomes with a minimum of 0.01X, representing several regions of Eurasia. A sample from Saudi Arabia was particularly interesting to process due to its age and origin; it represents the first ancient genome from Saudi Arabia. Through mitochondrial phylogenies and allele frequency-based analyses, characterisation of ancient cattle populations was made possible, with a deeper focus on the emergence of T1 in the Middle East, and introgression from wild and indicine populations. Specifically, the detection of an Eastern aurochs component in Late Neolithic and Early Bronze Age Central and Western Europe can be associated with the appearance of the Corded Ware culture that followed the Yamnaya migrations. It highlights parallels between human and animal histories shaped by migration, since both were influenced by the Yamnaya in Europe.

However, these conclusions will require future radiocarbon dating to confirm chronology and better define the indicine influx in the Middle East and the Caucasus. Attempting to radiocarbon-date Aak1, the Neolithic sample from Saudi Arabia, was unsuccessful, possibly due to its age, as suggested by the BEAST phylogeny, and the poor preservation state due to an arid environment. Addressing this issue would help improve our understanding of the T1 haplogroup and its uncertain origin.

Questions regarding T1 and Middle Eastern cattle populations in general require additional sampling because of the poor success rate of DNA extraction in this region. Environmental conditions are not favourable to paleogenomics, but this thesis

demonstrated that it is possible with additional sequencing and careful attention to extraction protocols. For samples already available, imputation might be the next step, as demonstrated by Erven et al. (2024).

7. Appendix

Supplementary Table S1. Additional ancient samples used in the maximum likelihood tree.

Sample	Species	Sex	Country	Site	Period	Coverage (X)	Haplogroup	Publication
Azer1	Bos taurus	Female	Iran	Kültepe	Late Chalcolithic	35.380	Q	Verdugo et al., 2019
Baikal1	Bos primigenius	Male	Russia	Baikal	Paleolithic	1697.960	C	Rossi et al., 2024
Bed3	Bos primigenius	Female	Germany	Bedburg-Königshoven	Mesolithic	5004.960	P	Erven et al., 2024
Bed4	Bos primigenius	Female	Germany	Bedburg-Königshoven	Mesolithic	702.440	P	Rossi et al., 2024
Bel1	Bos taurus	Female	Serbia	Belovode-Veliko	Neolithic	13.820	Q	Verdugo et al., 2019
Bel2	Bos taurus	Male	Serbia	Belovode-Veliko	Neolithic	28.680	Q	Verdugo et al., 2019
Bes1	Bos admixed	Female	Iraq	Bestansur	Iron age	160.410	T2	Verdugo et al., 2019
Bes2	Bos admixed	Male	Iraq	Bestansur	Medieval	626.780	T3	Verdugo et al., 2019
Borly1	Bos primigenius	Male	Kazakhstan	Borly	Neolithic	221.104	K	Rossi et al., 2024 (Borlya)
Borly2	Bos primigenius	Male	Kazakhstan	Borly	Neolithic	931.296	C	Rossi et al., 2024 (Borly4b)
Borly3	Bos primigenius	Female	Kazakhstan	Borly	Neolithic	80.357	P	Rossi et al., 2024 (Borly4a)
Bul1	Bos taurus	Male	Uzbekistan	Tilla Bulak	Late Bronze	14.800	T2	Verdugo et al., 2019
Cas1	Bos taurus	Female	Levant	Horvat Castra	Byzantine	10.590	T1	Verdugo et al., 2019
Fif1	Bos primigenius	Male	Scotland	Newburgh, Fife	Mesolithic	52.303	P	Rossi et al., 2024
Fre1	Bos primigenius	Male	Denmark	Frederiksborg, Alsønderup	Paleolithic	24.746	P	Unpublished
Gyu1	Bos primigenius	Male	Armenia	Gyumri	Paleolithic	1393.300	K	Rossi et al., 2024
Gyu2	Bos primigenius	Male	Armenia	Gyumri	Neolithic	479.450	Q	Verdugo et al., 2019

Has3	Bos taurus	Male	Iran	Hasanlu	Bronze Age	109.960	T2	Verdugo et al., 2019
Hjo1	Bos primigenius	Male	Denmark	Hjørring, Tofte Bæk	Neolithic	354.540	P	Rossi et al., 2024
Hxh2	Bos primigenius	Female	Germany	Herxheim	Neolithic	2529.150	P	Rossi et al., 2024
Kaz1	Bos admixed	Female	Georgia	Dariali Tamara Fort	Medieval	125.590	T3	Verdugo et al., 2019
Kaz2	Bos admixed	Female	Georgia	Dariali Tamara Fort	Medieval	29.990	T3	Verdugo et al., 2019
Kaz3	Bos admixed	Female	Georgia	Dariali Tamara Fort	Medieval	39.270	T2	Verdugo et al., 2019
Kaz4	Bos admixed	Female	Georgia	Dariali Tamara Fort	Medieval	30.040	T3	Verdugo et al., 2019
Kok1	Bos admixed	Male	Uzbekistan	Kok Tepe	Iron Age	94.910	P	Verdugo et al., 2019
KOL9	Bos taurus	Female	Russia	Kolyvanskoe, Altai Krai	Bronze Age	174.170	Q	Rossi et al., 2024
Men2	Bos taurus	Female	Anatolia	Mentese	Neolithic	88.000	Q	Verdugo et al., 2019
NVL1	Bos primigenius	Female	Russia	Novoilinka-6, Altai Krai	Bronze Age	10.940	C	Rossi et al., 2024
NVL3	Bos primigenius	Female	Russia	Novoilinka-6, Altai Krai	Bronze Age	51.360	C	Rossi et al., 2024
Plo1	Bos taurus	Female	Serbia	Plocnik	Middle Neolithic	10.000	Q	Verdugo et al., 2019
Plo2	Bos taurus	Female	Serbia	Plocnik	Late Neolithic	36.000	Q	Verdugo et al., 2019
Plo3	Bos taurus	Male	Serbia	Plocnik	Middle Neolithic	153.000	T3	Verdugo et al., 2019
Plo4	Bos taurus	Male	Serbia	Plocnik	Late Neolithic	154.000	T2	Verdugo et al., 2019
Plo5	Bos taurus	Male	Serbia	Plocnik	Late Neolithic	13.000	T3	Verdugo et al., 2019
Plo7	Bos taurus	Female	Serbia	Plocnik	Late Neolithic	19.000	T3	Verdugo et al., 2019
Plo8	Bos taurus	Female	Serbia	Plocnik	Late Neolithic	14.000	T2	Verdugo et al., 2019
Qaz1	Bos taurus	Female	Iran	Qazvin	Bronze Age	193.000	T2	Verdugo et al., 2019

Rhi3	Bos primigenius	Male	Germany	Upper Rhine Valley	Pleistocene	1075.220	P	Rossi et al., 2024
ROS002	Bos primigenius	Male	Kazakhstan	Roschinskoe	Chalcolithic	7.340	P	Rossi et al., 2024
Ska1	Bos	Male	Sweden	Skåne	Mesolithic	588.170	P	Rossi et al., 2024
Ska3	Bos	Male	Sweden	Skåne	Mesolithic	1105.770	P	Rossi et al., 2024
Sub1	Bos taurus	Male	Anatolia	Suberde	Neolithic	13.850	T2	Verdugo et al., 2019
Th7	Bos primigenius	Female	Morocco	Taghit Haddouch	Paleolithic	8.400	R	Verdugo et al., 2019
Tri1	Bos primigenius	Male	Germany	Trier-Brüderkrankenhaus	Mesolithic	822.200	P	Rossi et al., 2024
Tsa1	Bos admixed	Female	Levant	Tel es-Safi	Iron Age	160.000	Q	Verdugo et al., 2019
Tsa2	Bos admixed	Female	Levant	Tel es-Safi	Iron Age	1.000	T1	Verdugo et al., 2019
Tula1	Bos primigenius	Male	Russia	Tula district	Paleolithic	2443.070	C	Rossi et al., 2024
Tyq1	Bos admixed	Male	Levant	Tel Yoqneam	Early Hellenistic	11.000	T2	Verdugo et al., 2019
Var1	Bos primigenius	Male	Kazakhstan	Varfolomeevka	Neolithic	367.710	Q	Rossi et al., 2024
Var2	Bos primigenius	Male	Kazakhstan	Varfolomeevka	Neolithic	2469.010	K	Rossi et al., 2024
ALT232	Bos taurus	Female	Tunisia	Althiburos	Iron Age	8.19	R	Ginja et al., 2023
STK234	Bos taurus	Female	Tunisia	Althiburos	Iron Age	7.35	T1	Ginja et al., 2023
STK239	Bos taurus	Male	Tunisia	Althiburos	Iron Age	29.21	T1	Ginja et al., 2023
STK243	Bos taurus	Male	Tunisia	Althiburos	Iron Age	7.93	T1	Ginja et al., 2023

Supplementary Table S2. Modern samples used in the maximum likelihood tree (Figure 14).

NCBI ID	Species	Country	Haplogroup	Reference
EU177868	*Bos indicus*	Iraq	I	Achilli et al., 2008
EU177870	*Bos indicus*	Iran	I	Achilli et al., 2008
DQ124389	*Bos taurus*	Korea	P	Direct NCBI Submission
EU177867	*Bos taurus*	Italy	Q	Achilli et al., 2008
HQ184031	*Bos taurus*	Italy	Q	Achilli et al., 2009
HQ184034	*Bos taurus*	Italy	Q	Bonfiglio et al., 2010
HQ184039	*Bos taurus*	Italy	Q	Bonfiglio et al., 2010
KT184471	*Bos taurus*	Egypt	Q1	Olivieri et al., 2015
FJ971081	*Bos taurus*	Italy	Q2	Achilli et al., 2009
FJ971085	*Bos taurus*	Italy	R1	Achilli et al., 2009
HQ184040	*Bos taurus*	Italy	R1	Bonfiglio et al., 2010
FJ971087	*Bos taurus*	Italy	R2	Achilli et al., 2009
EU177841	*Bos taurus*	Italy	T1	Achilli et al., 2008
JN817304	*Bos taurus*	Ethiopia	T1	Bonfiglio et al., 2012
JN817320	*Bos taurus*	Italy	T1	Bonfiglio et al., 2012
JN817328	*Bos taurus*	Egypt	T1	Bonfiglio et al., 2012
KF163063	*Bos taurus*	South Africa	T1	Horsburgh et al., 2013
KF163093	*Bos taurus*	South Africa	T1	Horsburgh et al., 2013
MZ901441	*Bos taurus*	Croatia	T2	Cubric-Curik et al., 2021
DQ124383	*Bos taurus*	Korea	T2	Direct NCBI Submission
DQ124396	*Bos taurus*	Korea	T2	Direct NCBI Submission
EU177854	*Bos taurus*	Italy	T2	Achilli et al., 2008
EU177858	*Bos taurus*	Iran	T2	Achilli et al., 2008
EU177859	*Bos taurus*	Iran	T2	Achilli et al., 2008
AY676859	*Bos taurus*	Scotland	T3	Direct NCBI Submission
DQ124406	*Bos taurus*	Netherlands	T3	Direct NCBI Submission
EU177831	*Bos taurus*	Italy	T3	Achilli et al., 2008
EU177839	*Bos taurus*	Iran	T3	Achilli et al., 2008
V00654	*Bos taurus*	UK	T3	Anderson et al., 1982
DQ124372	*Bos taurus*	Korea	T4	Direct NCBI Submission
DQ124377	*Bos taurus*	Korea	T4	Direct NCBI Submission

DQ124401	*Bos taurus*	Korea	T4	Direct NCBI Submission
EU177862	*Bos taurus*	Italy	T5	Achilli et al., 2008

Supplementary Table S3. Dataset of modern and ancient samples used in the BEAST analyses (Figure 19).

Sample	Species	Country	BEAST calibrations	Reference
JN817298.1_A8	*Bos taurus*	Ethiopia	0	Bonfiglio et al., 2012
JN817299.1_Adt23	*Bos taurus*	Ethiopia	0	Bonfiglio et al., 2012
JN817300.1_Alt9280	*Bos taurus*	Portugal	0	Bonfiglio et al., 2012
JN817301.1_Alt9301	*Bos taurus*	Portugal	0	Bonfiglio et al., 2012
JN817302.1_Ar18	*Bos taurus*	Ethiopia	0	Bonfiglio et al., 2012
JN817303.1_Ar22	*Bos taurus*	Ethiopia	0	Bonfiglio et al., 2012
JN817304.1_Ar29	*Bos taurus*	Ethiopia	0	Bonfiglio et al., 2012
JN817305.1_B1	*Bos taurus*	Ethiopia	0	Bonfiglio et al., 2012
JN817306.1_Cal24	*Bos taurus*	Italy	0	Bonfiglio et al., 2012
JN817307.1_Cbc16	*Bos taurus*	Mexico	0	Bonfiglio et al., 2012
JN817308.1_CCh01	*Bos taurus*	Mexico	0	Bonfiglio et al., 2012
JN817309.1_Ccq3	*Bos taurus*	Paraguay	0	Bonfiglio et al., 2012
JN817310.1_Ccq31	*Bos taurus*	Paraguay	0	Bonfiglio et al., 2012
JN817311.1_Ccq40	*Bos taurus*	Paraguay	0	Bonfiglio et al., 2012
JN817312.1_Ch34	*Bos taurus*	Italy	0	Bonfiglio et al., 2012
JN817313.1_Chi336	*Bos taurus*	Italy	0	Bonfiglio et al., 2012
JN817314.1_Chi397	*Bos taurus*	Italy	0	Bonfiglio et al., 2012
JN817315.1_Chi481	*Bos taurus*	Italy	0	Bonfiglio et al., 2012
JN817316.1_Chi575	*Bos taurus*	Italy	0	Bonfiglio et al., 2012
JN817317.1_Cin11	*Bos taurus*	Italy	0	Bonfiglio et al., 2012
JN817318.1_Cin15	*Bos taurus*	Italy	0	Bonfiglio et al., 2012
JN817319.1_Cinl5b	*Bos taurus*	Italy	0	Bonfiglio et al., 2012
JN817320.1_Cinl14	*Bos taurus*	Italy	0	Bonfiglio et al., 2012
JN817321.1_Eg11	*Bos taurus*	Egypt	0	Bonfiglio et al., 2012
JN817322.1_Eg12	*Bos taurus*	Egypt	0	Bonfiglio et al., 2012
JN817323.1_Eg18	*Bos taurus*	Egypt	0	Bonfiglio et al., 2012
JN817324.1_Eg19	*Bos taurus*	Egypt	0	Bonfiglio et al., 2012
JN817325.1_Eg22	*Bos taurus*	Egypt	0	Bonfiglio et al., 2012

JN817326.1_Eg23	*Bos taurus*	Egypt	0	Bonfiglio et al., 2012
JN817328.1_Eg28	*Bos taurus*	Egypt	0	Bonfiglio et al., 2012
JN817329.1_Eg36	*Bos taurus*	Egypt	0	Bonfiglio et al., 2012
JN817330.1_H23	*Bos taurus*	Ethiopia	0	Bonfiglio et al., 2012
JN817331.1_Lmi50	*Bos taurus*	France	0	Bonfiglio et al., 2012
JN817332.1_Mar10	*Bos taurus*	Italy	0	Bonfiglio et al., 2012
JN817333.1_Mar12	*Bos taurus*	Italy	0	Bonfiglio et al., 2012
JN817334.1_Mcg358	*Bos taurus*	Italy	0	Bonfiglio et al., 2012
JN817335.1_Mcg363	*Bos taurus*	Italy	0	Bonfiglio et al., 2012
JN817336.1_Mcg378	*Bos taurus*	Italy	0	Bonfiglio et al., 2012
JN817337.1_Mcg383	*Bos taurus*	Italy	0	Bonfiglio et al., 2012
JN817338.1_Mcg452	*Bos taurus*	Italy	0	Bonfiglio et al., 2012
JN817339.1_Mcg469	*Bos taurus*	Italy	0	Bonfiglio et al., 2012
JN817340.1_Per5	*Bos taurus*	Italy	0	Bonfiglio et al., 2012
JN817341.1_Per10	*Bos taurus*	Italy	0	Bonfiglio et al., 2012
JN817342.1_Per17	*Bos taurus*	Italy	0	Bonfiglio et al., 2012
JN817343.1_Pod16N	*Bos taurus*	Italy	0	Bonfiglio et al., 2012
JN817344.1_Pod41	*Bos taurus*	Italy	0	Bonfiglio et al., 2012
JN817345.1_Rom242	*Bos taurus*	Italy	0	Bonfiglio et al., 2012
JN817346.1_Rom480	*Bos taurus*	Italy	0	Bonfiglio et al., 2012
JN817347.1_Rom558	*Bos taurus*	Italy	0	Bonfiglio et al., 2012
JN817348.1_S8	*Bos taurus*	Ethiopia	0	Bonfiglio et al., 2012
JN817349.1_S13	*Bos taurus*	Ethiopia	0	Bonfiglio et al., 2012
JN817350.1_Chi425	*Bos taurus*	Italy	0	Bonfiglio et al., 2012
JN817351.1_Mcg489	*Bos taurus*	Italy	0	Bonfiglio et al., 2012
Aak1	*Bos taurus*	Saudi Arabia	6755.0,7307.0	
Ace1	*Bos taurus*	Anatolia	4300.0,4900.0	Verdugo et al., 2019
Ash4	*Bos taurus*	Levant	3023.0,3223.0	Verdugo et al., 2019
Azer1	*Bos taurus*	Iran	6073.0,6283.0	Verdugo et al., 2019
Bel1	*Bos taurus*	Serbia	6823.0,7073.0	Verdugo et al., 2019
Bel2	*Bos taurus*	Serbia	6823.0,7073.0	Verdugo et al., 2019
Bes1	*Bos admixed*	Iraq	2632.0,2934.0	Verdugo et al., 2019
Bes2	*Bos admixed*	Iraq	624.0,731.0	Verdugo et al., 2019
Bub1	*Bos taurus*	Serbia	7623.0,8023.0	Verdugo et al., 2019
Bul1	*Bos taurus*	Uzbekistan	3723.0,4023.0	Verdugo et al., 2019
Bul2	*Bos taurus*	Uzbekistan	3500.0,4100.0	

Cas1	*Bos taurus*	Levant	1385.0,1699.0	Verdugo et al., 2019
Ch22	*Bos primigenius*	Anatolia	7623.0,7773.0	Verdugo et al., 2019
Gyu2	*Bos primigenius*	Armenia	7239.0,7394.0	Verdugo et al., 2019
Has3	*Bos taurus*	Iran	4200.0,4700.0	Verdugo et al., 2019
Has4	*Bos taurus*	Iran	2823.0,3073.0	Verdugo et al., 2019
Has5	*Bos taurus*	Iran	2298.0,2573.0	Verdugo et al., 2019
Kaz1	*Bos admixed*	Georgia	0.0,1123.0	Verdugo et al., 2019
Kaz2	*Bos admixed*	Georgia	1023.0,1223.0	Verdugo et al., 2019
Kaz3	*Bos admixed*	Georgia	1323.0,1423.0	Verdugo et al., 2019
Kaz4	*Bos admixed*	Georgia	1123.0,1423.0	Verdugo et al., 2019
KOL10	*Bos taurus*	Russia	3800.0,4200.0	Rossi et al., 2024
KOL9	*Bos taurus*	Russia	3800.0,4200.0	Rossi et al., 2024
Men1	*Bos taurus*	Anatolia	7723.0,8423.0	Verdugo et al., 2019
Men2	*Bos taurus*	Anatolia	7843.0,7998.0	Verdugo et al., 2019
Mon1	*Bos taurus*	Turkmenistan	6523.0,7123.0	Verdugo et al., 2019
Mon7	*Bos taurus*	Turkmenistan	6000.0,7000.0	Rossi et al., 2024
Plo1	*Bos taurus*	Serbia	7000.0,7275.0	Verdugo et al., 2019
Plo2	*Bos taurus*	Serbia	6650.0,7025.0	Verdugo et al., 2019
Plo3	*Bos taurus*	Serbia	6950.0,7100.0	Verdugo et al., 2019
Plo4	*Bos taurus*	Serbia	6950.0,7100.0	Verdugo et al., 2019
Plo5	*Bos taurus*	Serbia	6523.0,6873.0	Verdugo et al., 2019
Plo7	*Bos taurus*	Serbia	6950.0,7100.0	Verdugo et al., 2019
Plo8	*Bos taurus*	Serbia	6950.0,7100.0	Verdugo et al., 2019
Qaz1	*Bos taurus*	Iran	4659.0,4905.0	Verdugo et al., 2019
Shiderti1	*Bos taurus*	Kazakhstan	5800.0,5950.0	Rossi et al., 2024
Shimao02	*Bos taurus*	China	3835.0,3975.0	Chen et al., 2018
Shimao04	*Bos taurus*	China	3835.0,3975.0	Chen et al., 2018
Shimao05	*Bos taurus*	China	3835.0,3975.0	Chen et al., 2018
Shimao07	*Bos taurus*	China	3835.0,3975.0	Chen et al., 2018
Shimao08	*Bos taurus*	China	3835.0,3975.0	Chen et al., 2018
STK239	*Bos taurus*	Tunisia	2759.0,2877.0	Ginja et al., 2023
Stu1	*Bos taurus*	Serbia	6523.0,7123.0	Verdugo et al., 2019
Sub1	*Bos taurus*	Anatolia	8047.0,8244.0	Verdugo et al., 2019
Tango1	*Bos primigenius*	Russia	3914.0,4142.0	Rossi et al., 2024
Tango2	*Bos primigenius*	Russia	4379.0,4599.0	Rossi et al., 2024
TEF1	*Bos taurus*	Lebanon	3400.0,5150.0	

Tqa2	*Bos taurus*	Levant	4223.0,5173.0	Verdugo et al., 2019
Tsa1	*Bos admixed*	Levant	2743.0,2948.0	Verdugo et al., 2019
Tsa2	*Bos admixed*	Levant	3023.0,3223.0	Verdugo et al., 2019
Tyq1	*Bos admixed*	Levant	2190.0,2355.0	Verdugo et al., 2019
Var1	*Bos primigenius*	Russia	7500.0,7641.0	Rossi et al., 2024
VEM090	*Bos taurus*	Egypt	4003.0,4010.0	
VEM104A	*Bos taurus*	Egypt	4003.0,4010.0	
Yer1	*Bos taurus*	Uzbekistan	2000.0,2100.0	Verdugo et al., 2019
YoA	*Bos taurus*	England	4658.0,4909.0	Rossi et al., 2024

Supplementary Table S4. List of the 373 ancient samples analysed in Chapter 4. For each, the analysis they are included in is informed in the corresponding column.

Lab name	Species	Country	Context	Whole genome Coverage (X)	LASER PCA	Dstat	Admix-ture	f4-ratio	qpAdm	Origin
Aak1	*Bos taurus*	Saudi Arabia	Neolithic	0.002	X	X		X		This project
Abu1	*Bos taurus*	Israel	MPPNB	0.000		X		X		Verdugo et al., 2019
Abu2	*Bos taurus*	Israel	Mid-Late PPNB	0.000		X		X		Verdugo et al., 2019
Ace1	*Bos taurus*	Turkey	Bronze Age	0.310	X	X		X		Verdugo et al., 2019
ALT232	*Bos taurus*	Tunisia	Iron Age	0.010	X	X		X		Ginja et al., 2023
ALT235	*Bos taurus*	Tunisia	Iron Age	0.010	X	X		X	X	Ginja et al., 2023
Ash4	*Bos admixed*	Israel	Early Iron Age	0.010	X	X		X		Verdugo et al., 2019
Asp1	*Bos taurus*	UK	Medieval	1.850	X	X		X	X	This project
Azer1	*Bos taurus*	Iran	Late Chalcolithic	0.001	X	X		X		Verdugo et al., 2019
Baikal1	*Bos primigenius*	Russia	Paleolithic	13.583	X		X	X		Rossi et al., 2024
Bal1	*Bos taurus*	Ireland	Medieval	2.640	X	X		X	X	TCD dataset (Processed by Victoria Mullin)
Bal2	*Bos taurus*	Ireland	Medieval	2.030	X	X	X	X	X	TCD dataset (Processed by Victoria Mullin)
Bed3	*Bos primigenius*	Germany	Mesolithic	18.490	X			X		Erven et al., 2024
Bed4	*Bos primigenius*	Germany	Mesolithic	2.090	X			X		Rossi et al., 2024
Bel1	*Bos taurus*	Serbia	Late Neolithic	0.090	X	X		X	X	Verdugo et al., 2019
Bel2	*Bos taurus*	Serbia	Late Neolithic	0.130	X	X		X	X	Verdugo et al., 2019
Belarus1	*Bos primigenius*	Belarus	Late Neolithic	1.785	X			X	X	TCD dataset (processed by Marta Verdugo)

Bes1	*Bos admixed*	Iraq	Iron age	2.080	X	X	X	X		Verdugo et al., 2019
Bes2	*Bos admixed*	Iraq	Medieval	13.800	X	X	X	X		Verdugo et al., 2019
Bis1	*Bos taurus*	England	Neolithic	3.020	X	X		X	X	TCD dataset (Processed by Victoria Mullin)
Bis2	*Bos taurus*	England	Neolithic	2.740	X	X		X	X	TCD dataset (Processed by Victoria Mullin)
Bisk1	*Bos taurus*	Poland	Neolithic	2.430	X	X		X		TCD dataset (Processed by Victoria Mullin)
Bla1	*Bos taurus*	Serbia	Neolithic	0.300	X	X		X		Verdugo et al., 2019
Bla2	*Bos taurus*	Serbia	Early Neolithic	0.060	X	X		X		Verdugo et al., 2019
Bla3	*Bos taurus*	Serbia	Neolithic	0.020	X	X		X		TCD dataset (processed by Marta Verdugo)
Bol1	*Bos taurus*	Iran	Neolithic	0.690	X	X		X		This project
Bor1	*Bos taurus*	Scotland	Late Iron Age	2.680	X	X		X	X	TCD dataset (Processed by Victoria Mullin)
Borly1	*Bos primigenius*	Kazakhstan	Neolithic	0.300	X					Rossi et al., 2024
Borly2	*Bos primigenius*	Kazakhstan	Neolithic	1.700	X					Rossi et al., 2024
Borly3	*Bos primigenius*	Kazakhstan	Neolithic	0.190	X					Rossi et al., 2024
Bri1	*Bos taurus*	Netherlands	Roman	7.380	X	X		X		TCD dataset (processed by Marta Verdugo and Victoria Mullin)
Brn1	*Bos taurus*	France	Roman Period	1.000	X	X		X	X	This project
Brn2	*Bos taurus*	France	Roman Period	1.180	X	X		X	X	This project
Brn3	*Bos taurus*	France	Roman Period	2.790	X	X		X	X	This project
Brn4	*Bos taurus*	France	Roman Period	1.120	X	X		X	X	This project
Brn5	*Bos taurus*	France	Roman Period	1.430	X	X		X	X	This project
Bub1	*Bos taurus*	Serbia	Neolithic	0.800	X	X		X		Verdugo et al., 2019
Bub2	*Bos taurus*	Serbia	Neolithic	0.110	X	X		X		This project
Bul1	*Bos taurus*	Uzbekistan	Late Bronze Age	14.800	X	X	X	X		Verdugo et al., 2019
Bul2	*Bos taurus*	Uzbekistan	Late Bronze Age	1.090	X	X	X	X		This project
Bun1	*Bos taurus*	Netherlands	Iron age	1.760	X	X		X	X	TCD dataset (processed by Marta Verdugo and Victoria Mullin)
Bun2	*Bos taurus*	Netherlands	Iron age	2.380	X	X		X	X	TCD dataset (processed by Marta Verdugo and Victoria Mullin)
BZL1	*Bos taurus*	Russia	Bronze Age	1.940	X	X				Rossi et al., 2024
BZL2	*Bos taurus*	Russia	Bronze Age	2.430	X	X				Rossi et al., 2024
Cas1	*Bos admixed*	Israel	Byzantine	10.590	X	X	X	X		Verdugo et al., 2019

Cer1	*Bos taurus*	Germany	Neolithic?	0.005	X	X		X		This project
Cer2	*Bos taurus*	Germany	Neolithic	0.090	X	X		X		This project
Cer3	*Bos taurus*	Germany	Late Neolithic?	0.520	X	X	X	X	X	This project
Cer4	*Bos taurus*	Germany	Iron Age	0.150	X	X		X	X	This project
Ch22	*Bos primigenius*	Turkey	Chalcolithic	0.310	X			X		Verdugo et al., 2019
Cha1	*Bos taurus*	England	Early Medieval	2.170	X	X		X	X	TCD dataset (Processed by Victoria Mullin)
Che1	*Bos taurus*	Bulgaria	Neolithic	0.480	X	X		X	X	TCD dataset (Processed by Conor Rossi)
Cla1	*Bos taurus*	Scotland	Bronze Age	15.250	X	X		X	X	TCD dataset (Processed by Victoria Mullin)
Cla2	*Bos taurus*	Scotland	Bronze Age	1.860	X	X		X	X	TCD dataset (Processed by Victoria Mullin)
Cla3	*Bos taurus*	Scotland	Bronze Age	0.150	X	X		X	X	TCD dataset (Processed by Victoria Mullin)
Cla4	*Bos taurus*	Scotland	Bronze Age	0.320	X	X		X	X	TCD dataset (Processed by Victoria Mullin)
Cla5	*Bos taurus*	Scotland	Bronze Age	3.330	X	X		X	X	TCD dataset (Processed by Victoria Mullin)
Cla6	*Bos taurus*	Scotland	Iron Age	0.360	X	X		X	X	TCD dataset (Processed by Victoria Mullin)
Cla7	*Bos taurus*	Scotland	Bronze Age	0.290	X	X		X	X	TCD dataset (Processed by Victoria Mullin)
Cla8	*Bos taurus*	Scotland	Bronze Age	3.100	X	X		X	X	TCD dataset (Processed by Victoria Mullin)
Cla9	*Bos taurus*	Scotland	Bronze Age	2.920	X	X		X	X	TCD dataset (Processed by Victoria Mullin)
CPC98	*Bos primigenius*	England	Mesolithic	3.770	X					Park et al., 2015
Cri1	*Bos taurus*	UK	Roman Period	0.690	X	X		X	X	This project
Cuk1	*Bos taurus*	Turkey	Neolithic	0.090	X	X		X		This project
Cuk2	*Bos taurus*	Turkey	Neolithic	0.440	X	X		X		This project
Da1	*Bos taurus*	England	Iron age	7.280	X	X		X	X	TCD dataset (Processed by Victoria Mullin)
Da2	*Bos taurus*	England	Iron age	0.260	X	X		X		TCD dataset (Processed by Victoria Mullin)
Da3	*Bos taurus*	England	Iron age	2.510	X	X		X	X	TCD dataset (Processed by Victoria Mullin)
Da4	*Bos taurus*	England	Early Medieval	4.830	X	X		X	X	TCD dataset (Processed by Victoria Mullin)
Da5	*Bos taurus*	England	Iron age	5.440	X	X		X	X	TCD dataset (Processed by Victoria Mullin)

Da6	*Bos taurus*	England	Iron age	3.450	X	X		X	X	TCD dataset (Processed by Victoria Mullin)
Dam1	*Bos taurus*	North Caucasus	Bronze Age	0.010	X	X		X		This project
Dan1	*Bos taurus*	Israel	Early Bronze Age	0.000		X		X		Verdugo et al., 2019
Dav1	*Bos taurus*	Iran	Chalcolithic	1.000	X	X		X		This project
Dav2	*Bos taurus*	Iran	Chalcolithic	1.730	X	X		X		This project
Den1	*Bos taurus*	Denmark	Neolithic	3.550	X	X		X	X	TCD dataset (Processed by Victoria Mullin)
Den2	*Bos taurus*	Denmark	Neolithic	3.930	X	X		X	X	TCD dataset (Processed by Victoria Mullin)
Dro1	*Bos taurus*	Netherlands	Iron Age	2.840	X	X		X	X	TCD dataset (processed by Marta Verdugo and Victoria Mullin)
Dub1	*Bos taurus*	Ireland	Medieval	2.180	X	X		X	X	TCD dataset (Processed by Victoria Mullin)
Dur1	*Bos taurus*	England	Late Neolithic	2.440	X	X		X	X	TCD dataset (Processed by Victoria Mullin)
Dur2	*Bos taurus*	England	Late Neolithic	2.140	X	X		X	X	TCD dataset (Processed by Victoria Mullin)
Dur3	*Bos taurus*	England	Late Neolithic	0.380	X	X		X	X	TCD dataset (Processed by Victoria Mullin)
Dyr1	*Bos taurus*	Germany	Medieval	2.310	X	X	X	X		TCD dataset (Processed by Victoria Mullin)
Dzh1	*Bos taurus*	Bulgaria	Early Neolithic	9.680	X	X		X	X	TCD dataset (Processed by Victoria Mullin)
Dzh2	*Bos taurus*	Bulgaria	Early Neolithic	4.100	X	X		X		TCD dataset (Processed by Victoria Mullin)
Els1	*Bos taurus*	Netherlands	Roman	2.520	X	X		X	X	TCD dataset (processed by Marta Verdugo and Victoria Mullin)
ENK1	*Bos taurus*	Netherlands	Bronze Age	1.910	X	X		X	X	TCD dataset (processed by Conor Rossi and Victoria Mullin)
ENK2	*Bos taurus*	Netherlands	Bronze Age	0.280	X	X		X	X	TCD dataset (processed by Conor Rossi and Victoria Mullin)
ENK3	*Bos taurus*	Netherlands	Bronze Age	2.670	X	X		X	X	TCD dataset (processed by Conor Rossi and Victoria Mullin)
ENK4	*Bos taurus*	Netherlands	Bronze Age	2.140	X	X		X	X	TCD dataset (processed by Conor Rossi and Victoria Mullin)
ENK5	*Bos taurus*	Netherlands	Bronze Age	0.270	X	X		X	X	TCD dataset (processed by Conor Rossi and Victoria Mullin)

Far1	*Bos taurus*	Iran	Neolithic	0.110	X	X		X		Verdugo et al., 2019
Fif1	*Bos primigenius*	Scotland	Mesolithic	0.150	X					Rossi et al., 2024
Fir1	*Bos taurus*	Netherlands	Medieval	3.150	X	X		X	X	TCD dataset (processed by Marta Verdugo and Victoria Mullin)
Fis1	*Bos taurus*	Scotland	Iron Age	2.930	X	X		X	X	TCD dataset (processed by Victoria Mullin)
Fis2	*Bos taurus*	Scotland	Iron Age	2.690	X	X		X	X	TCD dataset (processed by Victoria Mullin)
For1	*Bos taurus*	Iran	Iron Age	0.410	X	X	X	X		This project
Galicia1	*Bos primigenius*	Spain	Mesolithic	17.080	X			X		Rossi et al., 2024
Galicia2	*Bos primigenius*	Spain	Mesolithic	0.846	X			X		Rossi et al., 2024
Galicia3	*Bos primigenius*	Spain	Mesolithic	2.926	X			X		Rossi et al., 2024
Gil1	*Bos taurus*	Israel	Chalcolithic	0.000		X		X		Verdugo et al., 2019
GQ1	*Bos taurus*	England	Bronze Age	0.070	X	X		X	X	TCD dataset (processed by Victoria Mullin)
Gub1	*Bos taurus*	Italy	Bronze Age	1.050	X	X		X	X	This project
Gub2	*Bos taurus*	Italy	Bronze Age	0.550	X	X		X	X	This project
Gub3	*Bos taurus*	Italy	Bronze Age	0.420	X	X		X	X	This project
Gub4	*Bos taurus*	Italy	Bronze Age	0.440	X	X		X	X	This project
Gub5	*Bos taurus*	Italy	Bronze Age	0.490	X	X		X	X	This project
Gyu1	*Bos primigenius*	Armenia	Paleolithic	17.650	X		X	X		Rossi et al., 2024
Gyu2	*Bos primigenius*	Armenia	Neolithic	1.700	X		X	X		Verdugo et al., 2019
Had1	*Bos taurus*	England	Roman	1.460	X	X		X	X	TCD dataset (processed by Victoria Mullin)
Has1	*Bos taurus*	Iran	Middle Bronze Age	0.030	X	X	X	X		Verdugo et al., 2019
Has3	*Bos taurus*	Iran	Bronze Age	1.500	X	X		X		Verdugo et al., 2019
Has4	*Bos taurus*	Iran	Iron Age II	1.670	X	X	X	X		Verdugo et al., 2019
Has5	*Bos taurus*	Iran	Achaemenid / Seleuco-Parthian	0.040	X	X	X	X		Verdugo et al., 2019
Has6	*Bos taurus*	Iran	Iron Age	0.050	X	X		X		This project
Has7	*Bos taurus*	Iran	Bronze Age	0.500	X	X	X	X		This project
Has8	*Bos taurus*	Iran	Iron Age	0.550	X	X		X		This project
Has9	*Bos taurus*	Iran	Iron Age	0.400	X	X	X	X		This project

HF1	*Bos taurus*	N. Ireland	Bronze Age?	0.170	X	X				TCD dataset (processed by Victoria Mullin)
Hjo1	*Bos primigenius*	Denmark	Neolithic	0.870	X			X		Rossi et al., 2024
Hou1	*Bos taurus*	Netherlands	Medieval	2.560	X	X		X	X	TCD dataset (processed by Marta Verdugo and Victoria Mullin)
Hou2	*Bos taurus*	Netherlands	Medieval	2.630	X	X		X	X	TCD dataset (processed by Marta Verdugo and Victoria Mullin)
Hxh1	*Bos taurus*	Germany	Neolithic	8.390	X	X		X	X	TCD dataset (processed by Victoria Mullin)
Hxh2	*Bos primigenius*	Germany	Neolithic	22.470	X			X		Rossi et al., 2024
Ipa1	*Bos taurus*	North Caucasus	Bronze Age	0.010	X	X		X		This project
Irt1	*Bos taurus*	England	Bronze Age	0.900	X	X		X	X	TCD dataset (processed by Victoria Mullin)
Irt2	*Bos taurus*	England	Bronze Age	0.510	X	X		X	X	TCD dataset (processed by Victoria Mullin)
Irt3	*Bos taurus*	England	Bronze Age	1.020	X	X		X	X	TCD dataset (processed by Victoria Mullin)
Irt4	*Bos taurus*	England	Bronze Age	0.890	X	X		X	X	TCD dataset (processed by Victoria Mullin)
Irt5	*Bos taurus*	England	Bronze Age	0.250	X	X		X	X	TCD dataset (processed by Victoria Mullin)
Irt6	*Bos taurus*	England	Bronze Age	1.290	X	X		X	X	TCD dataset (processed by Victoria Mullin)
Irt7	*Bos taurus*	England	Bronze Age	0.660	X	X	X	X	X	TCD dataset (processed by Victoria Mullin)
Irt8	*Bos taurus*	England	Bronze Age	1.370	X	X		X	X	TCD dataset (processed by Victoria Mullin)
Kas1	*Bos taurus*	Germany	Bronze Age	0.040	X	X	X	X	X	This project
Kaz1	*Bos admixed*	Georgia	Medieval	3.070	X	X	X	X		Verdugo et al., 2019
Kaz2	*Bos admixed*	Georgia	Medieval	0.280	X	X	X	X		Verdugo et al., 2019
Kaz3	*Bos admixed*	Georgia	Medieval	0.450	X	X	X	X		Verdugo et al., 2019
Kaz4	*Bos admixed*	Georgia	Medieval	0.460	X	X	X	X		Verdugo et al., 2019
Kaz5	*Bos admixed*	Georgia	Medieval	0.020	X	X	X	X		Verdugo et al., 2019
Kho1	*Bos admixed*	Iran	Parthes- Islamic	0.200	X	X	X	X		Verdugo et al., 2019
Kie1	*Bos taurus*	Poland	Neolithic	0.040	X	X		X	X	TCD dataset (processed by Marta Verdugo and Victoria Mullin)
Kie2	*Bos taurus*	Poland	Neolithic	0.970	X	X		X		TCD dataset (processed by Marta Verdugo and Victoria Mullin)

Kil1	*Bos taurus*	Scotland	Medieval	1.930	X	X		X	X	TCD dataset (processed by Victoria Mullin)
Kir4	*Bos taurus*	Germany	Neolithic	14.880	X	X	X	X	X	TCD dataset (processed by Victoria Mullin)
Kir5	*Bos taurus*	Germany	Neolithic	0.060	X	X		X		This project
Kir6	*Bos taurus*	Germany	Neolithic	0.030	X	X	X	X		This project
KK1	*Bos admixed*	Cyprus	Iron Age	0.020	X	X	X	X		TCD dataset (processed by Anna Spyrou and Victoria Mullin)
Kok1	*Bos admixed*	Uzbekistan	Iron Age	94.910	X	X	X	X		Verdugo et al., 2019
KOL10	*Bos taurus*	Russia	Bronze Age	0.710	X	X	X	X		Rossi et al., 2024
KOL9	*Bos taurus*	Russia	Bronze Age	2.280	X	X	X	X		Rossi et al., 2024
Kom1	*Bos taurus*	North Caucasus	Iron Age	0.400	X	X	X	X		This project
Kun1	*Bos taurus*	Germany	Neolithic	0.050	X	X		X	X	This project
Kun2	*Bos taurus*	Germany	Neolithic	1.000	X	X		X	X	This project
Kun3	*Bos taurus*	Germany	Neolithic	1.000	X	X		X	X	This project
LARK01	*Bos taurus*	UK	Early/Middle Bronze Age	0.700	X	X		X	X	TCD dataset (processed by Victoria Mullin)
LARK02	*Bos taurus*	UK	Neolithic	0.140	X	X		X	X	TCD dataset (processed by Victoria Mullin)
Lec1	*Bos taurus*	England	Bronze Age	1.770	X	X		X	X	TCD dataset (processed by Conor Rossi and Victoria Mullin)
Lec4	*Bos taurus*	England	Bronze Age	2.460	X	X		X	X	TCD dataset (processed by Conor Rossi and Victoria Mullin)
LISM01	*Bos taurus*	Scotland	Iron Age	0.000	X	X		X	X	TCD dataset (processed by Victoria Mullin)
LISM03	*Bos taurus*	Scotland	Iron Age	0.010	X	X		X	X	TCD dataset (processed by Victoria Mullin)
Lud1	*Bos taurus*	Poland	Neolithic	0.510	X	X		X	X	TCD dataset (processed by Marta Verdugo and Victoria Mullin)
Lud2	*Bos taurus*	Poland	Neolithic	0.350	X	X		X		TCD dataset (processed by Marta Verdugo and Victoria Mullin)
Lud3	*Bos taurus*	Poland	Neolithic	1.800	X	X		X	X	TCD dataset (processed by Marta Verdugo and Victoria Mullin)
Lud4	*Bos taurus*	Poland	Neolithic	0.190	X	X		X	X	TCD dataset (processed by Marta Verdugo and Victoria Mullin)
LW017	*Bos taurus*	Netherlands	Corded-Ware	0.010	X	X		X		TCD dataset (processed by Catherine Butt and Victoria Mullin)

LW018	*Bos taurus*	Netherlands	Corded-Ware	0.010	X	X		X	X	TCD dataset (processed by Catherine Butt and Victoria Mullin)
LW019	*Bos taurus*	Netherlands	Corded-Ware	0.170	X	X		X	X	TCD dataset (processed by Catherine Butt and Victoria Mullin)
LW020	*Bos taurus*	Netherlands	Corded-Ware	0.010	X	X		X	X	TCD dataset (processed by Catherine Butt and Victoria Mullin)
LW021	*Bos taurus*	Netherlands	Corded-Ware	0.410	X	X		X	X	TCD dataset (processed by Catherine Butt and Victoria Mullin)
LW023	*Bos taurus*	Netherlands	Corded-Ware	0.750	X	X		X		TCD dataset (processed by Catherine Butt and Victoria Mullin)
LW024	*Bos taurus*	Netherlands	Corded-Ware	0.570	X	X		X		TCD dataset (processed by Catherine Butt and Victoria Mullin)
LW025	*Bos taurus*	Netherlands	Corded-Ware	0.440	X	X		X		TCD dataset (processed by Catherine Butt and Victoria Mullin)
LW026	*Bos taurus*	Netherlands	Corded-Ware	0.010	X	X		X		TCD dataset (processed by Catherine Butt and Victoria Mullin)
LW027	*Bos taurus*	Netherlands	Corded-Ware	0.010	X	X		X		TCD dataset (processed by Catherine Butt and Victoria Mullin)
LW028	*Bos taurus*	Netherlands	Corded-Ware	0.010	X	X		X	X	TCD dataset (processed by Catherine Butt and Victoria Mullin)
Lym1	*Bos taurus*	England	Saxon/Norman	4.467	X	X		X	X	TCD dataset (processed by Victoria Mullin)
Lym2	*Bos taurus*	England	Saxon/Norman	2.970	X	X		X	X	TCD dataset (processed by Victoria Mullin)
Mam1	*Bos taurus*	Ukraine	Iron Age	0.410	X	X	X	X		This project
Man1	*Bos taurus*	Germany	Iron Age	0.100	X	X		X	X	This project
Man2	*Bos taurus*	Germany	Iron Age	0.440	X	X		X	X	This project
Man3	*Bos taurus*	Germany	Iron Age	1.060	X	X		X	X	This project
Man4	*Bos taurus*	Germany	Iron Age	1.040	X	X		X	X	This project
Man5	*Bos taurus*	Germany	Iron Age	0.520	X	X		X	X	This project
Man6	*Bos taurus*	Germany	Iron Age	1.050	X	X		X	X	This project
Mantova1	*Bos primigenius*	Italy	Mesolithic	0.210	X			X		Rossi et al., 2024
Mar1	*Bos taurus*	Iran	Chalcolithic	0.040	X	X		X		Verdugo et al., 2019
Mas1	*Bos taurus*	Israel	Iron Age	0.003	X	X		X		TCD dataset (processed

										by Marta Verdugo, Kevin Daly and Victoria Mullin)
MDR1	*Bos taurus*	Netherlands	Neolithic	0.370	X	X		X		TCD dataset (processed by Victoria Mullin)
Men1	*Bos taurus*	Turkey	Neolithic	2.920	X	X		X		Verdugo et al., 2019
Men2	*Bos taurus*	Turkey	Neolithic	1.850	X	X		X		Verdugo et al., 2019
Men3	*Bos taurus*	Turkey	Neolithic	0.330	X	X		X		This project
Mon1	*Bos taurus*	Turkmenistan	Chalcolithic	0.001	X	X		X		Verdugo et al., 2019
Mon10	*Bos taurus*	Turkmenistan	Chalcolithic	0.740	X	X	X	X		This project
Mon2	*Bos taurus*	Turkmenistan	Chalcolithic	0.060	X	X		X		This project
Mon3	*Bos taurus*	Turkmenistan	Chalcolithic	0.050	X	X		X		This project
Mon4	*Bos taurus*	Turkmenistan	Chalcolithic	0.490	X	X	X	X		Rossi et al., 2024
Mon5	*Bos taurus*	Turkmenistan	Chalcolithic	0.310	X	X		X		This project
Mon6	*Bos taurus*	Turkmenistan	Chalcolithic	0.220	X	X		X		This project
Mon7	*Bos taurus*	Turkmenistan	Chalcolithic	0.440	X	X	X	X		Rossi et al., 2024
Mon8	*Bos taurus*	Turkmenistan	Chalcolithic	1.270	X	X		X		This project
Mon9	*Bos taurus*	Turkmenistan	Chalcolithic	0.460	X	X		X		This project
Nad1	*Bos taurus*	Iran	Bronze Age	0.300	X	X		X		This project
Nad2	*Bos taurus*	Iran	Bronze Age	0.310	X	X		X		This project
Nah1	*Bos taurus*	Israel	Early Bronze	0.000		X		X		Verdugo et al., 2019
Ness5	*Bos taurus*	Scotland	Neolithic	2.730	X	X		X	X	TCD dataset (processed by Victoria Mullin)
New1	*Bos taurus*	Ireland	Neolithic	16.380	X	X		X	X	TCD dataset (processed by Victoria Mullin)
New2	*Bos taurus*	Ireland	Iron Age	3.930	X	X		X	X	TCD dataset (processed by Victoria Mullin)
New3	*Bos taurus*	Ireland	Neolithic	5.200	X	X		X	X	TCD dataset (processed by Victoria Mullin)
New4	*Bos taurus*	Ireland	Neolithic	4.230	X	X		X	X	TCD dataset (processed by Victoria Mullin)
New5	*Bos taurus*	Ireland	Neolithic	4.700	X	X		X	X	TCD dataset (processed by Victoria Mullin)
NVL1	*Bos primigenius*	Russia	Botai-Tersek-Late Eneolithic	0.010	X			X		Rossi et al., 2024
NVL3	*Bos primigenius*	Russia	Botai-Tersek-Late Eneolithic	0.060	X		X	X		Rossi et al., 2024
Otb1	*Bos taurus*	Netherlands	Roman	2.580	X	X		X	X	TCD dataset (processed by Marta Verdugo and Victoria Mullin)
Padova1	*Bos primigenius*	Italy		0.164	X			X		Rossi et al., 2024

Palidoro1	*Bos primigenius*	Italy		2.563	X			X		Rossi et al., 2024
Par1	*Bos taurus*	Ireland	Chalcolithic	1.710	X	X		X	X	TCD dataset (processed by Victoria Mullin)
Par2	*Bos taurus*	Ireland	Chalcolithic	2.020	X	X		X	X	TCD dataset (processed by Victoria Mullin)
Plo1	*Bos taurus*	Serbia	Middle Neolithic	0.130	X	X		X	X	Verdugo et al., 2019
Plo2	*Bos taurus*	Serbia	Late Neolithic	0.480	X	X		X		Verdugo et al., 2019
Plo3	*Bos taurus*	Serbia	Middle Neolithic	2.630	X	X		X	X	Verdugo et al., 2019
Plo4	*Bos taurus*	Serbia	Late Neolithic	2.860	X	X		X	X	Verdugo et al., 2019
Plo5	*Bos taurus*	Serbia	Late Neolithic	0.080	X	X		X	X	Verdugo et al., 2019
Plo6	*Bos taurus*	Serbia	Neolithic	0.460	X	X		X	X	Verdugo et al., 2019
Plo7	*Bos taurus*	Serbia	Late Neolithic	0.110	X	X		X	X	Verdugo et al., 2019
Plo8	*Bos taurus*	Serbia	Late Neolithic	0.200	X	X		X	X	Verdugo et al., 2019
Poi1	*Bos taurus*	France	Medieval	3.900	X	X		X	X	TCD dataset (processed by Victoria Mullin)
Poi2	*Bos taurus*	France	Medieval	1.390	X	X		X	X	TCD dataset (processed by Victoria Mullin)
Poi3	*Bos taurus*	France	Medieval	0.830	X	X		X	X	TCD dataset (processed by Victoria Mullin)
Poi4	*Bos taurus*	France	Medieval	1.980	X	X		X		TCD dataset (processed by Victoria Mullin)
Poi5	*Bos taurus*	France	Medieval	0.380	X	X		X	X	TCD dataset (processed by Victoria Mullin)
Pot1	*Bos taurus*	England	Early Iron Age	2.550	X	X				TCD dataset (processed by Victoria Mullin)
Pot2	*Bos taurus*	England	Early Iron Age	0.440	X	X		X	X	TCD dataset (processed by Victoria Mullin)
Pot3	*Bos taurus*	England	Early Iron Age	1.600	X	X		X	X	TCD dataset (processed by Victoria Mullin)
Pot4	*Bos taurus*	England	Early Iron Age	2.640	X	X		X	X	TCD dataset (processed by Victoria Mullin)
Pot5	*Bos taurus*	England	Early Iron Age	1.720	X	X		X	X	TCD dataset (processed by Victoria Mullin)
Pot6	*Bos taurus*	England	Early Iron Age	0.350	X	X		X	X	TCD dataset (processed by Victoria Mullin)
Qab1	*Bos taurus*	Iran	Chalcolithic	0.040	X	X		X		This project
Qaz1	*Bos taurus*	Iran	Bronze Age	1.650	X	X		X		Verdugo et al., 2019
Qaz2	*Bos taurus*	Iran	Chalcolithic	0.260	X	X		X		This project
Red1	*Bos taurus*	UK	Medieval	1.680	X	X		X	X	This project
Red2	*Bos taurus*	UK	Medieval	3.310	X	X		X	X	This project

Rhi1	*Bos primigenius*	Germany		4.610	X			X		Rossi et al., 2024
Rhi2	*Bos primigenius*	Germany		2.120	X			X		Rossi et al., 2024
Rhi3	*Bos primigenius*	Germany		3.140	X			X		Rossi et al., 2024
ROS002	*Bos primigenius*	Kazakhstan	Chalcolithic	7.340	X		X	X		Rossi et al., 2024
Rou1	*Bos taurus*	Ireland	Bronze Age	2.930	X	X		X	X	TCD dataset (processed by Victoria Mullin)
Rou2	*Bos taurus*	Ireland	Bronze Age	2.980	X	X		X	X	TCD dataset (processed by Victoria Mullin)
Rou3	*Bos taurus*	Ireland	Bronze Age	1.300	X	X		X	X	TCD dataset (processed by Victoria Mullin)
Rou4	*Bos taurus*	Ireland	Bronze Age	0.980	X	X		X	X	TCD dataset (processed by Victoria Mullin)
Ruc1	*Bos taurus*	Poland	Neolithic	0.008	X	X		X	X	This project
Ruc2	*Bos taurus*	Poland	Neolithic	0.490	X	X	X	X		This project
Rug1	*Bos taurus*	UK	Iron Age	9.140	X	X		X	X	This project
Sac3	*Bos taurus*	Iran	Late Neolithic	0.190	X	X		X		Verdugo et al., 2019
Sar38	*Bos taurus*	Greece	Middle Neolithic	0.480	X	X		X		Verdugo et al., 2019
Sch1	*Bos taurus*	Netherlands	Neolithic	1.710	X	X		X	X	TCD dataset (processed by Marta Verdugo and Victoria Mullin)
Sch2	*Bos taurus*	Netherlands	Neolithic	1.270	X	X		X	X	TCD dataset (processed by Marta Verdugo and Victoria Mullin)
Sch3	*Bos taurus*	Netherlands	Neolithic	0.040	X	X		X	X	This project
Sha1	*Bos primigenius*	Ukraine	Chalcolithic	0.370	X	X		X		This project
Sha2	*Bos taurus*	Ukraine	Chalcolithic	0.440	X	X		X		This project
Shiderti1	*Bos taurus*	Kazakhstan	Early Bronze Age	0.462	X	X	X	X		Rossi et al., 2024
Shimao02	*Bos taurus*	China	Neolithic	0.040	X	X	X	X		Chen et al., 2018
Shimao04	*Bos taurus*	China	Neolithic	0.140	X	X	X	X		Chen et al., 2018
Shimao05	*Bos taurus*	China	Neolithic	2.820	X	X	X	X		Chen et al., 2018
Shimao07	*Bos taurus*	China	Neolithic	0.060	X	X	X	X		Chen et al., 2018
Shr1	*Bos taurus*	North Caucasus	Kurgan	0.000	X	X		X		This project
Shr2	*Bos taurus*	North Caucasus	Bronze Age	0.050	X	X	X	X		This project
Shu1	*Bos taurus*	South Caucasus	Neolithic	1.530	X	X		X		TCD dataset (processed by Kevin Daly and Victoria Mullin)

Ska1	*Bos primigenius*	Sweden	Mesolithic	0.370	X			X		Rossi et al., 2024
Ska3	*Bos primigenius*	Sweden	Mesolithic	0.380	X			X		Rossi et al., 2024
Snu1	*Bos taurus*	Scotland	Norse	3.420	X	X		X	X	TCD dataset (processed by Victoria Mullin)
Soh1	*Bos taurus*	Iran	Chalcolithic	0.490	X	X		X		This project
Son1	*Bos taurus*	Germany	Middle Bronze Age	0.010	X	X		X	X	This project
STB1	*Bos taurus*	Netherlands	Bronze Age	1.650	X	X		X	X	TCD dataset (processed by Conor Rossi and Victoria Mullin)
STB2	*Bos taurus*	Netherlands	Bronze Age	1.690	X	X		X	X	TCD dataset (processed by Conor Rossi and Victoria Mullin)
STB3	*Bos taurus*	Netherlands	Bronze Age	1.280	X	X		X	X	TCD dataset (processed by Conor Rossi and Victoria Mullin)
Ste1	*Bos taurus*	Netherlands	Antiquity	0.950	X	X		X	X	TCD dataset (processed by Marta Verdugo and Victoria Mullin)
STK230	*Bos taurus*	Tunisia	Iron Age	0.010	X	X		X	X	Ginja et al., 2023
STK231	*Bos taurus*	Tunisia	Iron Age	0.010	X	X		X		Ginja et al., 2023
STK234	*Bos taurus*	Tunisia	Iron Age	0.090	X	X		X		Ginja et al., 2023
STK236	*Bos taurus*	Tunisia	Iron Age	0.010	X	X		X		Ginja et al., 2023
STK239	*Bos taurus*	Tunisia	Iron Age	0.100	X	X		X	X	Ginja et al., 2023
STK240		Tunisia	Iron Age	0.000	X	X		X		Ginja et al., 2023
STK243	*Bos taurus*	Tunisia	Iron Age	0.010	X	X		X		Ginja et al., 2023
Sto1	*Bos taurus*	North Caucasus	Bronze Age	0.450	X	X		X		This project
Sto2	*Bos taurus*	North Caucasus	Bronze Age	0.020	X	X		X		This project
Sto3	*Bos taurus*	North Caucasus	Bronze Age	0.080	X	X		X		This project
Stra1	*Bos taurus*	Scotland	Bronze Age	1.250	X	X	X	X	X	TCD dataset (processed by Victoria Mullin)
Stu1	*Bos taurus*	Serbia	Late Neolithic	0.150	X	X		X	X	Verdugo et al., 2019
Sub1	*Bos taurus*	Turkey	Neolithic	13.850	X	X		X		Verdugo et al., 2019
Sut1	*Bos taurus*	Switzerland	Neolithic	1.870	X	X		X	X	TCD dataset (processed by Marta Verdugo and Victoria Mullin)
Sza1	*Bos taurus*	Hungary	Middle Bronze Age	1.540	X	X	X	X	X	TCD dataset (processed by Andrew Hare, Conor Rossi and Victoria Mullin)
Sza10	*Bos taurus*	Hungary	Early Bronze Age	0.470	X	X	X	X		This project

Sza11	*Bos taurus*	Hungary	Early Bronze Age	0.030	X	X		X	X	This project
Sza12	*Bos taurus*	Hungary	Middle Bronze Age	0.010	X	X		X		This project
Sza13	*Bos taurus*	Hungary	Early Bronze Age	0.690	X	X	X	X		This project
Sza15	*Bos taurus*	Hungary	Middle Bronze Age	0.007	X	X		X		This project
Sza2	*Bos taurus*	Hungary	Middle Bronze Age	1.630	X	X	X	X		TCD dataset (processed by Andrew Hare, Conor Rossi and Victoria Mullin)
Sza3	*Bos taurus*	Hungary	Middle Bronze Age	2.240	X	X	X	X		TCD dataset (processed by Andrew Hare, Conor Rossi and Victoria Mullin)
Sza4	*Bos taurus*	Hungary	Middle Bronze Age	1.780	X	X	X	X		TCD dataset (processed by Andrew Hare, Conor Rossi and Victoria Mullin)
Sza5	*Bos taurus*	Hungary	Middle Bronze Age	1.850	X	X	X	X		TCD dataset (processed by Andrew Hare, Conor Rossi and Victoria Mullin)
Sza6	*Bos taurus*	Hungary	Middle Bronze Age	0.480	X	X	X	X		This project
Sza7	*Bos taurus*	Hungary	Middle Bronze Age	0.470	X	X	X	X		This project
Sza8	*Bos taurus*	Hungary	Middle Bronze Age	1.020	X	X	X	X		This project
Sza9	*Bos taurus*	Hungary	Early Bronze Age	0.430	X	X	X	X		This project
Tap1	*Bos taurus*	Iran	Bronze Age	0.940	X	X		X		This project
Tar1	*Bos taurus*	Italy	Antiquity	0.009	X	X		X		This project
Tar2	*Bos taurus*	Italy	Iron Age	0.009	X	X		X	X	This project
TAW1	*Bos taurus*	Belgium	Neolithic	0.050	X	X		X	X	TCD dataset (processed by Victoria Mullin)
TAW2	*Bos taurus*	Belgium	Neolithic	0.040	X	X		X	X	TCD dataset (processed by Victoria Mullin)
Tda1	*Bos taurus*	Israel	Early Bronze Age	0.001		X		X		Verdugo et al., 2019
TEF1	*Bos taurus*	Lebanon	Bronze Age	0.290	X	X		X		TCD dataset (processed by Conor Rossi and Victoria Mullin)
Tes1	*Bos taurus*	Israel	Early Bronze Age	0.040	X	X	X	X		This project
Th7	*Bos primigenius*	Morocco	Paleolithic	8.400	X		X	X		Verdugo et al., 2019
Thr1	*Bos taurus*	Israel	Neolithic	0.001	X	X		X		Verdugo et al., 2019

Til1	*Bos taurus*	Turkey	Chalcolithic/ Bronze Age	0.060	X	X	X	X		This project
Tmq1	*Bos admixed*	Israel	Iron Age	0.003		X		X		Verdugo et al., 2019
Tmq2	*Bos admixed*	Israel	Iron Age	0.004		X		X		Verdugo et al., 2019
Tmq3	*Bos admixed*	Israel	Late Bronze-Iron Age	0.020	X	X		X		Verdugo et al., 2019
Tmq4	*Bos admixed*	Israel	Late Bronze-Iron Age	0.020	X	X	X	X		Verdugo et al., 2019
Tqa1	*Bos taurus*	Israel	Early Bronze Age	0.003		X		X		Verdugo et al., 2019
Tqa2	*Bos taurus*	Israel	Bronze Age	0.560	X	X		X		Verdugo et al., 2019
Tqa3	*Bos taurus*	Israel	Bronze Age	0.020	X	X		X		Verdugo et al., 2019
Tri1	*Bos primigenius*	Germany	Mesolithic	2.240	X			X		Rossi et al., 2024
Tsa1	*Bos admixed*	Israel	Iron Age	3.470	X	X		X		Verdugo et al., 2019
Tsa2	*Bos admixed*	Israel	Iron Age	1.000	X	X		X		Verdugo et al., 2019
Tsa3	*Bos taurus*	Israel	Iron Age	0.020	X	X		X		Verdugo et al., 2019
Tula1	*Bos primigenius*	Russia	Paleolithic	3.470	X		X	X		Rossi et al., 2024
TWAN01	*Bos taurus*	Switzerland	Neolithic	0.010	X	X		X	X	TCD dataset (processed by Victoria Mullin)
TWAN02	*Bos taurus*	Switzerland	Neolithic	0.040	X	X		X	X	TCD dataset (processed by Victoria Mullin)
TWAN03	*Bos taurus*	Switzerland	Neolithic	0.010	X	X		X		TCD dataset (processed by Victoria Mullin)
TWAN04	*Bos taurus*	Switzerland	Neolithic	0.010	X	X		X	X	TCD dataset (processed by Victoria Mullin)
TWAN05	*Bos taurus*	Switzerland	Neolithic	0.010	X	X		X		TCD dataset (processed by Victoria Mullin)
Tyq1	*Bos admixed*	Israel	Early Hellenistic	11.000	X	X		X		Verdugo et al., 2019
Ulu1	*Bos taurus*	Turkey	Neolithic	0.190	X	X		X		This project
Ulu2	*Bos taurus*	Turkey	Neolithic	0.460	X	X		X		This project
Ulu3	*Bos taurus*	Turkey	Neolithic	0.550	X	X		X		This project
Ulu4	*Bos taurus*	Turkey	Neolithic	0.500	X	X		X		This project
Ulu5	*Bos taurus*	Turkey	Neolithic	0.090	X	X		X		This project
Uralsk1	*Bos primigenius*	Kazakhstan	Unknown	0.010	X			X		Rossi et al., 2024
Ust1	*Bos taurus*	North Caucasus	Bronze Age	0.020	X	X	X	X		This project
Uzzo1	*Bos primigenius*	Italy		1.996	X			X		Rossi et al., 2024
Var1	*Bos primigenius*	Russia	Neolithic	4.550	X		X	X		Rossi et al., 2024

Var2	*Bos primigenius*	Russia	Neolithic	8.250	X		X	X		Rossi et al., 2024
VEM329	*Bos taurus*	Syria	Early Bronze Age	0.470	X	X		X		TCD dataset (processed by Catherine Butt and Victoria Mullin)
VEM330	*Bos taurus*	Syria	Early Bronze Age	0.370	X	X		X		TCD dataset (processed by Catherine Butt and Victoria Mullin)
VEM331	*Bos taurus*	Syria	Chalcolithic	0.280	X	X		X		TCD dataset (processed by Catherine Butt and Victoria Mullin)
VEM337	*Bos taurus*	Ethiopia	Classical	0.010	X	X		X		TCD dataset (processed by Catherine Butt and Victoria Mullin)
VEM338	*Bos taurus*	Syria	Chalcolithic	0.000	X	X		X		TCD dataset (processed by Catherine Butt and Victoria Mullin)
VEM341	*Bos taurus*	Syria	Neolithic	0.000	X	X		X		TCD dataset (processed by Catherine Butt and Victoria Mullin)
Viz1	*Bos taurus*	Netherlands	Bronze Age	0.710	X	X		X	X	TCD dataset (processed by Marta Verdugo and Victoria Mullin)
Viz2	*Bos taurus*	Netherlands	Medieval	5.750	X	X		X	X	TCD dataset (processed by Marta Verdugo and Victoria Mullin)
Vratsa1	*Bos primigenius*	Bulgaria		9.026	X			X		Rossi et al., 2024
Vratsa2	*Bos primigenius*	Bulgaria		0.994	X			X		Rossi et al., 2024
Wal1	*Bos taurus*	UK	Medieval	0.940	X	X		X	X	This project
Win1	*Bos taurus*	Netherlands	Early Medieval	18.690	X	X		X	X	Erven et al., 2024
Wro1	*Bos taurus*	UK	Iron Age	1.630	X	X		X	X	This project
Yer1	*Bos admixed*	Uzbekistan	Achaemenid/Sasanid	1.680	X	X	X	X		Verdugo et al., 2019
YoA	*Bos primigenius*	England		6.530	X			X		Rossi et al., 2024
Yor1	*Bos taurus*	England	Roman	7.850	X	X		X	X	TCD dataset (processed by Victoria Mullin)
Yor11	*Bos taurus*	England	Roman	0.980	X	X		X	X	TCD dataset (processed by Victoria Mullin)
Yor12	*Bos taurus*	England	Medieval	4.480	X	X		X	X	TCD dataset (processed by Victoria Mullin)
Yor2	*Bos taurus*	England	Medieval	3.910	X	X		X	X	TCD dataset (processed by Victoria Mullin)
Yor3	*Bos taurus*	England	Medieval	2.500	X	X		X	X	TCD dataset (processed

										by Victoria Mullin)
Yor4	*Bos taurus*	England	Medieval	3.140	X	X		X	X	TCD dataset (processed by Victoria Mullin)
Yor5	*Bos taurus*	England	Medieval	2.960	X	X		X	X	TCD dataset (processed by Victoria Mullin)
Yor6	*Bos taurus*	England	Medieval	6.810	X	X		X	X	TCD dataset (processed by Victoria Mullin)
Yor7	*Bos taurus*	England	Medieval	6.630	X	X		X	X	TCD dataset (processed by Victoria Mullin)
Zag1	*Bos taurus*	Iran	Chalcolithic	0.210	X	X		X		This project
Zah1	*Bos admixed*	Israel	Islamic	0.030	X	X	X	X		Verdugo et al., 2019
Zea1	*Bos primigenius*	Denmark	Neolithic	0.220	X			X		Rossi et al., 2024
Zea2	*Bos primigenius*	Denmark	Neolithic	0.070	X			X		Rossi et al., 2024
Zha1	*Bos taurus*	Ukraine	Iron Age	0.390	X	X	X	X		This project
Zha2	*Bos taurus*	Ukraine	Iron Age	0.020	X	X	X	X		This project

Supplementary Table S5. List of modern taurine (*Bos taurus*) and indicine (*Bos indicus*) samples used as reference dataset for the LASER projections of Figures 22 and 23.

Sample ID	Breed	Context	Species	PC1	PC2
ERR3305588	Alentejana	Modern	Taurine	-565.151	3.124
SRR1355237	Angus	Modern	Taurine	-593.403	119.093
SRR1365144	Angus	Modern	Taurine	-602.896	117.701
SRR1425124	Angus	Modern	Taurine	-598.637	116.221
SRR1425145	Angus	Modern	Taurine	-583.152	103.535
SRR3656915	Ankole	Modern	Hybrid	548.892	-486.582
SRR3656921	Ankole	Modern	Hybrid	589.639	-473.518
ERR4472501	Baoule	Modern	Taurine	537.450	-635.087
ERR4472506	Baoule	Modern	Taurine	-438.734	-591.999
ERR4472507	Baoule	Modern	Taurine	-299.336	-613.171
ERR4472508	Baoule	Modern	Taurine	-324.250	-594.649
ERR4472510	Baoule	Modern	Taurine	-462.905	-584.565
ERR4472515	Baoule	Modern	Taurine	-417.210	-584.826
ERR4472516	Baoule	Modern	Taurine	-471.792	-601.512
XH8	Bashan	Modern	Hybrid	949.516	274.115

XH9	Bashan	Modern	Hybrid	911.135	258.993
XZ1	Bashan	Modern	Hybrid	736.415	229.323
XZ2	Bashan	Modern	Hybrid	796.632	216.681
XZ3	Bashan	Modern	Hybrid	698.621	220.709
SAMN05803879	BeltedGalloway	Modern	Taurine	-595.506	111.191
SAMN05803880	BeltedGalloway	Modern	Taurine	-591.246	109.553
BH1	BohaiBlack	Modern	Hybrid	318.271	203.812
BH4	BohaiBlack	Modern	Hybrid	-43.008	133.516
BH5	BohaiBlack	Modern	Hybrid	-136.007	130.528
BH8	BohaiBlack	Modern	Hybrid	-352.945	35.025
SRR8426534	Boskarin	Modern	Taurine	-348.523	13.561
SRR6650022	Brahman	Modern	Indicine	1890.730	-494.869
SRR6650023	Brahman	Modern	Indicine	1861.580	-490.239
SRR6650031	Brahman	Modern	Indicine	1885.360	-486.668
ERR1746314	BrownSwiss	Modern	Taurine	-559.632	64.555
ERR1746315	BrownSwiss	Modern	Taurine	-571.426	61.926
ERR1747023	BrownSwiss	Modern	Taurine	-575.659	67.062
ERR1766310	BrownSwiss	Modern	Taurine	-555.797	61.818
SRR5507272	ChaidamuYellow	Modern	Taurine	-218.221	19.289
SRR1343169	Charolais	Modern	Taurine	-478.221	55.235
SRR1348584	Charolais	Modern	Taurine	-476.372	56.751
SRR1355254	Charolais	Modern	Taurine	-473.028	53.767
SRR1355258	Charolais	Modern	Taurine	-543.302	65.501
SAMN05216091	Chianina	Modern	Taurine	-373.668	11.822
SRR5507191	Dabieshan	Modern	Hybrid	909.577	363.108
SRR5507192	Dabieshan	Modern	Hybrid	215.078	171.213
SRR1346378	Salers	Modern	Taurine	-557.965	96.342
SAMN05803883	Dexter	Modern	Taurine	-597.695	108.066
SAMN05803888	Dexter	Modern	Taurine	-573.261	98.839
ERR2734942	Finncattle	Modern	Taurine	-521.059	64.844
ERR2734943	Finncattle	Modern	Taurine	-514.888	60.387
ERR2734944	Finncattle	Modern	Taurine	-505.048	62.698

ERR2734945	Finncattle	Modern	Taurine	-516.359	68.777
ERR2734946	Finncattle	Modern	Taurine	-499.953	63.227
ERR2734948	Finncattle	Modern	Taurine	-509.373	69.655
ERR2734949	Finncattle	Modern	Taurine	-517.357	68.869
ERR2734950	Finncattle	Modern	Taurine	-498.861	66.776
ERR2734951	Finncattle	Modern	Taurine	-502.181	64.941
SRR1346388	Gelbvieh	Modern	Taurine	-542.696	72.155
SRR1348575	Gelbvieh	Modern	Taurine	-545.053	59.443
SRR1355236	Gelbvieh	Modern	Taurine	-500.632	74.793
SRR1365112	Gelbvieh	Modern	Taurine	-565.073	63.247
SRR2016752	Gir	Modern	Indicine	1720.110	-446.259
ERR4494254	Gourounsi	Modern	Taurine	566.761	-623.075
ERR4494255	Gourounsi	Modern	Taurine	-228.402	-593.961
ERR4494323	Gourounsi	Modern	Taurine	33.356	-598.555
SRR934415	Hanwoo	Modern	Taurine	-434.583	39.816
SRR934417	Hanwoo	Modern	Taurine	-417.744	40.051
SRR934418	Hanwoo	Modern	Taurine	-437.183	35.020
SRR934419	Hanwoo	Modern	Taurine	-439.812	28.948
SRR934432	Hanwoo	Modern	Taurine	-458.829	34.815
Har03	Hariana	Modern	Indicine	2056.270	-485.860
SRR1365124	Hereford	Modern	Taurine	-603.313	115.167
SRR1365126	Hereford	Modern	Taurine	-599.897	117.637
SRR1365128	Hereford	Modern	Taurine	-601.619	116.140
SRR1365131	Hereford	Modern	Taurine	-602.219	118.821
SRR1365137	Hereford	Modern	Taurine	-559.326	113.386
ERR3305587	ScottishHighland	Modern	Taurine	-597.813	111.329
SRR1346386	Holstein	Modern	Taurine	-566.933	85.565
SRR1348583	Holstein	Modern	Taurine	-591.828	85.301
SRR1425129	Holstein	Modern	Taurine	-594.509	93.728
SRR1425133	Holstein	Modern	Taurine	-591.517	93.881
SRR1425144	Holstein	Modern	Taurine	-587.865	92.337
SRR3497161	Jersey	Modern	Taurine	-587.818	97.268

SRR3497162	Jersey	Modern	Taurine	-583.180	99.527
SRR3497451	Jersey	Modern	Taurine	-561.955	90.248
SRR3497462	Jersey	Modern	Taurine	-558.886	90.656
SRR3497467	Jersey	Modern	Taurine	-590.759	97.649
SRR3497611	Jersey	Modern	Taurine	-569.888	95.075
SRR5507281	Jian	Modern	Indicine	1688.460	566.900
SRR5507282	Jian	Modern	Indicine	1824.090	626.615
SRR5507285	Jiangxi	Modern	Indicine	1625.250	517.893
SRR5507286	Jiangxi	Modern	Indicine	1515.170	512.689
JX10M	JiaxianRed	Modern	Hybrid	87.505	137.945
JX1M	JiaxianRed	Modern	Hybrid	497.252	228.577
JX21	JiaxianRed	Modern	Hybrid	260.258	154.301
JX6M	JiaxianRed	Modern	Hybrid	351.975	193.606
JX8M	JiaxianRed	Modern	Hybrid	286.185	163.037
HSK1	Kazakh	Modern	Taurine	-277.192	17.304
HSK25	Kazakh	Modern	Taurine	-349.687	39.548
HSK26	Kazakh	Modern	Taurine	-266.483	10.590
HSK29	Kazakh	Modern	Taurine	-354.012	20.766
HSK8	Kazakh	Modern	Taurine	-354.454	35.392
SRR3694651	Kenana	Modern	Hybrid	1018.170	-489.224
SRR3694652	Kenana	Modern	Hybrid	1009.170	-494.044
SRR3694653	Kenana	Modern	Hybrid	1050.650	-486.242
SRR3694654	Kenana	Modern	Hybrid	995.169	-466.954
SRR3694655	Kenana	Modern	Hybrid	1031.600	-489.520
ERR3293557	Lagune	Modern	Taurine	-339.498	-554.479
SRR1365134	Limousin	Modern	Taurine	-501.837	52.784
SRR1365136	Limousin	Modern	Taurine	-503.950	56.633
SXLN13	Lingnan	Modern	Hybrid	863.729	286.893
SXLN19	Lingnan	Modern	Hybrid	936.314	303.190
SXLN26	Lingnan	Modern	Hybrid	957.480	298.371
SXLN6	Lingnan	Modern	Hybrid	863.052	278.608
LX78	Luxi	Modern	Hybrid	660.749	319.588

LX80	Luxi	Modern	Hybrid	141.555	175.593
LX85	Luxi	Modern	Hybrid	417.212	267.338
SRR1355238	MaineAnjou	Modern	Taurine	-583.396	104.344
SRR1355245	MaineAnjou	Modern	Taurine	-591.212	99.420
SRR1355259	MaineAnjou	Modern	Taurine	-593.627	101.549
SRR1365108	MaineAnjou	Modern	Taurine	-586.278	103.471
SRR1365125	MaineAnjou	Modern	Taurine	-596.234	108.022
SRR8426543	Maremmana	Modern	Taurine	-316.953	-3.579
Mishima06	Mishima	Modern	Taurine	-400.995	25.878
MG12	Mongolian	Modern	Taurine	-215.436	10.191
MG14	Mongolian	Modern	Taurine	-78.940	-9.474
MG17	Mongolian	Modern	Taurine	-28.435	-13.626
MG37	Mongolian	Modern	Taurine	-188.493	0.053
ERR883698	Montbeliarde	Modern	Taurine	-505.330	55.749
ERR883699	Montbeliarde	Modern	Taurine	-531.723	59.012
ERR883701	Montbeliarde	Modern	Taurine	-499.356	61.599
SRR5507237	Nanyang	Modern	Taurine	247.069	191.458
SRR3693420	Ndama	Modern	Taurine	-26.701	-477.286
SRR3694478	Ndama	Modern	Taurine	-84.099	-488.176
SRR3694578	Ndama	Modern	Taurine	-96.428	-477.390
ERR883704	Normande	Modern	Taurine	-542.151	77.600
ERR883714	Normande	Modern	Taurine	-514.685	72.330
SRR3225526	Ogaden	Modern	Hybrid	1095.470	-491.719
SRR3490689	Ogaden	Modern	Hybrid	1066.540	-482.735
SRR3491019	Ogaden	Modern	Hybrid	1072.830	-497.714
SRR8426541	Pajuna	Modern	Taurine	-471.081	-5.215
SRR8426542	Pajuna	Modern	Taurine	-481.128	24.514
ERR454987	Rashoki	Modern	Hybrid	452.804	-174.754
ERR454988	Rashoki	Modern	Hybrid	-9.405	-42.681
ERR454989	Rashoki	Modern	Hybrid	331.379	-156.576
ERR454991	Rashoki	Modern	Hybrid	-84.045	-38.607
ERR454992	Rashoki	Modern	Hybrid	229.942	-112.879

SRR8426538	Sayaguesa	Modern	Taurine	-478.975	15.800
ERR1747025	ScottishHighland	Modern	Taurine	-586.524	106.433
ERR1747043	ScottishHighland	Modern	Taurine	-569.629	107.708
Sha3b	Sahiwal	Modern	Indicine	2045.210	-476.630
SAMN05216073	Shorthorm	Modern	Taurine	-513.977	97.849
ERR3293559	Sikia	Modern	Taurine	-299.351	-16.871
SRR1525700	Simmental	Modern	Taurine	-557.342	76.237
SRR1525701	Simmental	Modern	Taurine	-546.222	65.367
SRR1525702	Simmental	Modern	Taurine	-560.810	66.231
SRR1525703	Simmental	Modern	Taurine	-556.201	80.055
SRR1525705	Simmental	Modern	Taurine	-572.112	78.762
ERR3305591	Somba	Modern	Taurine	-405.564	-554.959
Thar1	Tharparkar	Modern	Indicine	2003.830	-487.298
ERR3293561	Wagyu	Modern	Taurine	-524.053	48.763
WN10	Wannan	Modern	Indicine	1719.480	643.403
WN11	Wannan	Modern	Indicine	1825.330	722.436
WN4	Wannan	Modern	Indicine	2022.560	749.840
WN8	Wannan	Modern	Indicine	2001.600	739.277
WN9	Wannan	Modern	Indicine	1692.770	639.752
Xizang11	Tibetan	Modern	Taurine	-335.126	29.102
Xizang19	Tibetan	Modern	Taurine	-294.237	18.712
Xizang21	Tibetan	Modern	Taurine	-316.118	30.775
ERR2734952	Yakutian	Modern	Taurine	-317.168	39.876
ERR2734953	Yakutian	Modern	Taurine	-321.042	32.296
ERR2734954	Yakutian	Modern	Taurine	-315.139	30.271
ERR2734955	Yakutian	Modern	Taurine	-299.661	27.201
ERR2734956	Yakutian	Modern	Taurine	-305.015	31.477
A83	Yanbian	Modern	Taurine	-382.312	32.907

Supplementary Table S6. List of modern taurine samples used as reference dataset for the LASER projections of Figure 23.

Sample ID	Breed	Context	Species	PC1	PC2
ERR3305588	Alentejana	Modern	Taurine	-81.938	115.015
ERR3305587	ScottishHighland	Modern	Taurine	-277.192	161.744
ERR3293557	Lagune	Modern	Taurine	879.386	216.241
ERR3293559	Sikia	Modern	Taurine	33.704	-115.696
ERR3305591	Somba	Modern	Taurine	872.496	242.398
ERR3293561	Wagyu	Modern	Taurine	-64.678	-370.603
SRR1365112	Gelbvieh	Modern	Taurine	-181.242	107.007
SRR1346388	Gelbvieh	Modern	Taurine	-194.308	111.644
SRR1348575	Gelbvieh	Modern	Taurine	-183.128	108.575
SRR1355236	Gelbvieh	Modern	Taurine	-186.380	112.972
SRR1355237	Angus	Modern	Taurine	-280.003	164.597
SRR1365144	Angus	Modern	Taurine	-286.222	172.151
SRR1425124	Angus	Modern	Taurine	-279.856	167.806
SRR1425145	Angus	Modern	Taurine	-258.999	163.088
ERR2734942	Finncattle	Modern	Taurine	-166.098	41.888
ERR2734943	Finncattle	Modern	Taurine	-156.839	39.515
ERR2734944	Finncattle	Modern	Taurine	-163.688	46.556
ERR2734945	Finncattle	Modern	Taurine	-171.705	42.780
ERR2734946	Finncattle	Modern	Taurine	-160.818	46.273
ERR2734948	Finncattle	Modern	Taurine	-172.236	59.964
ERR2734949	Finncattle	Modern	Taurine	-178.277	54.428
ERR2734950	Finncattle	Modern	Taurine	-171.790	59.318
ERR2734951	Finncattle	Modern	Taurine	-170.778	56.086
SRR1346386	SRR1346386	Modern	Taurine	-226.793	126.481
SRR1348583	SRR1348583	Modern	Taurine	-229.810	124.387
SRR1425129	SRR1425129	Modern	Taurine	-248.579	140.038
SRR1425133	SRR1425133	Modern	Taurine	-246.653	128.433
SRR1425144	SRR1425144	Modern	Taurine	-239.408	125.543
Mishima06	Mishima06	Modern	Taurine	-15.555	-417.318

SRR3497161	SRR3497161	Modern	Taurine	-283.694	299.183
SRR3497162	SRR3497162	Modern	Taurine	-286.933	306.968
SRR3497451	SRR3497451	Modern	Taurine	-266.557	293.102
SRR3497462	SRR3497462	Modern	Taurine	-264.021	278.958
SRR3497467	SRR3497467	Modern	Taurine	-292.351	318.034
SRR3497611	SRR3497611	Modern	Taurine	-277.948	303.648
SRR1365134	SRR1365134	Modern	Taurine	-146.019	105.742
SRR1525705	SRR1525705	Modern	Taurine	-219.165	155.257
SRR1525700	SRR1525700	Modern	Taurine	-219.607	141.725
SRR1525701	SRR1525701	Modern	Taurine	-189.513	121.370
SRR1525702	SRR1525702	Modern	Taurine	-193.954	124.700
SRR1525703	SRR1525703	Modern	Taurine	-204.945	128.528
SRR1343169	SRR1343169	Modern	Taurine	-157.073	105.401
SRR1348584	SRR1348584	Modern	Taurine	-157.582	104.316
SRR1355254	SRR1355254	Modern	Taurine	-149.878	96.158
SRR1355258	SRR1355258	Modern	Taurine	-181.206	97.152
Xizang11	Xizang11	Modern	Taurine	73.875	-502.354
Xizang19	Xizang19	Modern	Taurine	78.468	-465.863
Xizang21	Xizang21	Modern	Taurine	81.663	-513.472
ERR2734952	ERR2734952	Modern	Taurine	71.564	-810.084
ERR2734953	ERR2734953	Modern	Taurine	74.976	-826.534
ERR2734954	ERR2734954	Modern	Taurine	68.078	-779.022
ERR2734955	ERR2734955	Modern	Taurine	85.076	-815.650
ERR2734956	ERR2734956	Modern	Taurine	62.352	-749.743
SRR934415	SRR934415	Modern	Taurine	-1.289	-461.270
SRR934417	SRR934417	Modern	Taurine	-8.282	-450.777
SRR934418	SRR934418	Modern	Taurine	-1.819	-501.569
SRR934419	SRR934419	Modern	Taurine	0.266	-456.440
SRR934432	SRR934432	Modern	Taurine	0.477	-515.002
A83	A83	Modern	Taurine	2.403	-365.668
HSK25	HSK25	Modern	Taurine	-42.706	-134.429
HSK26	HSK26	Modern	Taurine	24.884	-224.813

HSK1	HSK1	Modern	Taurine	17.468	-177.963
HSK29	HSK29	Modern	Taurine	-46.702	-96.104
HSK8	HSK8	Modern	Taurine	-44.853	-123.638
SRR3693420	SRR3693420	Modern	Taurine	884.164	0.497
SRR3694478	SRR3694478	Modern	Taurine	880.642	14.820
SRR3694578	SRR3694578	Modern	Taurine	847.180	29.733
ERR1746314	ERR1746314	Modern	Taurine	-196.981	111.633
ERR1746315	ERR1746315	Modern	Taurine	-203.877	124.208
ERR1747023	ERR1747023	Modern	Taurine	-209.196	120.931
ERR1766310	ERR1766310	Modern	Taurine	-195.070	113.466
ERR4472506	ERR4472506	Modern	Taurine	925.748	304.457
ERR4472507	ERR4472507	Modern	Taurine	983.102	231.088
ERR4472508	ERR4472508	Modern	Taurine	977.730	239.813
ERR4472510	ERR4472510	Modern	Taurine	918.931	318.014
ERR4472515	ERR4472515	Modern	Taurine	921.686	286.415
ERR4472516	ERR4472516	Modern	Taurine	945.071	340.477
ERR4494255	ERR4494255	Modern	Taurine	989.236	168.478
ERR4494323	ERR4494323	Modern	Taurine	1050.160	-32.586
ERR883698	Montbeliarde	Modern	Taurine	-172.841	105.224
ERR883699	ERR883699	Modern	Taurine	-183.159	113.417
ERR883701	ERR883701	Modern	Taurine	-176.040	102.697
ERR883704	ERR883704	Modern	Taurine	-194.324	109.181
ERR883714	ERR883714	Modern	Taurine	-184.392	96.189
ERR1747025	ERR1747025	Modern	Taurine	-264.659	154.418
ERR1747043	ERR1747043	Modern	Taurine	-259.703	153.029
SRR1365136	SRR1365136	Modern	Taurine	-150.450	110.355
SRR8426534	SRR8426534	Modern	Taurine	-45.927	-32.887
SRR8426538	SRR8426538	Modern	Taurine	-87.558	96.056
SRR8426541	SRR8426541	Modern	Taurine	-57.766	95.509
SRR8426542	SRR8426542	Modern	Taurine	-108.234	99.576
SRR8426543	SRR8426543	Modern	Taurine	-4.486	-59.744
SRR5507272	SRR5507272	Modern	Taurine	86.740	-381.133

SAMN05803880	SAMN05803880	Modern	Taurine	-264.382	148.157
SAMN05803879	SAMN05803879	Modern	Taurine	-268.942	155.673
SAMN05803883	SAMN05803883	Modern	Taurine	-258.982	155.797
SAMN05803888	SAMN05803888	Modern	Taurine	-245.465	150.925
SAMN05216073	SAMN05216073	Modern	Taurine	-213.990	127.357
SRR1346378	SRR1346378	Modern	Taurine	-229.660	118.922
SAMN05216091	SAMN05216091	Modern	Taurine	-32.369	-31.314

Supplementary Table S7. D-statistics assessing indicine introgression in ancient taurine samples.

Individuals with at least 500 ABBA+BABA sites were plotted in Figure 26.

Sample	nABBA	nBABA	Total number of sites	D	SE	Z score
Aak1	161	99	260	0.238	0.061	3.883
Abu1	19	19	38	0.000	0.165	0.000
Abu2	15	19	34	-0.118	0.173	-0.680
Ace1	21263	21493	42756	-0.005	0.007	-0.820
ALT232	686	630	1316	0.043	0.028	1.510
ALT235	295	298	593	-0.005	0.042	-0.121
Ash4	1324	2597	3921	-0.325	0.017	-19.374
Asp1	74902	69655	144557	0.036	0.005	6.963
Azer1	156	129	285	0.095	0.060	1.568
Bal1	83101	77321	160422	0.036	0.005	6.907
Bal2	77017	71454	148471	0.037	0.006	6.640
Bel1	6787	6427	13214	0.027	0.010	2.663
Bel2	9985	9564	19549	0.022	0.008	2.558
Bes1	98673	196431	295104	-0.331	0.008	-40.735
Bes2	152054	394449	546503	-0.444	0.007	-60.953
Bis1	82762	77120	159882	0.035	0.005	6.426
Bis2	81532	76404	157936	0.032	0.006	5.867
Bisk1	79644	76948	156592	0.017	0.005	3.430
Bla1	19156	18315	37471	0.022	0.007	3.238

Bla2	4756	4184	8940	0.064	0.012	5.539
Bla3	1542	1308	2850	0.082	0.019	4.376
Bol1	37638	39088	76726	-0.019	0.006	-3.237
Bor1	85045	78331	163376	0.041	0.005	7.819
Bri1	92573	88972	181545	0.020	0.006	3.479
Brn1	55275	50569	105844	0.044	0.006	7.546
Brn2	59868	56409	116277	0.030	0.006	5.335
Brn3	84227	78526	162753	0.035	0.005	6.706
Brn4	59220	55612	114832	0.031	0.005	5.798
Brn5	64326	60517	124843	0.031	0.005	5.645
Bub1	42367	39086	81453	0.040	0.005	7.443
Bub2	7771	7704	15475	0.004	0.010	0.435
Bul1	5425	11056	16481	-0.342	0.011	-29.845
Bul2	46582	89593	136175	-0.316	0.009	-33.508
Bun1	71357	66726	138083	0.034	0.006	5.915
Bun2	78947	74592	153539	0.028	0.005	5.372
BZL1	73903	77076	150979	-0.021	0.006	-3.698
BZL2	80952	83096	164048	-0.013	0.006	-2.281
Cas1	7708	17515	25223	-0.389	0.010	-39.872
Cer1	399	432	831	-0.040	0.035	-1.141
Cer2	7485	7017	14502	0.032	0.011	3.047
Cer3	36471	34932	71403	0.022	0.006	3.438
Cer4	12910	11925	24835	0.040	0.009	4.543
Cha1	78786	72641	151427	0.041	0.006	7.149
Che1	27891	27786	55677	0.002	0.007	0.275
Cla1	93382	86687	180069	0.037	0.006	6.755
Cla2	75356	68773	144129	0.046	0.006	8.031
Cla3	12143	10840	22983	0.057	0.009	6.640
Cla4	24255	22428	46683	0.039	0.007	5.720
Cla5	85267	78284	163551	0.043	0.005	7.940
Cla6	26238	24335	50573	0.038	0.007	5.666
Cla7	22421	20720	43141	0.039	0.007	5.957

Cla8	83926	76676	160602	0.045	0.005	8.333
Cla9	85431	78208	163639	0.044	0.005	8.220
Cri1	45880	42093	87973	0.043	0.006	7.287
Cuk1	6085	6115	12200	-0.002	0.010	-0.238
Cuk2	24586	24467	49053	0.002	0.006	0.391
Da1	92725	85128	177853	0.043	0.005	7.991
Da2	20404	19196	39600	0.031	0.007	4.302
Da3	79502	74109	153611	0.035	0.006	6.062
Da4	90855	85060	175915	0.033	0.006	5.821
Da5	90748	84139	174887	0.038	0.005	6.924
Da6	86244	80312	166556	0.036	0.005	6.687
Dam1	646	639	1285	0.005	0.030	0.182
Dan1	33	31	64	0.031	0.121	0.259
Dav1	47056	49640	96696	-0.027	0.006	-4.626
Dav2	62573	64587	127160	-0.016	0.006	-2.795
Den1	84481	80995	165476	0.021	0.006	3.540
Den2	86750	81760	168510	0.030	0.006	5.322
Dro1	84275	78648	162923	0.035	0.005	6.407
Dub1	76621	70412	147033	0.042	0.005	8.596
Dur1	78649	73750	152399	0.032	0.005	6.191
Dur2	76622	71464	148086	0.035	0.005	6.583
Dur3	27093	25463	52556	0.031	0.007	4.632
Dyr1	77992	77921	155913	0.000	0.006	0.082
Dzh1	83528	82476	166004	0.006	0.005	1.174
Dzh2	79885	78864	158749	0.006	0.005	1.187
Els1	81216	77035	158251	0.026	0.005	5.046
Enk1	73585	68038	141623	0.039	0.006	6.767
Enk2	21125	19373	40498	0.043	0.007	6.378
Enk3	83761	77270	161031	0.040	0.005	7.826
Enk4	77871	73267	151138	0.030	0.005	5.668
Enk5	20784	19865	40649	0.023	0.007	3.386
Far1	8120	8536	16656	-0.025	0.008	-2.944

Fir1	84739	80593	165332	0.025	0.006	4.544
Fis1	83111	75828	158939	0.046	0.005	8.642
Fis2	82138	76019	158157	0.039	0.005	7.241
For1	32480	49136	81616	-0.204	0.009	-21.627
Gil1	11	14	25	-0.120	0.203	-0.592
GQ1	6068	5628	11696	0.038	0.011	3.528
Gub1	55185	53500	108685	0.016	0.006	2.695
Gub2	35330	34263	69593	0.015	0.006	2.510
Gub3	28434	27370	55804	0.019	0.006	3.066
Gub4	29524	28087	57611	0.025	0.006	4.166
Gub5	33762	31543	65305	0.034	0.006	5.502
Had1	70010	65541	135551	0.033	0.006	5.960
Has1	2760	3305	6065	-0.090	0.014	-6.440
Has3	56812	57907	114719	-0.010	0.005	-1.774
Has4	59325	71940	131265	-0.096	0.008	-11.568
Has5	2963	3408	6371	-0.070	0.014	-4.909
Has6	3881	4640	8521	-0.089	0.013	-7.020
Has7	46504	97748	144252	-0.355	0.008	-41.972
Has8	36100	39378	75478	-0.043	0.008	-5.545
Has9	27458	29949	57407	-0.043	0.007	-5.887
HF1	12426	10976	23402	0.062	0.008	7.488
Hou1	82223	79084	161307	0.019	0.005	3.547
Hou2	81385	79180	160565	0.014	0.006	2.368
Hxh1	85344	82878	168222	0.015	0.005	2.747
Ipa1	596	566	1162	0.026	0.030	0.847
Irt1	51224	47674	98898	0.036	0.005	6.530
Irt2	34202	31448	65650	0.042	0.006	6.797
Irt3	56280	53167	109447	0.028	0.005	5.275
Irt4	50096	47038	97134	0.031	0.006	5.508
Irt5	18604	17262	35866	0.037	0.007	5.675
Irt6	64987	59258	124245	0.046	0.005	8.753
Irt7	41659	38512	80171	0.039	0.005	7.188

Irt8	65267	60814	126081	0.035	0.006	6.230
Kas1	3038	2866	5904	0.029	0.014	2.111
Kaz1	91215	130565	221780	-0.177	0.008	-21.580
Kaz2	25184	41671	66855	-0.247	0.009	-27.841
Kaz3	34398	49816	84214	-0.183	0.009	-20.252
Kaz4	36491	56538	93029	-0.215	0.009	-24.722
Kaz5	2514	4414	6928	-0.274	0.014	-19.000
Kho1	26076	65388	91464	-0.430	0.008	-55.904
Kie1	3508	3320	6828	0.028	0.013	2.069
Kie2	52639	49809	102448	0.028	0.005	5.193
Kil1	75891	68911	144802	0.048	0.005	8.948
Kir4	89894	89408	179302	0.003	0.005	0.516
Kir5	4908	4852	9760	0.006	0.011	0.531
Kir6	2328	2317	4645	0.002	0.016	0.149
KK1	2421	4543	6964	-0.305	0.013	-23.016
Kok1	48685	117200	165885	-0.413	0.008	-52.657
KOL10	42683	44951	87634	-0.026	0.006	-4.206
KOL9	78566	83419	161985	-0.030	0.006	-5.338
Kom1	28744	31843	60587	-0.051	0.007	-6.966
Kun1	4228	3732	7960	0.062	0.012	5.118
Kun2	52884	48022	100906	0.048	0.005	9.043
Kun3	52911	49168	102079	0.037	0.005	6.709
LARK01	48117	40952	89069	0.080	0.005	15.048
LARK02	11891	10684	22575	0.053	0.008	6.527
Lec1	72642	67086	139728	0.040	0.005	7.590
Lec4	81908	75572	157480	0.040	0.005	7.355
LISM01	282	255	537	0.050	0.046	1.096
LISM03	946	795	1741	0.087	0.024	3.571
Lud1	29985	28982	58967	0.017	0.006	2.733
Lud2	23350	22699	46049	0.014	0.006	2.221
Lud3	66910	64982	131892	0.015	0.006	2.458
Lud4	13712	13070	26782	0.024	0.008	3.144

LW017	491	471	962	0.021	0.032	0.640
LW018	439	424	863	0.017	0.036	0.483
LW019	12707	11848	24555	0.035	0.008	4.345
LW020	798	685	1483	0.076	0.026	2.982
LW021	28925	27185	56110	0.031	0.006	5.257
LW023	44748	42953	87701	0.020	0.006	3.602
LW024	37665	35181	72846	0.034	0.006	5.924
LW025	29333	26796	56129	0.045	0.007	6.709
LW026	543	446	989	0.098	0.032	3.028
LW027	461	416	877	0.051	0.032	1.592
LW028	1046	961	2007	0.042	0.023	1.814
Lym1	89897	82972	172869	0.040	0.005	7.287
Lym2	85469	77636	163105	0.048	0.005	8.962
Mam1	30544	32383	62927	-0.029	0.008	-3.858
Man1	9269	8253	17522	0.058	0.010	6.001
Man2	32541	30128	62669	0.039	0.006	6.034
Man3	57435	53263	110698	0.038	0.005	7.433
Man4	57166	52423	109589	0.043	0.005	8.027
Man5	36957	32217	69174	0.069	0.006	11.847
Man6	56491	53198	109689	0.030	0.006	5.323
Mar1	3160	3223	6383	-0.010	0.014	-0.720
Mas1	373	816	1189	-0.373	0.028	-13.152
MDR1	25451	23638	49089	0.037	0.007	5.650
Men1	75991	75872	151863	0.001	0.006	0.139
Men2	61456	62296	123752	-0.007	0.006	-1.224
Men3	19615	20073	39688	-0.012	0.007	-1.722
Mon1	59	68	127	-0.071	0.090	-0.791
Mon10	37342	41430	78772	-0.052	0.006	-8.290
Mon2	4795	4920	9715	-0.013	0.011	-1.201
Mon3	3723	3808	7531	-0.011	0.014	-0.834
Mon4	30265	33830	64095	-0.056	0.007	-7.781
Mon5	19423	19972	39395	-0.014	0.007	-2.041

Mon6	13314	14547	27861	-0.044	0.008	-5.588
Mon7	25859	26917	52776	-0.020	0.007	-3.015
Mon8	51751	53833	105584	-0.020	0.005	-3.592
Mon9	25536	26657	52193	-0.021	0.006	-3.335
Nad1	20178	20926	41104	-0.018	0.007	-2.444
Nad2	21193	21963	43156	-0.018	0.007	-2.529
Nah1	31	32	63	-0.016	0.124	-0.129
Ness5	83432	77841	161273	0.035	0.005	6.350
New1	90693	85544	176237	0.029	0.005	5.465
New2	86790	79220	166010	0.046	0.005	8.744
New3	89387	83177	172564	0.036	0.005	6.794
New4	86193	79437	165630	0.041	0.005	7.657
New5	88391	81708	170099	0.039	0.006	7.007
Nor1	55188	40623	95811	0.152	0.005	29.862
Otb1	82368	78723	161091	0.023	0.005	4.240
Par1	70064	65159	135223	0.036	0.005	6.649
Par2	74601	70156	144757	0.031	0.005	5.620
Plo1	9792	9275	19067	0.027	0.008	3.240
Plo2	29023	27980	57003	0.018	0.006	2.945
Plo3	75052	73860	148912	0.008	0.006	1.443
Plo4	76819	74895	151714	0.013	0.005	2.474
Plo5	6441	6186	12627	0.020	0.010	1.954
Plo6	28655	26658	55313	0.036	0.006	5.738
Plo7	8185	8088	16273	0.006	0.009	0.653
Plo8	14113	13614	27727	0.018	0.007	2.418
Poi1	88225	85972	174197	0.013	0.006	2.225
Poi2	63366	61194	124560	0.017	0.005	3.231
Poi3	47088	44498	91586	0.028	0.006	4.873
Poi4	72506	70093	142599	0.017	0.006	3.028
Poi5	28557	26484	55041	0.038	0.007	5.584
Pot1	79686	72758	152444	0.045	0.006	8.112
Pot2	79473	72931	152404	0.043	0.005	7.925

Pot3	69308	64451	133759	0.036	0.005	6.706
Pot4	81076	75415	156491	0.036	0.005	7.467
Pot5	72856	66922	139778	0.042	0.006	7.512
Pot6	25368	23196	48564	0.045	0.006	7.209
Qab1	3239	2993	6232	0.039	0.012	3.233
Qaz1	58749	60914	119663	-0.018	0.005	-3.329
Qaz2	19187	18577	37764	0.016	0.007	2.248
Red1	73571	67937	141508	0.040	0.006	6.777
Red2	87994	81742	169736	0.037	0.005	7.043
Rou1	83080	77765	160845	0.033	0.005	6.099
Rou2	82903	77293	160196	0.035	0.005	6.537
Rou3	63534	56761	120295	0.056	0.006	10.055
Rou4	54289	50603	104892	0.035	0.006	6.260
Ruc1	753	696	1449	0.039	0.027	1.468
Ruc2	31113	30297	61410	0.013	0.007	2.043
Rug1	94972	86809	181781	0.045	0.005	8.542
Sac3	14373	14622	28995	-0.009	0.008	-1.120
Sar38	30258	29000	59258	0.021	0.006	3.278
Sch1	67901	63898	131799	0.030	0.006	5.383
Sch2	56857	52439	109296	0.040	0.006	7.060
Sch3	3632	3375	7007	0.037	0.013	2.818
Sha1	26863	25600	52463	0.024	0.006	3.720
Sha2	30898	29481	60379	0.023	0.006	3.957
Shiderti1	27223	28496	55719	-0.023	0.007	-3.234
Shimao02	3394	2934	6328	0.073	0.014	5.261
Shimao04	11125	10364	21489	0.035	0.009	4.043
Shimao05	60837	58075	118912	0.023	0.006	4.102
Shimao07	4767	3815	8582	0.111	0.011	9.679
Shr1	197	217	414	-0.048	0.050	-0.967
Shr2	4640	5832	10472	-0.114	0.012	-9.388
Shu1	55680	57628	113308	-0.017	0.006	-3.044
Snu1	87278	81870	169148	0.032	0.005	5.891

Soh1	31430	31922	63352	-0.008	0.006	-1.224
Son1	1056	924	1980	0.067	0.023	2.849
Stb1	69687	65338	135025	0.032	0.005	6.146
Stb2	71454	66586	138040	0.035	0.005	6.703
Stb3	61153	57995	119148	0.027	0.006	4.590
Ste1	51979	47717	99696	0.043	0.005	7.953
STK230	1030	968	1998	0.031	0.023	1.323
STK231	457	380	837	0.092	0.035	2.618
STK234	7175	7064	14239	0.008	0.010	0.779
STK236	572	532	1104	0.036	0.030	1.207
STK239	7608	7385	14993	0.015	0.009	1.648
STK240	310	269	579	0.071	0.043	1.646
STK243	1999	805	2804	0.426	0.017	24.463
Sto1	27259	27094	54353	0.003	0.006	0.482
Sto2	1575	1604	3179	-0.009	0.019	-0.485
Sto3	6830	6481	13311	0.026	0.010	2.750
Stra1	61946	56518	118464	0.046	0.006	8.314
Stu1	11090	10680	21770	0.019	0.008	2.314
Sub1	37869	37918	75787	-0.001	0.004	-0.172
Sut1	70200	64394	134594	0.043	0.005	8.115
Sza1	65048	64182	129230	0.007	0.006	1.209
Sza10	31334	30749	62083	0.009	0.006	1.601
Sza11	2699	2618	5317	0.015	0.016	0.981
Sza12	590	541	1131	0.043	0.029	1.471
Sza13	41133	40117	81250	0.013	0.006	2.132
Sza15	619	663	1282	-0.034	0.029	-1.199
Sza2	67349	66063	133412	0.010	0.005	1.840
Sza3	76784	75641	152425	0.007	0.005	1.485
Sza4	69044	69611	138655	-0.004	0.005	-0.744
Sza5	70959	70143	141102	0.006	0.005	1.095
Sza6	33545	32522	66067	0.015	0.006	2.660
Sza7	33136	31378	64514	0.027	0.006	4.620

Sza8	54367	52631	106998	0.016	0.006	2.752
Sza9	28418	25676	54094	0.051	0.006	7.977
Tap1	45046	47387	92433	-0.025	0.006	-4.340
Tar1	913	905	1818	0.004	0.025	0.175
Tar2	936	841	1777	0.053	0.024	2.191
TAW1	4388	4079	8467	0.036	0.012	3.133
TAW2	3790	3706	7496	0.011	0.012	0.897
Tda1	8	18	26	-0.385	0.180	-2.141
Tef1	19383	20279	39662	-0.023	0.007	-3.256
Tes1	5227	11125	16352	-0.361	0.011	-34.168
Thr1	93	68	161	0.155	0.077	2.011
Til1	6922	15334	22256	-0.378	0.010	-37.296
Tmq1	38	75	113	-0.327	0.089	-3.669
Tmq2	44	79	123	-0.285	0.087	-3.278
Tmq3	2214	4689	6903	-0.359	0.013	-27.514
Tmq4	2598	4996	7594	-0.316	0.014	-23.083
Tqa1	18	10	28	0.286	0.184	1.549
Tqa2	32509	34056	66565	-0.023	0.006	-3.871
Tqa3	1210	1255	2465	-0.018	0.021	-0.888
Tsa1	62341	123913	186254	-0.331	0.008	-38.917
Tsa2	224	444	668	-0.329	0.037	-8.802
Tsa3	1480	1722	3202	-0.076	0.019	-3.893
TWAN01	504	438	942	0.070	0.033	2.138
TWAN02	3383	2974	6357	0.064	0.013	4.872
TWAN03	513	388	901	0.139	0.033	4.231
TWAN04	687	580	1267	0.084	0.028	2.992
TWAN05	620	536	1156	0.073	0.029	2.518
Tyq1	6567	14168	20735	-0.367	0.010	-36.542
Ulu1	12601	12693	25294	-0.004	0.008	-0.442
Ulu2	26330	26573	52903	-0.005	0.006	-0.709
Ulu3	30750	30633	61383	0.002	0.006	0.309
Ulu4	29209	28559	57768	0.011	0.006	1.793

Ulu5	6031	6073	12104	-0.003	0.010	-0.338
Ust1	1623	1501	3124	0.039	0.019	2.026
VEM329	29484	31241	60725	-0.029	0.006	-4.563
VEM330	22795	23775	46570	-0.021	0.007	-3.107
VEM331	19010	20175	39185	-0.030	0.007	-4.364
VEM337	1050	1338	2388	-0.121	0.021	-5.630
VEM338	366	390	756	-0.032	0.037	-0.851
VEM341	307	360	667	-0.079	0.039	-2.030
Viz1	43888	39020	82908	0.059	0.006	10.586
Viz2	93295	86995	180290	0.035	0.005	6.808
Wal1	54142	49273	103415	0.047	0.006	7.966
Win1	93184	88335	181519	0.027	0.005	5.305
Wro1	72305	65149	137454	0.052	0.006	9.332
Yer1	108378	266724	375102	-0.422	0.007	-59.611
Yor1	93232	84861	178093	0.047	0.006	8.520
Yor11	56194	50339	106533	0.055	0.005	10.054
Yor12	88622	81892	170514	0.039	0.005	7.290
Yor2	89505	82747	172252	0.039	0.005	7.482
Yor3	81526	75484	157010	0.038	0.005	7.194
Yor4	86548	79147	165695	0.045	0.005	8.437
Yor5	86679	78521	165200	0.049	0.005	9.314
Yor6	93606	86119	179725	0.042	0.005	7.768
Yor7	94081	85535	179616	0.048	0.005	9.294
Zag1	15216	15253	30469	-0.001	0.008	-0.161
Zah1	3528	7876	11404	-0.381	0.011	-35.781
Zha1	30394	30470	60864	-0.001	0.007	-0.178
Zha2	1869	1947	3816	-0.020	0.017	-1.217

Supplementary Table S8. Admixture results for K=6, plotted in Figures 32 and 33.

Sample	European Taurine	African/ Levantine Taurine	European Aurochs	Eastern aurochs	Indian/ Pakistani Indicines	Chinese Indicines
Ace1	0.42901	0.550774	0.020186	0.00001	0.00001	0.00001
Asp1	0.935131	0.00001	0.064829	0.00001	0.00001	0.00001
Baikal1	0.00001	0.00001	0.00001	0.99995	0.00001	0.00001
Bal1	0.957809	0.00001	0.04215	0.00001	0.00001	0.00001
Bal2	0.95234	0.00001	0.047618	0.000012	0.00001	0.00001
Bed3	0.000011	0.00001	0.999949	0.00001	0.00001	0.00001
Bed4	0.00001	0.00001	0.99995	0.00001	0.00001	0.00001
Bel1	0.75786	0.205803	0.036307	0.00001	0.00001	0.00001
Bel2	0.735598	0.215142	0.04923	0.00001	0.00001	0.00001
Belarus1	0.00001	0.00001	0.99995	0.00001	0.00001	0.00001
Bes1	0.313945	0.206788	0.00001	0.14403	0.335217	0.00001
Bes2	0.135518	0.136936	0.00001	0.001059	0.726467	0.00001
Bis1	0.916034	0.00001	0.083926	0.00001	0.00001	0.00001
Bis2	0.881023	0.00001	0.118937	0.00001	0.00001	0.00001
Bisk1	0.677106	0.229219	0.093645	0.00001	0.00001	0.00001
Bla1	0.835306	0.164652	0.00001	0.00001	0.000012	0.00001
Bla2	0.81265	0.18731	0.00001	0.00001	0.00001	0.00001
Bla3	0.921033	0.078927	0.00001	0.00001	0.00001	0.00001
Bol1	0.399598	0.600362	0.00001	0.00001	0.00001	0.00001
Bor1	0.946206	0.00001	0.053754	0.00001	0.00001	0.00001
Borly1	0.00001	0.00001	0.035695	0.964265	0.00001	0.00001
Borly2	0.00001	0.00001	0.16402	0.83594	0.00001	0.00001
Borly3	0.000011	0.00001	0.165982	0.833977	0.00001	0.00001
Bri1	0.840975	0.097644	0.061351	0.00001	0.00001	0.00001
Brn1	0.87096	0.085073	0.043937	0.00001	0.00001	0.00001
Brn2	0.904665	0.00001	0.095295	0.00001	0.00001	0.00001
Brn3	0.897593	0.022636	0.079741	0.00001	0.00001	0.00001
Brn4	0.891951	0.02737	0.08065	0.00001	0.00001	0.00001

Brn5	0.88304	0.024211	0.092719	0.00001	0.00001	0.00001
Bub1	0.797778	0.202182	0.00001	0.00001	0.00001	0.00001
Bub2	0.816218	0.183742	0.00001	0.00001	0.00001	0.00001
Bul1	0.28379	0.234287	0.00001	0.160685	0.321218	0.00001
Bul2	0.292373	0.306856	0.00001	0.122327	0.278423	0.00001
Bun1	0.874047	0.028192	0.097729	0.00001	0.00001	0.000012
Bun2	0.865022	0.045036	0.08991	0.00001	0.00001	0.000012
Cas1	0.15599	0.382516	0.00001	0.027029	0.434445	0.00001
Cer2	0.89192	0.013596	0.094454	0.00001	0.00001	0.00001
Cer3	0.606964	0.273197	0.091272	0.028547	0.00001	0.00001
Cer4	0.814321	0.090144	0.095504	0.00001	0.00001	0.00001
Ch22	0.469308	0.472816	0.057846	0.00001	0.00001	0.00001
Cha1	0.94146	0.00001	0.0585	0.00001	0.00001	0.00001
Che1	0.66892	0.33104	0.00001	0.00001	0.00001	0.00001
Cla1	0.891317	0.00001	0.108643	0.00001	0.00001	0.00001
Cla2	0.905035	0.00001	0.094925	0.00001	0.00001	0.00001
Cla3	0.931514	0.00001	0.068446	0.00001	0.00001	0.00001
Cla4	0.906288	0.00001	0.093672	0.00001	0.00001	0.00001
Cla5	0.876384	0.00001	0.123576	0.00001	0.00001	0.00001
Cla6	0.899576	0.00001	0.100384	0.00001	0.00001	0.00001
Cla7	0.907322	0.00001	0.092638	0.00001	0.00001	0.00001
Cla8	0.882742	0.00001	0.117218	0.00001	0.00001	0.00001
Cla9	0.879336	0.00001	0.120624	0.00001	0.00001	0.00001
Cri1	0.912347	0.035134	0.05249	0.00001	0.00001	0.00001
Cuk1	0.596139	0.403818	0.00001	0.00001	0.00001	0.000013
Cuk2	0.628435	0.371525	0.00001	0.00001	0.00001	0.00001
Da1	0.917955	0.00001	0.082003	0.00001	0.00001	0.000012
Da2	0.925406	0.00001	0.074554	0.00001	0.00001	0.00001
Da3	0.918701	0.00001	0.081259	0.00001	0.00001	0.00001
Da4	0.939096	0.00001	0.060864	0.00001	0.00001	0.00001
Da5	0.894523	0.00001	0.105437	0.00001	0.00001	0.00001
Da6	0.907103	0.00001	0.092857	0.00001	0.00001	0.00001

Dav1	0.4158	0.58416	0.00001	0.00001	0.00001	0.00001
Dav2	0.412644	0.587316	0.00001	0.00001	0.00001	0.00001
Den1	0.907194	0.010931	0.081845	0.00001	0.00001	0.00001
Den2	0.885253	0.009288	0.105429	0.00001	0.00001	0.00001
Dro1	0.863685	0.035323	0.100963	0.00001	0.00001	0.00001
Dub1	0.967888	0.00001	0.032063	0.00001	0.00001	0.000018
Dur1	0.827708	0.00001	0.172252	0.00001	0.00001	0.00001
Dur2	0.818155	0.00001	0.181805	0.00001	0.00001	0.00001
Dur3	0.830054	0.00001	0.169906	0.00001	0.00001	0.00001
Dyr1	0.808641	0.112953	0.059995	0.018392	0.00001	0.00001
Dzh1	0.789382	0.210578	0.00001	0.00001	0.00001	0.00001
Dzh2	0.693597	0.306363	0.00001	0.00001	0.00001	0.00001
Els1	0.85384	0.050962	0.095168	0.00001	0.00001	0.00001
Enk1	0.827011	0.051532	0.121422	0.00001	0.00001	0.000014
Enk2	0.839765	0.036919	0.123286	0.00001	0.00001	0.00001
Enk3	0.851731	0.036729	0.111507	0.00001	0.000013	0.00001
Enk4	0.843439	0.038385	0.118143	0.00001	0.00001	0.000013
Enk5	0.844367	0.042963	0.11264	0.00001	0.00001	0.00001
Far1	0.385853	0.614107	0.00001	0.00001	0.00001	0.00001
Fir1	0.866489	0.070906	0.062576	0.00001	0.00001	0.00001
Fis1	0.942138	0.00001	0.057822	0.00001	0.00001	0.00001
Fis2	0.932499	0.00001	0.067461	0.00001	0.00001	0.00001
For1	0.388104	0.423337	0.00001	0.06617	0.12237	0.00001
Galicia1	0.00001	0.00001	0.99995	0.00001	0.00001	0.00001
Galicia2	0.00001	0.00001	0.99995	0.00001	0.00001	0.00001
Galicia3	0.00001	0.00001	0.99995	0.00001	0.00001	0.00001
GQ1	0.820924	0.00001	0.179036	0.00001	0.00001	0.00001
Gub1	0.781633	0.166659	0.051678	0.00001	0.00001	0.00001
Gub2	0.792166	0.153603	0.054201	0.00001	0.00001	0.00001
Gub3	0.823568	0.143978	0.032424	0.00001	0.00001	0.00001
Gub4	0.809126	0.150301	0.040543	0.00001	0.00001	0.00001
Gub5	0.806897	0.147665	0.045408	0.00001	0.00001	0.00001

Gyu1	0.00001	0.00001	0.00001	0.99995	0.00001	0.00001
Gyu2	0.207467	0.454584	0.256256	0.081674	0.00001	0.00001
Had1	0.898886	0.030721	0.07036	0.00001	0.00001	0.000013
Has1	0.37987	0.595412	0.00001	0.024688	0.00001	0.00001
Has3	0.501837	0.498123	0.00001	0.00001	0.00001	0.00001
Has4	0.443225	0.451124	0.00001	0.070787	0.034844	0.00001
Has5	0.444909	0.464537	0.00001	0.07171	0.018824	0.00001
Has6	0.404576	0.595384	0.00001	0.00001	0.00001	0.00001
Has7	0.264587	0.274923	0.00001	0.055669	0.404801	0.00001
Has8	0.40041	0.59955	0.00001	0.00001	0.00001	0.00001
Has9	0.39629	0.540294	0.00001	0.063386	0.00001	0.00001
Hjo1	0.00001	0.00001	0.99995	0.00001	0.00001	0.00001
Hou1	0.852132	0.082869	0.064969	0.00001	0.00001	0.00001
Hou2	0.854023	0.074849	0.071097	0.00001	0.00001	0.00001
Hxh1	0.899623	0.100337	0.00001	0.00001	0.00001	0.00001
Hxh2	0.000013	0.00001	0.999947	0.00001	0.00001	0.00001
Irt1	0.813589	0.00001	0.186371	0.00001	0.00001	0.00001
Irt2	0.837893	0.00001	0.162067	0.00001	0.00001	0.00001
Irt3	0.804126	0.00001	0.195834	0.00001	0.00001	0.00001
Irt4	0.817912	0.00001	0.182048	0.00001	0.00001	0.00001
Irt5	0.819823	0.00001	0.180137	0.00001	0.00001	0.00001
Irt6	0.821666	0.00001	0.178294	0.00001	0.00001	0.00001
Irt7	0.827073	0.00001	0.172886	0.000011	0.00001	0.00001
Irt8	0.839167	0.00001	0.160793	0.00001	0.00001	0.00001
Kas1	0.701025	0.101757	0.186331	0.010867	0.00001	0.00001
Kaz1	0.438884	0.352114	0.00001	0.114889	0.094093	0.00001
Kaz2	0.399708	0.353235	0.00001	0.084234	0.162803	0.00001
Kaz3	0.44615	0.33262	0.00001	0.115403	0.105808	0.00001
Kaz4	0.46655	0.319889	0.00001	0.072108	0.141433	0.00001
Kaz5	0.372707	0.301936	0.00001	0.083146	0.242191	0.00001
Kho1	0.174961	0.119853	0.00001	0.015758	0.689408	0.00001
Kie1	0.861064	0.086536	0.052369	0.00001	0.00001	0.00001

Kie2	0.609199	0.073754	0.317018	0.00001	0.00001	0.00001
Kil1	0.911569	0.00001	0.088391	0.00001	0.00001	0.00001
Kir4	0.617709	0.262706	0.096563	0.023002	0.00001	0.00001
Kir5	0.611025	0.297683	0.091262	0.00001	0.00001	0.00001
Kir6	0.625933	0.281515	0.027919	0.064612	0.00001	0.00001
KK1	0.223505	0.453297	0.00001	0.011273	0.311905	0.00001
KO201411	0.46086	0.243228	0.018308	0.277584	0.00001	0.00001
KO201415	0.469393	0.22668	0.014705	0.289203	0.00001	0.00001
Kok1	0.249455	0.097629	0.00001	0.072607	0.580289	0.00001
KOL9	0.434283	0.249462	0.031639	0.284597	0.00001	0.00001
KOL10	0.449233	0.220338	0.025762	0.304647	0.00001	0.00001
Kom1	0.47754	0.266033	0.00001	0.256397	0.00001	0.00001
Kun1	0.90585	0.076707	0.017413	0.00001	0.00001	0.00001
Kun2	0.89076	0.1092	0.00001	0.00001	0.00001	0.00001
Kun3	0.855915	0.106811	0.037244	0.00001	0.00001	0.00001
LARK01	0.901592	0.00001	0.098368	0.00001	0.00001	0.00001
LARK02	0.915443	0.00001	0.084517	0.00001	0.00001	0.00001
Lec1	0.83611	0.00001	0.16385	0.00001	0.00001	0.00001
Lec4	0.828744	0.00001	0.171216	0.00001	0.00001	0.00001
Lud1	0.852377	0.147583	0.00001	0.00001	0.00001	0.00001
Lud2	0.87206	0.1279	0.00001	0.00001	0.00001	0.00001
Lud3	0.878677	0.121283	0.00001	0.00001	0.00001	0.00001
Lud4	0.87858	0.12138	0.00001	0.00001	0.00001	0.00001
LW019	0.63094	0.15594	0.21309	0.00001	0.00001	0.00001
LW021	0.607177	0.155074	0.237719	0.00001	0.00001	0.00001
LW023	0.63136	0.159789	0.208821	0.00001	0.00001	0.00001
LW024	0.59164	0.119104	0.289225	0.00001	0.00001	0.00001
LW025	0.6222	0.135946	0.241824	0.00001	0.00001	0.00001
Lym1	0.92603	0.00001	0.073928	0.00001	0.000012	0.00001
Lym2	0.939188	0.00001	0.060772	0.00001	0.00001	0.00001
Mam1	0.467847	0.312837	0.00001	0.219286	0.00001	0.00001
Man1	0.871265	0.03265	0.096055	0.00001	0.00001	0.00001

Man2	0.851619	0.06027	0.088081	0.00001	0.00001	0.00001
Man3	0.845049	0.054708	0.100214	0.00001	0.00001	0.00001
Man4	0.840188	0.063817	0.095961	0.00001	0.00001	0.000014
Man5	0.844846	0.060471	0.094653	0.00001	0.00001	0.00001
Man6	0.859857	0.049111	0.091002	0.00001	0.00001	0.00001
Mantova1	0.00001	0.00001	0.99995	0.00001	0.00001	0.00001
Mar1	0.517525	0.482435	0.00001	0.00001	0.00001	0.00001
MDR1	0.878952	0.00001	0.121008	0.00001	0.00001	0.00001
Men1	0.539807	0.460153	0.00001	0.00001	0.00001	0.00001
Men2	0.533132	0.466828	0.00001	0.00001	0.00001	0.00001
Men3	0.503191	0.496769	0.00001	0.00001	0.00001	0.00001
Mon2	0.426892	0.573068	0.00001	0.00001	0.00001	0.00001
Mon3	0.373562	0.626398	0.00001	0.00001	0.00001	0.00001
Mon4	0.292482	0.653539	0.00001	0.053949	0.00001	0.00001
Mon5	0.39952	0.60044	0.00001	0.00001	0.00001	0.00001
Mon6	0.353829	0.646131	0.00001	0.00001	0.00001	0.00001
Mon7	0.38821	0.584597	0.003774	0.023399	0.00001	0.00001
Mon8	0.416161	0.583799	0.00001	0.00001	0.00001	0.00001
Mon9	0.372007	0.627953	0.00001	0.00001	0.00001	0.00001
Mon10	0.354735	0.63723	0.00001	0.008005	0.00001	0.00001
Nad1	0.39645	0.60351	0.00001	0.00001	0.00001	0.00001
Nad2	0.397636	0.602324	0.00001	0.00001	0.00001	0.00001
Ness5	0.847786	0.00001	0.152174	0.00001	0.00001	0.00001
New1	0.998616	0.00001	0.001344	0.00001	0.00001	0.00001
New2	0.940543	0.00001	0.059417	0.00001	0.00001	0.00001
New3	0.99995	0.00001	0.00001	0.00001	0.00001	0.00001
New4	0.997238	0.00001	0.002722	0.00001	0.00001	0.00001
New5	0.994936	0.00001	0.005024	0.00001	0.00001	0.00001
NVL3	0.00001	0.00001	0.185163	0.814797	0.00001	0.00001
Otb1	0.885461	0.048445	0.066064	0.00001	0.00001	0.00001
Padova1	0.00001	0.00001	0.99995	0.00001	0.00001	0.00001
Palidoro1	0.00001	0.077844	0.922116	0.00001	0.00001	0.00001

Par1	0.99995	0.00001	0.00001	0.00001	0.00001	0.00001
Par2	0.99995	0.00001	0.00001	0.00001	0.00001	0.00001
Par3	0.99995	0.00001	0.00001	0.00001	0.00001	0.00001
Plo1	0.744541	0.176381	0.079049	0.00001	0.00001	0.00001
Plo2	0.784013	0.215947	0.00001	0.00001	0.00001	0.00001
Plo3	0.808707	0.187023	0.004238	0.00001	0.00001	0.000012
Plo4	0.785718	0.213402	0.00085	0.00001	0.00001	0.00001
Plo5	0.825183	0.174777	0.00001	0.00001	0.00001	0.00001
Plo6	0.82895	0.17101	0.00001	0.00001	0.00001	0.00001
Plo7	0.799688	0.200272	0.00001	0.00001	0.00001	0.00001
Plo8	0.8011	0.19886	0.00001	0.00001	0.00001	0.00001
Poi1	0.875611	0.082875	0.041484	0.00001	0.00001	0.00001
Poi2	0.812558	0.140272	0.047141	0.00001	0.00001	0.00001
Poi3	0.908645	0.057869	0.033456	0.00001	0.00001	0.00001
Poi4	0.845794	0.128827	0.02535	0.00001	0.00001	0.00001
Poi5	0.864996	0.082956	0.052018	0.00001	0.00001	0.00001
Pot1	0.904935	0.00001	0.095025	0.00001	0.00001	0.00001
Pot2	0.905973	0.00001	0.093987	0.00001	0.00001	0.00001
Pot3	0.901342	0.00001	0.098618	0.00001	0.00001	0.00001
Pot4	0.904468	0.00001	0.095492	0.00001	0.00001	0.00001
Pot5	0.908132	0.00001	0.091828	0.00001	0.00001	0.00001
Pot6	0.893112	0.00001	0.106848	0.00001	0.00001	0.00001
Qab1	0.462906	0.537054	0.00001	0.00001	0.00001	0.00001
Qaz1	0.409976	0.589984	0.00001	0.00001	0.00001	0.00001
Qaz2	0.357229	0.642731	0.00001	0.00001	0.00001	0.00001
Red1	0.945354	0.00001	0.054606	0.00001	0.00001	0.00001
Red2	0.937138	0.00001	0.062822	0.00001	0.00001	0.00001
Rhi1	0.00001	0.00001	0.99995	0.00001	0.00001	0.00001
Rhi2	0.00001	0.00001	0.99995	0.00001	0.00001	0.00001
Rhi3	0.000011	0.00001	0.999949	0.00001	0.00001	0.00001
ROS002	0.00001	0.00001	0.146733	0.853227	0.00001	0.00001
Rou1	0.944011	0.00001	0.055949	0.00001	0.00001	0.00001

Rou2	0.976289	0.00001	0.023671	0.00001	0.00001	0.00001
Rou3	0.996007	0.00001	0.003953	0.00001	0.00001	0.00001
Rou4	0.963213	0.00001	0.036747	0.00001	0.00001	0.00001
Ruc2	0.615945	0.260969	0.084021	0.039045	0.00001	0.00001
Rug1	0.942914	0.00001	0.057046	0.00001	0.00001	0.00001
SAC3	0.432327	0.567633	0.00001	0.00001	0.00001	0.00001
Sar38	0.836697	0.163263	0.00001	0.00001	0.00001	0.00001
Sch1	0.91955	0.00001	0.08041	0.00001	0.00001	0.00001
Sch2	0.923775	0.00001	0.076185	0.00001	0.00001	0.00001
Sch3	0.956043	0.00001	0.043917	0.00001	0.00001	0.00001
Sha1	0.550024	0.313581	0.136364	0.00001	0.00001	0.00001
Sha2	0.52394	0.28792	0.18811	0.00001	0.00001	0.00001
Shiderti1	0.417	0.259396	0.025173	0.298411	0.00001	0.00001
Shimao02	0.520319	0.00001	0.00001	0.479641	0.00001	0.00001
Shimao04	0.492723	0.000908	0.00001	0.506338	0.00001	0.00001
Shimao07	0.501419	0.00001	0.00001	0.498541	0.00001	0.00001
Shimao05	0.463414	0.052535	0.00001	0.484022	0.00001	0.00001
Shr2	0.544237	0.170325	0.00001	0.230495	0.054923	0.00001
Shu1	0.419958	0.580002	0.00001	0.00001	0.00001	0.00001
Ska1	0.00001	0.00001	0.99995	0.00001	0.00001	0.00001
Ska3	0.00001	0.00001	0.99995	0.00001	0.00001	0.00001
Snu1	0.916854	0.00001	0.083106	0.00001	0.00001	0.00001
Soh1	0.389972	0.609988	0.00001	0.00001	0.00001	0.00001
Stb1	0.871447	0.018541	0.109982	0.00001	0.00001	0.00001
Stb2	0.842658	0.043266	0.114046	0.00001	0.00001	0.00001
Stb3	0.837756	0.037662	0.124551	0.00001	0.00001	0.00001
Ste1	0.856795	0.039094	0.104073	0.00001	0.00001	0.000019
STK234	0.122608	0.877352	0.00001	0.00001	0.00001	0.00001
STK239	0.165016	0.834944	0.00001	0.00001	0.00001	0.00001
Sto1	0.445281	0.554679	0.00001	0.00001	0.00001	0.00001
Sto2	0.419094	0.580866	0.00001	0.00001	0.00001	0.00001
Sto3	0.449032	0.550928	0.00001	0.00001	0.00001	0.00001

Stra1	0.906004	0.00001	0.093947	0.000019	0.00001	0.00001
Stu1	0.813084	0.186876	0.00001	0.00001	0.00001	0.00001
Sub1	0.55972	0.44024	0.00001	0.00001	0.00001	0.00001
Sut1	0.924783	0.00001	0.075177	0.00001	0.00001	0.00001
Sza1	0.650968	0.264547	0.082382	0.002082	0.00001	0.00001
Sza2	0.647644	0.253673	0.093326	0.005337	0.00001	0.00001
Sza3	0.637383	0.260944	0.078248	0.023405	0.00001	0.00001
Sza4	0.656943	0.246946	0.075404	0.020687	0.00001	0.00001
Sza5	0.658967	0.259823	0.054657	0.026533	0.00001	0.00001
Sza6	0.6606	0.255997	0.065685	0.017699	0.00001	0.00001
Sza7	0.661867	0.267782	0.062987	0.007343	0.00001	0.00001
Sza8	0.654384	0.239654	0.097037	0.008905	0.00001	0.00001
Sza9	0.661616	0.24516	0.073126	0.020078	0.00001	0.00001
Sza10	0.641063	0.257953	0.090591	0.010373	0.00001	0.00001
Sza11	0.613679	0.263911	0.122379	0.00001	0.00001	0.00001
Sza13	0.649641	0.233742	0.115381	0.001216	0.00001	0.00001
Tap1	0.42199	0.57797	0.00001	0.00001	0.00001	0.00001
TAW1	0.929377	0.021207	0.049386	0.00001	0.00001	0.00001
TAW2	0.883545	0.00244	0.113984	0.00001	0.00001	0.00001
Tef1	0.152783	0.847177	0.00001	0.00001	0.00001	0.00001
Tes1	0.186514	0.370456	0.00001	0.001159	0.44185	0.00001
Th7	0.018778	0.639387	0.269846	0.071969	0.00001	0.00001
Til1	0.273497	0.260709	0.00001	0.032103	0.433671	0.00001
Tmq3	0.210391	0.377833	0.00001	0.00001	0.411746	0.00001
Tmq4	0.20023	0.458703	0.00001	0.00702	0.334027	0.00001
Tqa2	0.05243	0.94753	0.00001	0.00001	0.00001	0.00001
Tri1	0.000015	0.00001	0.999945	0.00001	0.00001	0.00001
Tsa1	0.219376	0.470286	0.00001	0.00001	0.310308	0.00001
Tsa3	0.00001	0.99995	0.00001	0.00001	0.00001	0.00001
Tula1	0.00001	0.00001	0.00001	0.99995	0.00001	0.00001
TWAN02	0.85617	0.078574	0.065226	0.00001	0.00001	0.00001
Tyq1	0.144877	0.434595	0.00001	0.00001	0.420497	0.00001

Ulu1	0.55301	0.44695	0.00001	0.00001	0.00001	0.00001
Ulu2	0.577183	0.422777	0.00001	0.00001	0.00001	0.00001
Ulu3	0.621418	0.378542	0.00001	0.00001	0.00001	0.00001
Ulu4	0.533965	0.465995	0.00001	0.00001	0.00001	0.00001
Ulu5	0.59551	0.40445	0.00001	0.00001	0.00001	0.00001
Ust1	0.504966	0.404708	0.082306	0.008	0.00001	0.00001
Uzzo1	0.00001	0.00001	0.99995	0.00001	0.00001	0.00001
Var1	0.00001	0.100051	0.322793	0.577125	0.00001	0.00001
Var2	0.00001	0.113685	0.35727	0.529014	0.00001	0.00001
VEM329	0.302333	0.697627	0.00001	0.00001	0.00001	0.00001
VEM330	0.343746	0.656214	0.00001	0.00001	0.00001	0.00001
VEM331	0.295159	0.704801	0.00001	0.00001	0.00001	0.00001
Viz1	0.843833	0.039896	0.116241	0.00001	0.00001	0.00001
Viz2	0.854043	0.06554	0.080388	0.00001	0.00001	0.00001
Vratsa1	0.183841	0.090095	0.726033	0.00001	0.00001	0.00001
Vratsa2	0.165657	0.085711	0.748601	0.00001	0.00001	0.00001
Wal1	0.937543	0.00001	0.062417	0.00001	0.00001	0.00001
Win1	0.862411	0.051668	0.08589	0.00001	0.00001	0.000011
Wro1	0.904676	0.00001	0.095284	0.00001	0.00001	0.00001
Yer1	0.214944	0.073298	0.00001	0.094198	0.617541	0.00001
YoA	0.00001	0.00001	0.99995	0.00001	0.00001	0.00001
Yor1	0.901844	0.00001	0.098116	0.00001	0.00001	0.00001
Yor2	0.93752	0.00001	0.06244	0.00001	0.00001	0.00001
Yor3	0.927489	0.00001	0.072471	0.00001	0.00001	0.00001
Yor4	0.938568	0.00001	0.061392	0.00001	0.00001	0.00001
Yor5	0.932938	0.00001	0.067022	0.00001	0.00001	0.00001
Yor6	0.944674	0.00001	0.055286	0.00001	0.00001	0.00001
Yor7	0.92571	0.00001	0.07425	0.00001	0.00001	0.00001
Yor11	0.913764	0.00001	0.086196	0.00001	0.00001	0.000011
Yor12	0.926314	0.00001	0.073646	0.00001	0.00001	0.00001
Zag1	0.443231	0.556729	0.00001	0.00001	0.00001	0.00001
Zah1	0.228696	0.303284	0.00001	0.057693	0.410307	0.00001

Zea1	0.00001	0.00001	0.99995	0.00001	0.00001	0.00001
Zea2	0.00001	0.00001	0.99995	0.00001	0.00001	0.00001
Zha1	0.486076	0.412324	0.012799	0.088782	0.00001	0.00001
Zha2	0.648247	0.266094	0.008231	0.077408	0.00001	0.00001

Supplementary Table S9. Whole genome D-statistics results for indicine introgression in samples from the Middle East. The Sahiwal individual Sha3b was used to test D(test; Sub1; Sha3b). A positive D-statistics indicates introgression from indicine cattle into the tested individual. Samples that have nABBA+nBABA > 500 were then plotted in Figure 36 and 37 (with the exception of Aak1 which was kept in Figure 37). Aurochs are represented with an * at the end of their name.

Sample	nABBA	nBABA	Total number of sites	D	SE	Z score
Aak1	161	99	260	-0.238	0.061	3.883
Cuk1	6085	6115	12200	0.002	0.01	-0.238
Cuk2	24586	24467	49053	-0.002	0.006	0.391
Men1	75991	75872	151863	-0.001	0.006	0.139
Men2	61456	62296	123752	0.007	0.006	-1.224
Men3	19615	20073	39688	0.012	0.007	-1.722
Sub1	37869	37918	75787	0.001	0.004	-0.172
Ulu1	12601	12693	25294	0.004	0.008	-0.442
Ulu2	26330	26573	52903	0.005	0.006	-0.709
Ulu3	30750	30633	61383	-0.002	0.006	0.309
Ulu4	29209	28559	57768	-0.011	0.006	1.793
Ulu5	6031	6073	12104	0.003	0.01	-0.338
Ch22	22029	21414	43443	-0.014	0.007	1.990
Ace1	21263	21493	42756	0.005	0.007	-0.820
Til1	6922	15334	22256	0.378	0.01	-37.296
Shu1	55680	57628	113308	0.017	0.006	-3.044
Dam1	646	639	1285	-0.005	0.03	0.182
Ipa1	596	566	1162	-0.026	0.03	0.847
Shr2	4640	5832	10472	0.114	0.012	-9.388
Sto1	27259	27094	54353	-0.003	0.006	0.482
Sto2	1575	1604	3179	0.009	0.019	-0.485
Sto3	6830	6481	13311	-0.026	0.01	2.750
Ust1	1623	1501	3124	-0.039	0.019	2.026

Kom1	28744	31843	60587	0.051	0.007	-6.966
Kaz1	91215	130565	221780	0.177	0.008	-21.580
Kaz2	25184	41671	66855	0.247	0.009	-27.841
Kaz3	34398	49816	84214	0.183	0.009	-20.252
Kaz4	36491	56538	93029	0.215	0.009	-24.722
Kaz5	2514	4414	6928	0.274	0.014	-19.000
Bol1	37638	39088	76726	0.019	0.006	-3.237
Far1	8120	8536	16656	0.025	0.008	-2.944
Soh1	31430	31922	63352	0.008	0.006	-1.224
Dav1	47056	49640	96696	0.027	0.006	-4.626
Dav2	62573	64587	127160	0.016	0.006	-2.795
Mar1	3160	3223	6383	0.01	0.014	-0.720
Qab1	3239	2993	6232	-0.039	0.012	3.233
Qaz2	19187	18577	37764	-0.016	0.007	2.248
Sac3	14373	14622	28995	0.009	0.008	-1.120
Zag1	15216	15253	30469	0.001	0.008	-0.161
Has1	2760	3305	6065	0.09	0.014	-6.440
Has3	56812	57907	114719	0.01	0.005	-1.774
Has7	46504	97748	144252	0.355	0.008	-41.972
Nad1	20178	20926	41104	0.018	0.007	-2.444
Nad2	21193	21963	43156	0.018	0.007	-2.529
Qaz1	58749	60914	119663	0.018	0.005	-3.329
Tap1	45046	47387	92433	0.025	0.006	-4.340
For1	32480	49136	81616	0.204	0.009	-21.627
Has4	59325	71940	131265	0.096	0.008	-11.568
Has6	3881	4640	8521	0.089	0.013	-7.020
Has8	36100	39378	75478	0.043	0.008	-5.545
Has9	27458	29949	57407	0.043	0.007	-5.887
Has5	2963	3408	6371	0.07	0.014	-4.909
Kho1	26076	65388	91464	0.43	0.008	-55.904
Mon10	37342	41430	78772	0.052	0.006	-8.290
Mon2	4795	4920	9715	0.013	0.011	-1.201
Mon3	3723	3808	7531	0.011	0.014	-0.834
Mon4	30265	33830	64095	0.056	0.007	-7.781
Mon5	19423	19972	39395	0.014	0.007	-2.041
Mon6	13314	14547	27861	0.044	0.008	-5.588

Mon7	25859	26917	52776	0.02	0.007	-3.015
Mon8	51751	53833	105584	0.02	0.005	-3.592
Mon9	25536	26657	52193	0.021	0.006	-3.335
VEM341	307	360	667	0.079	0.039	-2.030
VEM331	19010	20175	39185	0.03	0.007	-4.364
VEM338	366	390	756	0.032	0.037	-0.851
Tef1	19383	20279	39662	0.023	0.007	-3.256
VEM329	29484	31241	60725	0.029	0.006	-4.563
VEM330	22795	23775	46570	0.021	0.007	-3.107
Bes1	98673	196431	295104	0.331	0.008	-40.735
Bes2	152054	394449	546503	0.444	0.007	-60.953
KK1	2421	4543	6964	0.305	0.013	-23.016
Tes1	5227	11125	16352	0.361	0.011	-34.168
Tqa2	32509	34056	66565	0.023	0.006	-3.871
Tqa3	1210	1255	2465	0.018	0.021	-0.888
Ash4	1324	2597	3921	0.325	0.017	-19.374
Mas1	373	816	1189	0.373	0.028	-13.152
Tmq3	2214	4689	6903	0.359	0.013	-27.514
Tmq4	2598	4996	7594	0.316	0.014	-23.083
Tsa1	62341	123913	186254	0.331	0.008	-38.917
Tsa2	224	444	668	0.329	0.037	-8.802
Tsa3	1480	1722	3202	0.076	0.019	-3.893
Cas1	7708	17515	25223	0.389	0.01	-39.872
Tyq1	6567	14168	20735	0.367	0.01	-36.542
Zah1	3528	7876	11404	0.381	0.011	-35.781

8. Bibliography

Achilli, A., Bonfiglio, S., Olivieri, A., Malusà, A., Pala, M., Kashani, B. H., Perego, U. A., Ajmone-Marsan, P., Liotta, L., Semino, O., Bandelt, H.-J., Ferretti, L., & Torroni, A. (2009). The Multifaceted Origin of Taurine Cattle Reflected by the Mitochondrial Genome. *PLOS ONE*, *4*(6), e5753. https://doi.org/10.1371/journal.pone.0005753

Achilli, A., Olivieri, A., Pellecchia, M., Uboldi, C., Colli, L., Al-Zahery, N., Accetturo, M., Pala, M., Kashani, B. H., Perego, U. A., Battaglia, V., Fornarino, S., Kalamati, J., Houshmand, M., Negrini, R., Semino, O., Richards, M., Macaulay, V., Ferretti, L., … Torroni, A. (2008). Mitochondrial genomes of extinct aurochs survive in domestic cattle. *Current Biology*, *18*(4), R157–R158. https://doi.org/10.1016/j.cub.2008.01.019

Alders, R. G., Campbell, A., Costa, R., Guèye, E. F., Ahasanul Hoque, M., Perezgrovas-Garza, R., Rota, A., & Wingett, K. (2021). Livestock across the world: Diverse animal species with complex roles in human societies and ecosystem services. *Animal Frontiers*, *11*(5), 20–29. https://doi.org/10.1093/af/vfab047

Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, *19*(9), 1655–1664. https://doi.org/10.1101/gr.094052.109

Allentoft, M. E., Collins, M., Harker, D., Haile, J., Oskam, C. L., Hale, M. L., Campos, P. F., Samaniego, J. A., Gilbert, M. T. P., Willerslev, E., Zhang, G., Scofield, R. P., Holdaway, R. N., & Bunce, M. (2012). The half-life of DNA in bone: Measuring decay kinetics in 158 dated fossils. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1748), 4724–4733. https://doi.org/10.1098/rspb.2012.1745

Anderson, S., de Bruijn, M. H. L., Coulson, A. R., Eperon, I. C., Sanger, F., & Young, I. G. (1982). Complete sequence of bovine mitochondrial DNA conserved features of the mammalian mitochondrial genome. *Journal of Molecular Biology*, *156*(4), 683–717. https://doi.org/10.1016/0022-2836(82)90137-1

Andrews, S. (2010). *FastQC A Quality Control tool for High Throughput Sequence Data*. Babraham Bioinformatics. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Arbuckle, B. S., Price, M. D., Hongo, H., & Öksüz, B. (2016). Documenting the initial appearance of domestic cattle in the Eastern Fertile Crescent (northern Iraq and western Iran). *Journal of Archaeological Science*, *72*, 1–9. https://doi.org/10.1016/j.jas.2016.05.008

Arnold, E., Greenfield, H., Hartman, G., Greenfield, T., Shai, I., Carter-McGee, P., & Maeir, A. (2018). *Provisioning the Early Bronze Age City of Tell es-Safi/Gath, Israel: Isotopic Analyses of Domestic Livestock Management Patterns | Open Quaternary*. https://doi.org/10.5334/oq.35

Bailey, J. F., Richards, M. B., Macaulay, V. A., Colson, I. B., James I, T., Bradley, D. G., Hedges, R. E. M., & Sykes, B. C. (1997). Ancient DNA suggests a recent expansion of European cattle from

a diverse wild progenitor species. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *263*(1376), 1467–1473. https://doi.org/10.1098/rspb.1996.0214

Bar-Yosef, O. (1998). On the Nature of Transitions: The Middle to Upper Palaeolithic and the Neolithic Revolution. *Cambridge Archaeological Journal*, *8*(2), 141–163. https://doi.org/10.1017/S0959774300001815

Bar-Yosef, O., & Belfer-Cohen, A. (1989). The origins of sedentism and farming communities in the Levant. *Journal of World Prehistory*, *3*(4), 447–498. https://doi.org/10.1007/BF00975111

Berthon, R. J.-P., & Mashkour, M. (2008). Animal Remains from Tilbeşar Excavations, Southeast Anatolia, Turkey. *Anatolia Antiqua. Eski Anadolu*, *16*(1), 23–51. https://doi.org/10.3406/anata.2008.1248

Besser, J., Carleton, H. A., Gerner-Smidt, P., Lindsey, R. L., & Trees, E. (2018). Next-Generation Sequencing Technologies and their Application to the Study and Control of Bacterial Infections. *Clinical Microbiology and Infection : The Official Publication of the European Society of Clinical Microbiology and Infectious Diseases*, *24*(4), 335–341. https://doi.org/10.1016/j.cmi.2017.10.013

Boessenkool, S., Hanghøj, K., Nistelberger, H. M., Der Sarkissian, C., Gondek, A. T., Orlando, L., Barrett, J. H., & Star, B. (2017). Combining bleach and mild predigestion improves ancient DNA recovery from bones. *Molecular Ecology Resources*, *17*(4), 742–751. https://doi.org/10.1111/1755-0998.12623

Bollongino, R., Burger, J., Powell, A., Mashkour, M., Vigne, J.-D., & Thomas, M. G. (2012). Modern Taurine Cattle Descended from Small Number of Near-Eastern Founders. *Molecular Biology and Evolution*, *29*(9), 2101–2104. https://doi.org/10.1093/molbev/mss092

Bollongino, R., Edwards, C. J., Alt, K. W., Burger, J., & Bradley, D. G. (2006). Early history of European domestic cattle as revealed by ancient DNA. *Biology Letters*, *2*(1), 155–159. https://doi.org/10.1098/rsbl.2005.0404

Bollongino, R., Tresset, A., & Vigne, J.-D. (2008a). Environment and excavation: Pre-lab impacts on ancient DNA analyses. *Comptes Rendus Palevol*, *7*(2), 91–98. https://doi.org/10.1016/j.crpv.2008.02.002

Bollongino, R., Elsner, J., Vigne, J.-D., & Burger, J. (2008b). Y-SNPs Do Not Indicate Hybridisation between European Aurochs and Domestic Cattle. *PLOS ONE*, *3*(10), e3418. https://doi.org/10.1371/journal.pone.0003418

Bonfiglio, S., Achilli, A., Olivieri, A., Negrini, R., Colli, L., Liotta, L., Ajmone-Marsan, P., Torroni, A., & Ferretti, L. (2010). The Enigmatic Origin of Bovine mtDNA Haplogroup R: Sporadic Interbreeding or an Independent Event of Bos primigenius Domestication in Italy? *PLOS ONE*, *5*(12), e15760. https://doi.org/10.1371/journal.pone.0015760

Bonfiglio, S., Ginja, C., Gaetano, A. D., Achilli, A., Olivieri, A., Colli, L., Tesfaye, K., Agha, S. H., Gama, L. T., Cattonaro, F., Penedo, M. C. T., Ajmone-Marsan, P., Torroni, A., & Ferretti, L. (2012). Origin and Spread of Bos taurus: New Clues from Mitochondrial Genomes Belonging to Haplogroup T1. *PLOS ONE*, *7*(6), e38601. https://doi.org/10.1371/journal.pone.0038601

Bouckaert, R., Vaughan, T. G., Barido-Sottani, J., Duchêne, S., Fourment, M., Gavryushkina, A., Heled, J., Jones, G., Kühnert, D., Maio, N. D., Matschiner, M., Mendes, F. K., Müller, N. F., Ogilvie, H. A., Plessis, L. du, Popinga, A., Rambaut, A., Rasmussen, D., Siveroni, I., … Drummond, A. J. (2019). BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *PLOS Computational Biology*, *15*(4), e1006650. https://doi.org/10.1371/journal.pcbi.1006650

Bradley, D. G., Loftus, R. T., Cunningham, P., & MacHugh, D. E. (1998). Genetics and domestic cattle origins. *Evolutionary Anthropology: Issues, News, and Reviews*, *6*(3), 79–86. https://doi.org/10.1002/(SICI)1520-6505(1998)6:3%253C79::AID-EVAN2%253E3.0.CO;2-R

Bradley, D. G., MacHugh, D. E., Cunningham, P., & Loftus, R. T. (1996). Mitochondrial diversity and the origins of African and European cattle. *Proceedings of the National Academy of Sciences*, *93*(10), 5131–5135. https://doi.org/10.1073/pnas.93.10.5131

Broad Institute. (2019). Picard toolkit. *Broad Institute, GitHub Repository*.

Cai, D., Sun, Y., Tang, Z., Hu, S., Li, W., Zhao, X., Xiang, H., & Zhou, H. (2014). The origins of Chinese domestic cattle as revealed by ancient DNA analysis. *Journal of Archaeological Science*, *41*, 423–434. https://doi.org/10.1016/j.jas.2013.09.003

Çakırlar, C., & Ikram, S. (2016). 'When elephants battle, the grass suffers.' Power, ivory and the Syrian elephant. *Levant*, *48*(2), 167–183. https://doi.org/10.1080/00758914.2016.1198068

Cavalli-Sforza, L. L., & Edwards, A. W. F. (1967). Phylogenetic analysis. Models and estimation procedures. *American Journal of Human Genetics*, *19*(3 Pt 1), 233–257.

Chen, N., Cai, Y., Chen, Q., Li, R., Wang, K., Huang, Y., Hu, S., Huang, S., Zhang, H., Zheng, Z., Song, W., Ma, Z., Ma, Y., Dang, R., Zhang, Z., Xu, L., Jia, Y., Liu, S., Yue, X., … Lei, C. (2018). Whole-genome resequencing reveals world-wide ancestry and adaptive introgression events of domesticated cattle in East Asia. *Nature Communications*, *9*(1), 2337. https://doi.org/10.1038/s41467-018-04737-0

Chen, S., Lin, B.-Z., Baig, M., Mitra, B., Lopes, R. J., Santos, A. M., Magee, D. A., Azevedo, M., Tarroso, P., Sasazaki, S., Ostrowski, S., Mahgoub, O., Chaudhuri, T. K., Zhang, Y., Costa, V., Royo, L. J., Goyache, F., Luikart, G., Boivin, N., … Beja-Pereira, A. (2010). Zebu Cattle Are an Exclusive Legacy of the South Asia Neolithic. *Molecular Biology and Evolution*, *27*(1), 1–6. https://doi.org/10.1093/molbev/msp213

Colominas, L., Edwards, C. J., Beja-Pereira, A., Vigne, J.-D., Silva, R. M., Castanyer, P., Tremoleda, J., Saña Seguí, M., Pérez-Ripoll, M., Goyache, F., Howe, C. J., Barker, G., & Bower, M. A. (2015). Detecting the T1 cattle haplogroup in the Iberian Peninsula from Neolithic to medieval times:

New clues to continuous cattle migration through time. *Journal of Archaeological Science*, *59*, 110–117. https://doi.org/10.1016/j.jas.2015.04.014

Cubric-Curik, V., Novosel, D., Brajkovic, V., Rota Stabelli, O., Krebs, S., Sölkner, J., Šalamon, D., Ristov, S., Berger, B., Trivizaki, S., Bizelis, I., Ferenčaković, M., Rothammer, S., Kunz, E., Simčič, M., Dovč, P., Bunevski, G., Bytyqi, H., Marković, B., … Medugorac, I. (2022). Large-scale mitogenome sequencing reveals consecutive expansions of domestic taurine cattle and supports sporadic aurochs introgression. *Evolutionary Applications*, *15*(4), 663–678. https://doi.org/10.1111/eva.13315

Cymbron, T., Loftus, R. T., Malheiro, M. I., & Bradley, D. G. (1999). Mitochondrial sequence variation suggests an African influence in Portuguese cattle. *Proceedings. Biological Sciences*, *266*(1419), 597–603. https://doi.org/10.1098/rspb.1999.0678

Dabney, J., & Meyer, M. (2019). Extraction of Highly Degraded DNA from Ancient Bones and Teeth. *Methods in Molecular Biology (Clifton, N.J.)*, *1963*, 25–29. https://doi.org/10.1007/978-1-4939-9176-1_4

Daly, K. G., Maisano Delser, P., Mullin, V. E., Scheu, A., Mattiangeli, V., Teasdale, M. D., Hare, A. J., Burger, J., Verdugo, M. P., Collins, M. J., Kehati, R., Erek, C. M., Bar-Oz, G., Pompanon, F., Cumer, T., Çakırlar, C., Mohaseb, A. F., Decruyenaere, D., Davoudi, H., … Bradley, D. G. (2018). Ancient goat genomes reveal mosaic domestication in the Fertile Crescent. *Science (New York, N.Y.)*, *361*(6397), 85–88. https://doi.org/10.1126/science.aas9411

Daly, K. G., Mattiangeli, V., Hare, A. J., Davoudi, H., Fathi, H., Doost, S. B., Amiri, S., Khazaeli, R., Decruyenaere, D., Nokandeh, J., Richter, T., Darabi, H., Mortensen, P., Pantos, A., Yeomans, L., Bangsgaard, P., Mashkour, M., Zeder, M. A., & Bradley, D. G. (2021). Herded and hunted goat genomes from the dawn of domestication in the Zagros Mountains. *Proceedings of the National Academy of Sciences*, *118*(25), e2100901118. https://doi.org/10.1073/pnas.2100901118

Daly, K. G., Mullin, V. E., Hare, A. J., Halpin, Á., Mattiangeli, V., Teasdale, M. D., Rossi, C., Geiger, S., Krebs, S., Medugorac, I., Sandoval-Castellanos, E., Özbaşaran, M., Duru, G., Gülcür, S., Pöllath, N., Collins, M., Frantz, L., Vila, E., Zidarov, P., … Bradley, D. G. (2025). Ancient genomics and the origin, dispersal, and development of domestic sheep. *Science*, *387*(6733), 492–497. https://doi.org/10.1126/science.adn2094

Decker, J. E., McKay, S. D., Rolf, M. M., Kim, J., Alcalá, A. M., Sonstegard, T. S., Hanotte, O., Götherström, A., Seabury, C. M., Praharani, L., Babar, M. E., Regitano, L. C. de A., Yildiz, M. A., Heaton, M. P., Liu, W.-S., Lei, C.-Z., Reecy, J. M., Saif-Ur-Rehman, M., Schnabel, R. D., & Taylor, J. F. (2014). Worldwide Patterns of Ancestry, Divergence, and Admixture in Domesticated Cattle. *PLOS Genetics*, *10*(3), e1004254. https://doi.org/10.1371/journal.pgen.1004254

Drummond, A. J., & Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, *7*, 214. https://doi.org/10.1186/1471-2148-7-214

Dutta, P., Talenti, A., Young, R., Jayaraman, S., Callaby, R., Jadhav, S. K., Dhanikachalam, V., Manikandan, M., Biswa, B. B., Low, W. Y., Williams, J. L., Cook, E., Toye, P., Wall, E., Djikeng, A., Marshall, K., Archibald, A. L., Gokhale, S., Kumar, S., … Prendergast, J. G. D. (2020). Whole genome analysis of water buffalo and global cattle breeds highlights convergent signatures of domestication. *Nature Communications*, *11*(1), 4739. https://doi.org/10.1038/s41467-020-18550-1

Edgar, R. C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, *32*(5), 1792–1797. https://doi.org/10.1093/nar/gkh340

Erven, J. A. M., Scheu, A., Verdugo, M. P., Cassidy, L., Chen, N., Gehlen, B., Street, M., Madsen, O., & Mullin, V. E. (2024). A High-Coverage Mesolithic Aurochs Genome and Effective Leveraging of Ancient Cattle Genomes Using Whole Genome Imputation. *Molecular Biology and Evolution*, *41*(5), msae076. https://doi.org/10.1093/molbev/msae076

Evershed, R. P., Payne, S., Sherratt, A. G., Copley, M. S., Coolidge, J., Urem-Kotsu, D., Kotsakis, K., Özdoğan, M., Özdoğan, A. E., Nieuwenhuyse, O., Akkermans, P. M. M. G., Bailey, D., Andeescu, R.-R., Campbell, S., Farid, S., Hodder, I., Yalman, N., Özbaşaran, M., Bıçakcı, E., … Burton, M. M. (2008). Earliest date for milk use in the Near East and southeastern Europe linked to cattle herding. *Nature*, *455*(7212), Article 7212. https://doi.org/10.1038/nature07180

Fages, A., Hanghøj, K., Khan, N., Gaunitz, C., Seguin-Orlando, A., Leonardi, M., McCrory Constantz, C., Gamba, C., Al-Rasheid, K. A. S., Albizuri, S., Alfarhan, A. H., Allentoft, M., Alquraishi, S., Anthony, D., Baimukhanov, N., Barrett, J. H., Bayarsaikhan, J., Benecke, N., Bernáldez-Sánchez, E., … Orlando, L. (2019). Tracking Five Millennia of Horse Management with Extensive Ancient Genome Time Series. *Cell*, *177*(6), 1419-1435.e31. https://doi.org/10.1016/j.cell.2019.03.049

Felsenstein, J. (1981). Evolutionary trees from DNA sequences: A maximum likelihood approach. *Journal of Molecular Evolution*, *17*(6), 368–376. https://doi.org/10.1007/BF01734359

Frantz, L. A. F., Bradley, D. G., Larson, G., & Orlando, L. (2020). Animal domestication in the era of ancient genomics. *Nature Reviews Genetics*, *21*(8), Article 8. https://doi.org/10.1038/s41576-020-0225-0

Frantz, L. A. F., Haile, J., Lin, A. T., Scheu, A., Geörg, C., Benecke, N., Alexander, M., Linderholm, A., Mullin, V. E., Daly, K. G., Battista, V. M., Price, M., Gron, K. J., Alexandri, P., Arbogast, R.-M., Arbuckle, B., Bălăşescu, A., Barnett, R., Bartosiewicz, L., … Larson, G. (2019). Ancient pigs reveal a near-complete genomic turnover following their introduction to Europe. *Proceedings of the National Academy of Sciences*, *116*(35), 17231–17238. https://doi.org/10.1073/pnas.1901169116

Gamba, C., Jones, E. R., Teasdale, M. D., McLaughlin, R. L., Gonzalez-Fortes, G., Mattiangeli, V., Domboróczki, L., Kővári, I., Pap, I., Anders, A., Whittle, A., Dani, J., Raczky, P., Higham, T. F. G., Hofreiter, M., Bradley, D. G., & Pinhasi, R. (2014). Genome flux and stasis in a five millennium

transect of European prehistory. *Nature Communications*, *5*(1), Article 1. https://doi.org/10.1038/ncomms6257

Gaunitz, C., Fages, A., Hanghøj, K., Albrechtsen, A., Khan, N., Schubert, M., Seguin-Orlando, A., Owens, I. J., Felkel, S., Bignon-Lau, O., de Barros Damgaard, P., Mittnik, A., Mohaseb, A. F., Davoudi, H., Alquraishi, S., Alfarhan, A. H., Al-Rasheid, K. A. S., Crubézy, E., Benecke, N., … Orlando, L. (2018). Ancient genomes revisit the ancestry of domestic and Przewalski's horses. *Science (New York, N.Y.)*, *360*(6384), 111–114. https://doi.org/10.1126/science.aao3297

Gifford-Gonzalez, D., & Hanotte, O. (2011). Domesticating Animals in Africa: Implications of Genetic and Archaeological Findings. *Journal of World Prehistory*, *24*(1), 1–23. https://doi.org/10.1007/s10963-010-9042-2

Gilbert, M., Nicolas, G., Cinardi, G., Van Boeckel, T. P., Vanwambeke, S. O., Wint, G. R. W., & Robinson, T. P. (2018). Global distribution data for cattle, buffaloes, horses, sheep, goats, pigs, chickens and ducks in 2010. *Scientific Data*, *5*(1), 180227. https://doi.org/10.1038/sdata.2018.227

Ginja, C., Guimarães, S., Fonseca, R. R. da, Rasteiro, R., Rodríguez-Varela, R., Simões, L. G., Sarmento, C., Belarte, M. C., Kallala, N., Torres, J. R., Sanmartí, J., Arruda, A. M., Detry, C., Davis, S., Matos, J., Götherström, A., Pires, A. E., & Valenzuela-Lamas, S. (2023). Iron age genomic data from Althiburos – Tunisia renew the debate on the origins of African taurine cattle. *iScience*, *26*(7). https://doi.org/10.1016/j.isci.2023.107196

Götherström, A., Anderung, C., Hellborg, L., Elburg, R., Smith, C., Bradley, D. G., & Ellegren, H. (2005). Cattle domestication in the Near East was followed by hybridization with aurochs bulls in Europe. *Proceedings of the Royal Society B: Biological Sciences*, *272*(1579), 2345–2351. https://doi.org/10.1098/rspb.2005.3243

Gouy, M., Tannier, E., Comte, N., & Parsons, D. P. (2021). Seaview Version 5: A Multiplatform Software for Multiple Sequence Alignment, Molecular Phylogenetic Analyses, and Tree Reconciliation. *Methods in Molecular Biology (Clifton, N.J.)*, *2231*, 241–260. https://doi.org/10.1007/978-1-0716-1036-7_15

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

Grigson, C. (1991). An African origin for African cattle? — Some archaeological evidence. *African Archaeological Review*, *9*(1), 119–144. https://doi.org/10.1007/BF01117218

Grigson, C. (2015). The fauna of Tell Nebi Mend (Syria) in the Bronze and Iron Age—a diachronic overview. Part 1: Stability and change—animal husbandry. *Levant*, *47*(1), 5–29. https://doi.org/10.1179/0075891415Z.00000000055

Groenen, M. A. M. (2016). A decade of pig genome sequencing: A window on pig domestication and evolution. *Genetics Selection Evolution*, *48*(1), 23. https://doi.org/10.1186/s12711-016-0204-2

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., & Gascuel, O. (2010). New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Systematic Biology*, *59*(3), 307–321. https://doi.org/10.1093/sysbio/syq010

Haak, W., Lazaridis, I., Patterson, N., Rohland, N., Mallick, S., Llamas, B., Brandt, G., Nordenfelt, S., Harney, E., Stewardson, K., Fu, Q., Mittnik, A., Bánffy, E., Economou, C., Francken, M., Friederich, S., Pena, R. G., Hallgren, F., Khartanovich, V., … Reich, D. (2015). Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature*, *522*(7555), Article 7555. https://doi.org/10.1038/nature14317

Haarkötter, C., Vinueza-Espinosa, D. C., Gálvez, X., Saiz, M., Medina-Lozano, M. I., Lorente, J. A., & Álvarez, J. C. (2023). A comparison between petrous bone and tooth, femur and tibia DNA analysis from degraded skeletal remains. *ELECTROPHORESIS*, *44*(19–20), 1559–1568. https://doi.org/10.1002/elps.202300097

Haber, M., Doumet-Serhal, C., Scheib, C., Xue, Y., Danecek, P., Mezzavilla, M., Youhanna, S., Martiniano, R., Prado-Martinez, J., Szpak, M., Matisoo-Smith, E., Schutkowski, H., Mikulski, R., Zalloua, P., Kivisild, T., & Tyler-Smith, C. (2017). Continuity and Admixture in the Last Five Millennia of Levantine History from Ancient Canaanite and Present-Day Lebanese Genome Sequences. *The American Journal of Human Genetics*, *101*(2), 274–282. https://doi.org/10.1016/j.ajhg.2017.06.013

Hanotte, O., Bradley, D. G., Ochieng, J. W., Verjee, Y., Hill, E. W., & Rege, J. E. O. (2002). African Pastoralism: Genetic Imprints of Origins and Migrations. *Science (New York, N.Y.)*, *296*(5566), 336–339. https://doi.org/10.1126/science.1069878

Hansen, H. B., Damgaard, P. B., Margaryan, A., Stenderup, J., Lynnerup, N., Willerslev, E., & Allentoft, M. E. (2017). Comparing Ancient DNA Preservation in Petrous Bone and Tooth Cementum. *PLOS ONE*, *12*(1), e0170940. https://doi.org/10.1371/journal.pone.0170940

Hansen, P. J. (2004). Physiological and cellular adaptations of zebu cattle to thermal stress. *Animal Reproduction Science*, *82–83*, 349–360. https://doi.org/10.1016/j.anireprosci.2004.04.011

Hassink, J., De Bruin, S. R., Berget, B., & Elings, M. (2017). Exploring the Role of Farm Animals in Providing Care at Care Farms. *Animals*, *7*(6), 45. https://doi.org/10.3390/ani7060045

Helmer, D., Blaise, É., Gourichon, L., & Saña-Seguí, M. (2018). Using cattle for traction and transport during the Neolithic period: Contribution of the study of the first and second phalanxes. *Bulletin de La Société Préhistorique Française*, *115*(1), 71–98.

Helmer, D., Gourichon, L., Monchot, H., Peters, J., & Saña, M. (2005). Identifying early domestic cattle from Pre-Pottery Neolithic sites on the Middle Euphrates using sexual dimorphism. *The First Steps of Animal Domestication*, 86–95.

Higuchi, R., Bowman, B., Freiberger, M., Ryder, O. A., & Wilson, A. C. (1984). DNA sequences from the quagga, an extinct member of the horse family. *Nature*, *312*(5991), Article 5991. https://doi.org/10.1038/312282a0

Hole, F. (1984). A Reassessment of the Neolithic Revolution. *Paléorient*, *10*(2), 49–60.

Hongo, H., Pearson, J., Öksüz, B., & Ilgezdi, G. (2009). The Process of Ungulate Domestication at Çayönü, Southeastern Turkey: A Multidisciplinary Approach focusing on *Bos* sp. and *Cervus elaphus*. *Anthropozoologica*, *44*(1), 63–78. https://doi.org/10.5252/az2009n1a3

Horsburgh, K. A., Prost, S., Gosling, A., Stanton, J.-A., Rand, C., & Matisoo-Smith, E. A. (2013). The Genetic Diversity of the Nguni Breed of African Cattle (Bos spp.): Complete Mitochondrial Genomes of Haplogroup T1. *PLOS ONE*, *8*(8), e71956. https://doi.org/10.1371/journal.pone.0071956

Horwitz, L. K., Tchernov, E., Ducos, P., Becker, C., Von Den Driesch, A., Martin, L., & Garrard, A. (1999). Animal Domestication in the southern Levant. *Paléorient*, *25*(2), 63–80.

Huelsenbeck, J. P., Ronquist, F., Nielsen, R., & Bollback, J. P. (2001). Bayesian Inference of Phylogeny and Its Impact on Evolutionary Biology. *Science*, *294*(5550), 2310–2314. https://doi.org/10.1126/science.1065889

Irving-Pease, E. K., Ryan, H., Jamieson, A., Dimopoulos, E. A., Larson, G., & Frantz, L. A. F. (2019). Paleogenomics of Animal Domestication. In C. Lindqvist & O. P. Rajora (Eds), *Paleogenomics: Genome-Scale Analysis of Ancient DNA* (pp. 225–272). Springer International Publishing. https://doi.org/10.1007/13836_2018_55

Juras, A., Chyleński, M., Ehler, E., Malmström, H., Żurkiewicz, D., Włodarczak, P., Wilk, S., Peška, J., Fojtík, P., Králík, M., Libera, J., Bagińska, J., Tunia, K., Klochko, V. I., Dabert, M., Jakobsson, M., & Kośko, A. (2018). Mitochondrial genomes reveal an east to west cline of steppe ancestry in Corded Ware populations. *Scientific Reports*, *8*(1), 11603. https://doi.org/10.1038/s41598-018-29914-5

Keller, A., Graefen, A., Ball, M., Matzas, M., Boisguerin, V., Maixner, F., Leidinger, P., Backes, C., Khairat, R., Forster, M., Stade, B., Franke, A., Mayer, J., Spangler, J., McLaughlin, S., Shah, M., Lee, C., Harkins, T. T., Sartori, A., … Zink, A. (2012). New insights into the Tyrolean Iceman's origin and phenotype as inferred by whole-genome sequencing. *Nature Communications*, *3*(1), 698. https://doi.org/10.1038/ncomms1701

Kennedy, M., Strolin, L., McMahon, J., Franklin, D., Flavel, A., Noble, J., Swift, L., Nassr, A., Fallon, S., & Thomas, H. (2023). Cult, herding, and 'pilgrimage' in the Late Neolithic of north-west Arabia: Excavations at a mustatil east of AlUla. *PLOS ONE*, *18*(3), e0281904. https://doi.org/10.1371/journal.pone.0281904

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

Koufariotis, L., Hayes, B. J., Kelly, M., Burns, B. M., Lyons, R., Stothard, P., Chamberlain, A. J., & Moore, S. (2018). Sequencing the mosaic genome of Brahman cattle identifies historic and recent introgression including polled. *Scientific Reports*, *8*(1), 17761. https://doi.org/10.1038/s41598-018-35698-5

Larson, G., Albarella, U., Dobney, K., Rowley-Conwy, P., Schibler, J., Tresset, A., Vigne, J.-D., Edwards, C. J., Schlumbaum, A., Dinu, A., Bălăçsescu, A., Dolman, G., Tagliacozzo, A., Manaseryan, N., Miracle, P., Van Wijngaarden-Bakker, L., Masseti, M., Bradley, D. G., & Cooper, A. (2007). Ancient DNA, pig domestication, and the spread of the Neolithic into Europe. *Proceedings of the National Academy of Sciences*, *104*(39), 15276–15281. https://doi.org/10.1073/pnas.0703411104

Lazaridis, I., Alpaslan-Roodenberg, S., Acar, A., Açıkkol, A., Agelarakis, A., Aghikyan, L., Akyüz, U., Andreeva, D., Andrijašević, G., Antonović, D., Armit, I., Atmaca, A., Avetisyan, P., Aytek, A. İ., Bacvarov, K., Badalyan, R., Bakardzhiev, S., Balen, J., Bejko, L., … Reich, D. (2022a). Ancient DNA from Mesopotamia suggests distinct Pre-Pottery and Pottery Neolithic migrations into Anatolia. *Science*, *377*(6609), 982–987. https://doi.org/10.1126/science.abq0762

Lazaridis, I., Alpaslan-Roodenberg, S., Acar, A., Açıkkol, A., Agelarakis, A., Aghikyan, L., Akyüz, U., Andreeva, D., Andrijašević, G., Antonović, D., Armit, I., Atmaca, A., Avetisyan, P., Aytek, A. İ., Bacvarov, K., Badalyan, R., Bakardzhiev, S., Balen, J., Bejko, L., … Reich, D. (2022b). The genetic history of the Southern Arc: A bridge between West Asia and Europe. *Science (New York, N.Y.)*, *377*(6609), eabm4247. https://doi.org/10.1126/science.abm4247

Lazaridis, I., Patterson, N., Anthony, D., Vyazov, L., Fournier, R., Ringbauer, H., Olalde, I., Khokhlov, A. A., Kitov, E. P., Shishlina, N. I., Ailincăi, S. C., Agapov, D. S., Agapov, S. A., Batieva, E., Bauyrzhan, B., Bereczki, Z., Buzhilova, A., Changmai, P., Chizhevsky, A. A., … Reich, D. (2025). The genetic origin of the Indo-Europeans. *Nature*, *639*(8053), 132–142. https://doi.org/10.1038/s41586-024-08531-5

Lefort, V., Longueville, J.-E., & Gascuel, O. (2017). SMS: Smart Model Selection in PhyML. *Molecular Biology and Evolution*, *34*(9), 2422–2424. https://doi.org/10.1093/molbev/msx149

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

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

Librado, P., Khan, N., Fages, A., Kusliy, M. A., Suchan, T., Tonasso-Calvière, L., Schiavinato, S., Alioglu, D., Fromentier, A., Perdereau, A., Aury, J.-M., Gaunitz, C., Chauvey, L., Seguin-Orlando, A., Sarkissian, C. D., Southon, J., Shapiro, B., Tishkin, A. A., Kovalev, A. A., … Orlando, L. (2021). The origins and spread of domestic horses from the Western Eurasian steppes. *Nature*, *598*(7882), 634. https://doi.org/10.1038/s41586-021-04018-9

Lindahl, T. (1993). Instability and decay of the primary structure of DNA. *Nature*, *362*(6422), 709–715. https://doi.org/10.1038/362709a0

Loftus, R. T., MacHugh, D. E., Bradley, D. G., Sharp, P. M., & Cunningham, E. P. (1994a). Evidence for two independent domestications of cattle. *Proceedings of the National Academy of Sciences*, *91*(7), 2757–2761. https://doi.org/10.1073/pnas.91.7.2757

Loftus, R. T., MacHugh, D. E., Ngere, L. O., Balain, D. S., Badi, A. M., Bradley, D. G., & Cunningham, E. P. (1994b). Mitochondrial genetic variation in European, African and Indian cattle populations. *Animal Genetics*, *25*(4), 265–271. https://doi.org/10.1111/j.1365-2052.1994.tb00203.x

MacHugh, D. E., Larson, G., & Orlando, L. (2017). Taming the Past: Ancient DNA and the Study of Animal Domestication. *Annual Review of Animal Biosciences*, *5*(1), 329–351. https://doi.org/10.1146/annurev-animal-022516-022747

MacHugh, D. E., Shriver, M. D., Loftus, R. T., Cunningham, P., & Bradley, D. G. (1997). Microsatellite DNA Variation and the Evolution, Domestication and Phylogeography of Taurine and Zebu Cattle (Bos taurus and Bos indicus). *Genetics*, *146*(3), 1071–1086. https://doi.org/10.1093/genetics/146.3.1071

MacHugh, D. E., Troy, C. S., McCormick, F., Olsaker, I., Eythórsdóttir, E., & Bradley, D. G. (1999). Early medieval cattle remains from a Scandinavian settlement in Dublin: Genetic analysis and comparison with extant breeds. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *354*(1379), 99–108; discussion 108-109. https://doi.org/10.1098/rstb.1999.0363

Maier, R., Flegontov, P., Flegontova, O., Işıldak, U., Changmai, P., & Reich, D. (2023). On the limits of fitting complex models of population history to f-statistics. *eLife*, *12*, e85492. https://doi.org/10.7554/eLife.85492

Margulies, M., Egholm, M., Altman, W. E., Attiya, S., Bader, J. S., Bemben, L. A., Berka, J., Braverman, M. S., Chen, Y.-J., Chen, Z., Dewell, S. B., Du, L., Fierro, J. M., Gomes, X. V., Godwin, B. C., He, W., Helgesen, S., Ho, C. H., Irzyk, G. P., … Rothberg, J. M. (2005). Genome sequencing in microfabricated high-density picolitre reactors. *Nature*, *437*(7057), Article 7057. https://doi.org/10.1038/nature03959

Marshall, F., & Hildebrand, E. (2002). Cattle Before Crops: The Beginnings of Food Production in Africa. *Journal of World Prehistory*, *16*(2), 99–143. https://doi.org/10.1023/A:1019954903395

Matthews, R. (2002). Zebu: Harbingers of doom in Bronze Age western Asia? *Antiquity*, *76*(292), 438–446. https://doi.org/10.1017/S0003598X00090542

Mattiangeli, V., cassidl, Daly, K. G., & Mullin, V. E. (2023). *Bleach extraction protocol: Damaged or degraded DNA recovery from bone or tooth powder*.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., & DePristo, M. A. (2010). The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, *20*(9), 1297–1303. https://doi.org/10.1101/gr.107524.110

McTavish, E. J., Decker, J. E., Schnabel, R. D., Taylor, J. F., & Hillis, D. M. (2013). New World cattle show ancestry from multiple independent domestication events. *Proceedings of the National Academy of Sciences*, *110*(15), E1398–E1406. https://doi.org/10.1073/pnas.1303367110

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

Naderi, S., Rezaei, H.-R., Pompanon, F., Blum, M. G. B., Negrini, R., Naghash, H.-R., Balkız, Ö., Mashkour, M., Gaggiotti, O. E., Ajmone-Marsan, P., Kence, A., Vigne, J.-D., & Taberlet, P. (2008). The goat domestication process inferred from large-scale mitochondrial DNA analysis of wild and domestic individuals. *Proceedings of the National Academy of Sciences*, *105*(46), 17659–17664. https://doi.org/10.1073/pnas.0804782105

Okonechnikov, K., Conesa, A., & García-Alcalde, F. (2016). Qualimap 2: Advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics (Oxford, England)*, *32*(2), 292–294. https://doi.org/10.1093/bioinformatics/btv566

Olivieri, A., Gandini, F., Achilli, A., Fichera, A., Rizzi, E., Bonfiglio, S., Battaglia, V., Brandini, S., Gaetano, A. D., El-Beltagi, A., Lancioni, H., Agha, S., Semino, O., Ferretti, L., & Torroni, A. (2015). Mitogenomes from Egyptian Cattle Breeds: New Clues on the Origin of Haplogroup Q and the Early Spread of Bos taurus from the Near East. *PLOS ONE*, *10*(10), e0141170. https://doi.org/10.1371/journal.pone.0141170

Orlando, L., Allaby, R., Skoglund, P., Der Sarkissian, C., Stockhammer, P. W., Ávila-Arcos, M. C., Fu, Q., Krause, J., Willerslev, E., Stone, A. C., & Warinner, C. (2021). Ancient DNA analysis. *Nature Reviews Methods Primers*, *1*(1), 1–26. https://doi.org/10.1038/s43586-020-00011-0

Outram, A. K., Stear, N. A., Bendrey, R., Olsen, S., Kasparov, A., Zaibert, V., Thorpe, N., & Evershed, R. P. (2009). The Earliest Horse Harnessing and Milking. *Science (New York, N.Y.)*, *323*(5919), 1332–1335. https://doi.org/10.1126/science.1168594

Park, S. D. E., Magee, D. A., McGettigan, P. A., Teasdale, M. D., Edwards, C. J., Lohan, A. J., Murphy, A., Braud, M., Donoghue, M. T., Liu, Y., Chamberlain, A. T., Rue-Albrecht, K., Schroeder, S., Spillane, C., Tai, S., Bradley, D. G., Sonstegard, T. S., Loftus, B. J., & MacHugh, D.

E. (2015). Genome sequencing of the extinct Eurasian wild aurochs, Bos primigenius, illuminates the phylogeography and evolution of cattle. *Genome Biology*, *16*(1), 234. https://doi.org/10.1186/s13059-015-0790-2

Parker, C., Rohrlach, A. B., Friederich, S., Nagel, S., Meyer, M., Krause, J., Bos, K. I., & Haak, W. (2020). A systematic investigation of human DNA preservation in medieval skeletons. *Scientific Reports*, *10*(1), 18225. https://doi.org/10.1038/s41598-020-75163-w

Patel, A. K., & Meadow, R. H. (2017). South Asian contributions to animal domestication and pastoralism: Bones, genes, and archaeology. In U. Albarella, M. Rizzetto, H. Russ, K. Vickers, & S. Viner-Daniels (Eds), *The Oxford Handbook of Zooarchaeology* (p. 0). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199686476.013.19

Perry-Gal, L., Erlich, A., Gilboa, A., & Bar-Oz, G. (2015). Earliest economic exploitation of chicken outside East Asia: Evidence from the Hellenistic Southern Levant. *Proceedings of the National Academy of Sciences*, *112*(32), 9849–9854. https://doi.org/10.1073/pnas.1504236112

Peters, J., & Arbuckle, B. (2012). Subsistence and beyond: Animals in Neolithic Anatolia. *The Neolithic in Turkey*, *VOL. 6*, 1–65.

Pigière, F., & Smyth, J. (2023). First evidence for cattle traction in Middle Neolithic Ireland: A pivotal element for resource exploitation. *PLOS ONE*, *18*(1), e0279556. https://doi.org/10.1371/journal.pone.0279556

Pinhasi, R., Fernandes, D., Sirak, K., Novak, M., Connell, S., Alpaslan-Roodenberg, S., Gerritsen, F., Moiseyev, V., Gromov, A., Raczky, P., Anders, A., Pietrusewsky, M., Rollefson, G., Jovanovic, M., Trinhhoang, H., Bar-Oz, G., Oxenham, M., Matsumura, H., & Hofreiter, M. (2015). Optimal Ancient DNA Yields from the Inner Ear Part of the Human Petrous Bone. *PLOS ONE*, *10*(6), e0129102. https://doi.org/10.1371/journal.pone.0129102

Posit Team. (2023). *RStudio: Integrated Development Environment for R* [Computer software]. Posit Software, PBC. http://www.posit.co/

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., Bakker, P. I. W. de, Daly, M. J., & Sham, P. C. (2007). PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *The American Journal of Human Genetics*, *81*(3), 559–575. https://doi.org/10.1086/519795

Rambaut, A. (2018). *FigTree v1.4.4*. https://tree.bio.ed.ac.uk/software/figtree/

Rambaut, A., Drummond, A. J., Xie, D., Baele, G., & Suchard, M. A. (2018). Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. *Systematic Biology*, *67*(5), 901–904. https://doi.org/10.1093/sysbio/syy032

Reich, D., Green, R. E., Kircher, M., Krause, J., Patterson, N., Durand, E. Y., Viola, B., Briggs, A. W., Stenzel, U., Johnson, P. L. F., Maricic, T., Good, J. M., Marques-Bonet, T., Alkan, C., Fu, Q., Mallick, S., Li, H., Meyer, M., Eichler, E. E., … Pääbo, S. (2010). Genetic history of an archaic

hominin group from Denisova Cave in Siberia. *Nature*, *468*(7327), Article 7327. https://doi.org/10.1038/nature09710

Rossi, C., Sinding, M.-H. S., Mullin, V. E., Scheu, A., Erven, J. A. M., Verdugo, M. P., Daly, K. G., Ciucani, M. M., Mattiangeli, V., Teasdale, M. D., Diquelou, D., Manin, A., Bangsgaard, P., Collins, M., Lord, T. C., Zeibert, V., Zorzin, R., Vinter, M., Timmons, Z., … Bradley, D. G. (2024). The genomic natural history of the aurochs. *Nature*, *635*(8037), 136–141. https://doi.org/10.1038/s41586-024-08112-6

Saitou, N., & Nei, M. (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, *4*(4), 406–425. https://doi.org/10.1093/oxfordjournals.molbev.a040454

Scheu, A., Powell, A., Bollongino, R., Vigne, J.-D., Tresset, A., Çakırlar, C., Benecke, N., & Burger, J. (2015). The genetic prehistory of domesticated cattle from their origin to the spread across Europe. *BMC Genetics*, *16*(1), 54. https://doi.org/10.1186/s12863-015-0203-2

Schibler, J., Elsner, J., & Schlumbaum, A. (2014). Incorporation of aurochs into a cattle herd in Neolithic Europe: Single event or breeding? *Scientific Reports*, *4*(1), 5798. https://doi.org/10.1038/srep05798

Schmid, C., & Schiffels, S. (2023). Estimating human mobility in Holocene Western Eurasia with large-scale ancient genomic data. *Proceedings of the National Academy of Sciences*, *120*(9), e2218375120. https://doi.org/10.1073/pnas.2218375120

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

Scott, A., Reinhold, S., Hermes, T., Kalmykov, A. A., Belinskiy, A., Buzhilova, A., Berezina, N., Kantorovich, A. R., Maslov, V. E., Guliyev, F., Lyonnet, B., Gasimov, P., Jalilov, B., Eminli, J., Iskandarov, E., Hammer, E., Nugent, S. E., Hagan, R., Majander, K., … Warinner, C. (2022). Emergence and intensification of dairying in the Caucasus and Eurasian steppes. *Nature Ecology & Evolution*, *6*(6), 813–822. https://doi.org/10.1038/s41559-022-01701-6

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

Skoglund, P., Storå, J., Götherström, A., & Jakobsson, M. (2013). Accurate sex identification of ancient human remains using DNA shotgun sequencing. *Journal of Archaeological Science*, *40*(12), 4477–4482. https://doi.org/10.1016/j.jas.2013.07.004

Taylor, W. T. T., Pruvost, M., Posth, C., Rendu, W., Krajcarz, M. T., Abdykanova, A., Brancaleoni, G., Spengler, R., Hermes, T., Schiavinato, S., Hodgins, G., Stahl, R., Min, J., Alisher kyzy, S., Fedorowicz, S., Orlando, L., Douka, K., Krivoshapkin, A., Jeong, C., … Shnaider, S. (2021).

Evidence for early dispersal of domestic sheep into Central Asia. *Nature Human Behaviour*, *5*(9), Article 9. https://doi.org/10.1038/s41562-021-01083-y

Teasdale, M. (2024). *karyotypeR* [R]. https://github.com/teasdalm/karyotypeR (Original work published 2018)

Teasdale, M. D., & Bradley, D. G. (2012). The Origins of Cattle. In J. E. Womack (Ed.), *Bovine Genomics* (1st edn, pp. 1–10). Wiley. https://doi.org/10.1002/9781118301739.ch1

Thalmann, O., & Perri, A. R. (2019). Paleogenomic Inferences of Dog Domestication. In C. Lindqvist & O. P. Rajora (Eds), *Paleogenomics: Genome-Scale Analysis of Ancient DNA* (pp. 273–306). Springer International Publishing. https://doi.org/10.1007/13836_2018_27

Thorvaldsdóttir, H., Robinson, J. T., & Mesirov, J. P. (2013). Integrative Genomics Viewer (IGV): High-performance genomics data visualization and exploration. *Briefings in Bioinformatics*, *14*(2), 178–192. https://doi.org/10.1093/bib/bbs017

Troy, C. S., MacHugh, D. E., Bailey, J. F., Magee, D. A., Loftus, R. T., Cunningham, P., Chamberlain, A. T., Sykes, B. C., & Bradley, D. G. (2001). Genetic evidence for Near-Eastern origins of European cattle. *Nature*, *410*(6832), Article 6832. https://doi.org/10.1038/35074088

Upadhyay, M. R., Chen, W., Lenstra, J. A., Goderie, C. R. J., MacHugh, D. E., Park, S. D. E., Magee, D. A., Matassino, D., Ciani, F., Megens, H.-J., van Arendonk, J. a. M., & Groenen, M. a. M. (2017). Genetic origin, admixture and population history of aurochs (Bos primigenius) and primitive European cattle. *Heredity*, *118*(2), 169–176. https://doi.org/10.1038/hdy.2016.79

Verdugo, M. P., Mullin, V. E., Scheu, A., Mattiangeli, V., Daly, K. G., Maisano Delser, P., Hare, A. J., Burger, J., Collins, M. J., Kehati, R., Hesse, P., Fulton, D., Sauer, E. W., Mohaseb, F. A., Davoudi, H., Khazaeli, R., Lhuillier, J., Rapin, C., Ebrahimi, S., … Bradley, D. G. (2019). Ancient cattle genomics, origins, and rapid turnover in the Fertile Crescent. *Science (New York, N.Y.)*, *365*(6449), 173–176. https://doi.org/10.1126/science.aav1002

Vigne, J.-D. (2015). Early domestication and farming: What should we know or do for a better understanding? *Anthropozoologica*, *50*(2), 123–150. https://doi.org/10.5252/az2015n2a5

Vigne, J.-D., Carrère, I., Briois, F., & Guilaine, J. (2011). The Early Process of Mammal Domestication in the Near East: New Evidence from the Pre-Neolithic and Pre-Pottery Neolithic in Cyprus. *Current Anthropology*, *52*(S4), S255–S271. https://doi.org/10.1086/659306

Wang, C., Zhan, X., Bragg-Gresham, J., Kang, H. M., Stambolian, D., Chew, E. Y., Branham, K. E., Heckenlively, J., Fulton, R., Wilson, R. K., Mardis, E. R., Lin, X., Swaroop, A., Zöllner, S., & Abecasis, G. R. (2014). Ancestry Estimation and Control of Population Stratification for Sequence-based Association Studies. *Nature Genetics*, *46*(4), 409–415. https://doi.org/10.1038/ng.2924

Wang, C., Zhan, X., Liang, L., Abecasis, G. R., & Lin, X. (2015). Improved Ancestry Estimation for both Genotyping and Sequencing Data using Projection Procrustes Analysis and Genotype

Imputation. *American Journal of Human Genetics*, *96*(6), 926–937. https://doi.org/10.1016/j.ajhg.2015.04.018

Weisdorf, J. L. (2005). From Foraging To Farming: Explaining The Neolithic Revolution. *Journal of Economic Surveys*, *19*(4), 561–586. https://doi.org/10.1111/j.0950-0804.2005.00259.x

Weiss, H., Courty, M.-A., Wetterstrom, W., Guichard, F., Senior, L., Meadow, R., & Curnow, A. (1993). The Genesis and Collapse of Third Millennium North Mesopotamian Civilization. *Science*, *261*(5124), 995–1004. https://doi.org/10.1126/science.261.5124.995

Wendorf, F., & Schild, R. (1994). Are the early holocene cattle in the eastern sahara domestic or wild? *Evolutionary Anthropology: Issues, News, and Reviews*, *3*(4), 118–128. https://doi.org/10.1002/evan.1360030406

Wiener, M. H. (2014). The Interaction of Climate Change and Agency in the Collapse of Civilizations ca. 2300–2000 BC. *Radiocarbon*, *56*(4), S1–S16. https://doi.org/10.2458/azu_rc.56.18325

Wilkin, S., Ventresca Miller, A., Fernandes, R., Spengler, R., Taylor, W. T.-T., Brown, D. R., Reich, D., Kennett, D. J., Culleton, B. J., Kunz, L., Fortes, C., Kitova, A., Kuznetsov, P., Epimakhov, A., Zaibert, V. F., Outram, A. K., Kitov, E., Khokhlov, A., Anthony, D., & Boivin, N. (2021). Dairying enabled Early Bronze Age Yamnaya steppe expansions. *Nature*, *598*(7882), Article 7882. https://doi.org/10.1038/s41586-021-03798-4

Wingett, S. W., & Andrews, S. (2018). FastQ Screen: A tool for multi-genome mapping and quality control. *F1000Research*, *7*, 1338. https://doi.org/10.12688/f1000research.15931.2

Wint, W., & Robinson, T. P. (2007). *Gridded livestock of the world 2007*. Food and Agriculture Organization of the United Nations.

Xia, X.-T., Achilli, A., Lenstra, J. A., Tong, B., Ma, Y., Huang, Y.-Z., Han, J.-L., Sun, Z.-Y., Chen, H., Lei, C.-Z., Hu, S.-M., & Chen, N.-B. (2021). Mitochondrial genomes from modern and ancient Turano-Mongolian cattle reveal an ancient diversity of taurine maternal lineages in East Asia. *Heredity*, *126*(6), 1000–1008. https://doi.org/10.1038/s41437-021-00428-7

Zeder, M. A. (2008). Domestication and early agriculture in the Mediterranean Basin: Origins, diffusion, and impact. *Proceedings of the National Academy of Sciences*, *105*(33), 11597–11604. https://doi.org/10.1073/pnas.0801317105

Zeder, M. A. (2012). The Domestication of Animals. *Journal of Anthropological Research*. https://doi.org/10.3998/jar.0521004.0068.201

Zhang, X., Yang, L., Hou, L., Li, H., Xiang, H., & Zhao, X. (2021). Ancient Mitogenomes Reveal the Domestication and Distribution of Cattle During the Longshan Culture Period in North China. *Frontiers in Genetics*, *12*, 759827. https://doi.org/10.3389/fgene.2021.759827