UPPSALA
UNIVERSITET

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 2228

Uncovering the Past through ancient DNA

The Fate and Legacy of the last hunter-gatherers in Western Europe and Northwestern Africa

LUCIANA G. SIMÕES

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2023

ISSN 1651-6214
ISBN 978-91-513-1684-0
URN urn:nbn:se:uu:diva-493382

Dissertation presented at Uppsala University to be publicly examined in Ekmansalen, Evolutionary Biology Centre, Norbyvägen 14, Uppsala, Friday, 3 March 2023 at 13:15 for the degree of Doctor of Philosophy. The examination will be conducted in English. Faculty examiner: Associate Professor Martin Sikora (University of Copenhagen (Globe Institute), Denmark).

Abstract

Simões, L. G. 2023. Uncovering the Past through ancient DNA. The Fate and Legacy of the last hunter-gatherers in Western Europe and Northwestern Africa. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology* 2228. 65 pp. Uppsala: Acta Universitatis Upsaliensis. ISBN 978-91-513-1684-0.

The genetic analysis of ancient human remains has revolutionized the study of human history, opening a direct window onto the demographic and evolutionary events that shaped our species' past. I use state-of-the-art ancient DNA methodologies, from sample collection and DNA extraction to data analysis, to study the ancient past of the western Mediterranean region, where climate does not favour DNA preservation. After the Last Glacial Maximum, amenable climatic conditions enabled the development of agriculture in the Levant, initiating the Neolithic period. In Europe, the transition from foraging to farming was driven by the migration of people from Anatolia, but in North Africa, evidence indicates a cultural diffusion, instead of population replacement. In this thesis, I show that this transition was in fact ignited by the migration of early farmers from Iberia. Moreover, a different migration wave, originating in the Levant and expanding within Africa, was associated with pastoralism in that region during the Neolithic. While the Neolithic transition is one of the most studied periods of pre-history, earlier periods are comparatively under-studied. Using whole genome sequencing data for 36 hunter-gatherers from Iberia and France, I observed that genetic lineages rooted in the Palaeolithic, survived throught the Mesolithic. Mesolithic hunter-gatherer populations formed social units that were not based on familial bonds; exchanges between groups avoided consanguinity. Coexistence with the first farming communities resulted in unidirectional admixture patterns, as we do not find gene flow from farmers to the last hunter-gatherers. Finally, using a multidisciplinary approach to study an exceptional individual of African descent buried in a Mesolithic shell midden, we find that the burial of this man during the transatlantic slave trade period could be an example of the maintenance of African cultural practices by African people displaced to Europe. My thesis highlights the power of ancient DNA analysis to uncover events and patterns of the human evolutionary history, but also that integrative approaches, where different lines of evidence are combined, can lead to exciting findings.

Keywords: human evolution, archaeogenomics, ancient DNA, population genetics, hunter-gatherer

Luciana G. Simões, Department of Organismal Biology, Human Evolution, Norbyvägen 18 A, Uppsala University, SE-752 36 Uppsala, Sweden.

ISSN 1651-6214
ISBN 978-91-513-1684-0
URN urn:nbn:se:uu:diva-493382 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-493382)

To my husband, my children and my parents.

List of papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.

I **Simões, L. G.**, Günther, T. Who were the Mesolithic people: the archaeogenomic perspective. *Oxford Handbook of Mesolithic Europe*, *in press*

II **Simões, L. G.***, Peyroteo-Stjerna, R.*, Bernhardsson, C., Günther, T., Marchand, G., Edlund, H., Chetty, D., Bouquin, D., Garmond, N., Vialet, A., Jakobsson, M. Genomic ancestry and social dynamics of the last hunter-gatherers of Atlantic France. *manuscript*

III Peyroteo-Stjerna, R.*, **Simões, L. G.***, Araújo, A.C., Diniz, M., Arias, P., Cubas, M., Valdiosera, C., Günther, T., Jakobsson, M. The genomic history of the Iberian hunter-gatherers, from the Late Pleistocene to the Early Holocene. *manuscript*

IV Peyroteo-Stjerna, R.*, **Simões, L. G.***, Fernandes, R., Lopes, G., Günther, T. and Jakobsson, M. (2022) Multidisciplinary investigation reveals an individual of West African origin buried in a Portuguese Mesolithic shell midden four centuries ago. *Journal of Archaeological Science: Reports*, 42: 103370

V **Simões, L. G.**, Günther, T., Martínez-Sánchez, R. M., Vera-Rodríguez, J. C., Iriarte, E. Rodríguez-Varela, R., Bokbot, Y., Valdiosera, C., Jakobsson, M. Neolithization of Northwestern Africa was ignited by migrants from Iberia and Levant. *in revision*

*These authors contributed equally to the study

The following papers were also published during my graduate studies but are not part of this dissertation.

Silva, A.P., Curveira-Santos, G., Kilshaw, K., Newman, C., Macdonald, D. W., **Simões, L. G.**, Rosalino, L. M. (2017) Climate and anthropogenic factors determine site occupancy in Scotland's Northern-range badger population: implications of context-dependent responses under environmental change, *Diversity and Distributions*, 23(6), 627-639.

Günther, T., Malmström, H., Svensson, E. M., Omrak, A., Sánchez-Quinto, F., Kılınç, G. M., Krzewińska, M., Eriksson, G., Fraser, M., Edlund, H., Munters, A. R., Coutinho, A., **Simões, L. G.**, Vicente, M., Sjölander, A., Sellevold, B. J., JÃ¸rgensen, R., Claes, P., Shriver, M. D., Valdiosera, C., Netea, M. G., Apel, J., Lidén, K., Skar, B., Storå, J., Götherström, A., Jakobsson, M. (2018) Population genomics of Mesolithic Scandinavia: Investigating early postglacial migration routes and high-latitude adaptation. *PLoS biology*, 16(1):e2003703.

Valdiosera, C., Günther, T., Vera-Rodríguez, J. C., Ureña, I., Iriarte, E., Rodríguez-Varela, R., Simões, **Simões, L. G.**, Martínez-Sánchez, R. M., Svensson, E. M., Malmström, H., Rodríguez, L., Bermúdez de Castro, J. M., Carbonell, E., Alday, A., Hernández Vera, J. A., Götherström ,A., Carretero, J. M., Arsuaga, J. L., Smith, C. I., Jakobsson, M. (2018) Four millennia of Iberian biomolecular prehistory illustrate the impact of prehistoric migrations at the far end of Eurasia. *Proceedings of the National Academy of Sciences*, 115(13), 3428-3433.

Sánchez-Quinto, F., Malmström, H., Fraser, M. Girdland-Flink, L., Svensson, E. M., **Simões, L. G.**, George, R., Hollfelder, N., Burenhult, G., Noble, G., Britton, K., Talamo, S., Curtis, N., Brzobohata, H., Sumberova, R., Götherström, A., Storå, J., Jakobsson, M. (2019) Megalithic tombs in western and northern Neolithic Europe were linked to a kindred society. *Proceedings of the National Academy of Sciences*, 116(19), 9469–9474.

Contents

1. Introduction

The Mediterranean region, extending from the Levant to the Atlantic Ocean, has witnessed some of the most revolutionary chapters in human prehistory. During glacial periods, Mediterranean refugia harboured innumerable species of flora and fauna, on which human populations relied, and from those regions started the recolonization of other latitudes during interstadial periods. These special environmental dynamics, with rich biodiversity and relatively hospitable climate, make the Mediterranean a region of interest for human evolutionary history, especially since the Last Glacial Maximum (LGM).

For millennia, the Mediterranean Sea, enclaved between southern Europe, North Africa and the Levant, has impacted human societies: it either posed a barrier or, upon the development of the necessary means, acted as a migration corridor for human populations. The Iberian Peninsula, at the far west European Mediterranean, bridges the Mediterranean Sea and the Atlantic Ocean. In the European Atlantic façade, which extends from the current territory of Portugal to Great Britain, Ireland and reaching Scandinavia, long-standing human contacts have resulted in shared demographic and cultural dynamics. These links are rooted in populations inhabiting the European Atlantic coast since the Stone Age and are tangled with the ocean and its resources.

The field of population genetics aims at understanding demographic and evolutionary processes by investigating patterns of genetic variation over time, within and across populations. The analysis of ancient DNA (aDNA) under the principles of population genetics has revolutionized the study of human evolution. It opened a direct window onto the study of human evolutionary history, allowing the reconstruction of past demographic events [1, 2, 3] that is no longer limited to the genomic data of present-day humans. With my work, I wish to contribute to understanding how, when and where (and ultimately why) evolutionary events in the human past happened in the Mediterranean region, shaping the genetic diversity patterns of our species. For that, I retrieved aDNA data from archaeological remains originating from western Europe and northwestern Africa, where warm and wet environmental conditions do not favour DNA preservation.

From a young age, I was fascinated by human history. The decision to follow a scientific education was based on the desire to use science and technology to contribute to the future of mankind. My academic journey laid the path and when the opportunity presented to join the young and challenging research field of aDNA, I took it with excitement - I had found the combination of both my biggest interests.

2. Ancient DNA

The analysis of the genetic profiles of ancient individuals has been a powerful tool to reconstruct demographic events in the human past [1, 2, 3]. aDNA data provide the opportunity to look directly into the evolutionary past without relying on modern-day data to do so, and many surprising discoveries have been made since the beginning of this field. Questions that have been traditionally posed by linguists, archaeologists, anthropologists are now approached with this set of tools and methods. Unlike those traditional disciplines, aDNA has the advantage of not relying on artifacts, tools or scriptures for studying the human past. Instead, these studies can be taken directly on biological human remains, such as bones, teeth, hair or even feces. Moreover, as a result of technological advances, aDNA analyses require very small fragments of biological tissue, as opposed to near-complete and morphologically informative skeletons. aDNA studies are enabled by theoretic principles and computational developments of the discipline of population genetics, but pose additional challenges deriving from intrinsic characteristics of aDNA, discussed in section 3, such as damage, contamination, small sample sizes.

Since the first attempts of retrieving aDNA in the 1980's [4, 5], aDNA development as a field was prompted by a series of technological advancements, the most relevant being Next Generation Sequencing (NGS). NGS revolutionised the study of genomics and molecular biology and was the real breakthrough for aDNA studies given the quick, cheap and massive throughput of sequences, compared to the previous Sanger sequencing. NGS allows for the generation of billions of independent sequence reads, and depending on the DNA preservation in each specimen, each position in the genome can be covered several times, conferring analytical power.

DNA, the double-stranded macromolecule that encodes genetic information, is thought to be preserved in bones and teeth because of its affinity to hydroxyapatite and collagen, the main components of inorganic and organic fractions in osseous materials [6]. But as a result of ceased DNA repairing systems after the death of an organism, DNA accumulates damages as it degrades over time [7]. One of aDNA's resulting characteristics is highly fragmented molecules, which rarely reach 100 bp. Fragmentation was a problem for previous methods, based on Polymerade Chain Reaction (PCR) amplification. NGS technology, in turn, actually requires that long DNA strands are cut into short fragments prior to sequencing. This happens naturally in aDNA, making NGS very suited for this purpose.

Successful aDNA studies are increasing due to the combination of reducing costs of NGS technology, the development of specifically tailored molecular methods for aDNA retrieval, and the use of good DNA reservoirs. Molecular methods have been developed to suit the specific characteristics of aDNA: short fragments, single-stranded sequences, chemical alterations, abasic sites, etc. Several kinds of bone samples pre-digestion have been tested to improve the recovery of endogenous molecules [8, 9, 10] and minimize human and microbial contamination [11]. Extraction methods were optimized for retrieving short fragments of aDNA, with silica columns and varying volumes of guanidine-based binding buffers to increase the recovery of fragments as short as 20bp, and library building methods were also adapted to aDNA [12]. Moreover, the identification of the best substrates as sources of aDNA added to the success of aDNA research. Tooth roots have been identified as a good source [13], and particularly the outer layer of the tooth - cementum - more than the inner dentine [14, 9]. The petrous bone, which is part of the temporal bone and the hardest and most dense bone in the mammal body, is one of the best sources of endogenous aDNA, exceeding those obtained from teeth by 4- to 16-fold [15], even though these differences may be heavily dependent on specific inter-skeleton/-site variation [16]. Lastly, the optimization of bioinformatic pipelines and analytical methods (reviewed in [17]) are crucial technical advances that enable the exponential growth of the archaeogenetics field.

Many early aDNA studies relied on mitochondrial DNA (mtDNA), which is more readily available for extraction, as there are usually several copies available in the cell, whereas nuclear DNA is limited to one copy per cell. However, mtDNA is only inherited maternally and is haploid, while the human nuclear genome is inherited from both parents (diploid) and consists of more than three billion base pairs (bp). On average, between the two sets of chromosomes, only a couple in a thousand of sites differ. Moreover, the variation found within a single human and between different individuals is within the same range. This means that a lot more resolution can be obtained from genomic studies (or the study of thousands or millions of genetic variants) than single locus studies (such as mtDNA): the study of many of such variable positions allows for powerful inferences about population structure [18].

3. aDNA research challenges and opportunities

NGS has brought aDNA to the centre of evolutionary biology through high-throughput sequencing (Knapp 2010). The short fragment sizes, typical of aDNA, are of ideal length for the preparation of sequencing libraries, which include the ligation of universal sequencing adapters with intrinsic primer annealing regions to both ends of DNA molecules in a DNA extracts [19]. The untargeted cluster amplification of NGS also allows for massive parallel sequencing of DNA molecules, and impressive amounts of sequencing data can be generated per ancient individual, at an ever reducing cost [20]. The amount of data generated confers statistical power to the downstream analysis and allows for distinguishing modern contamination and endogenous portions of DNA, for example.

During NGS library preparation, adapter ligation is not species-specific and the DNA library built from a aDNA extract will include the metagenomic pool of DNA molecules extracted from the ancient sample. That is usually a combination of endogenous DNA, environmental DNA from fungi and bacteria and possible modern contaminant human DNA. Except for some rare cases of preservation, microbial contaminants are often the most frequent source of DNA extracted from an ancient human remain [12]. Microorganisms colonize bones and teeth during the targeted individuals' decomposition to such an extent that the endogenous proportion of DNA extracted from a human remain rarely exceeds 5%. Modern human contamination, a particularly challenging aspect of aDNA studies on humans (because contaminant DNA and endogenous DNA originate from the same species), can be easily introduced during handling at archaeological excavations or at museums, for example. To amend the effects of contamination, several methods have been developed. The most important way to handle microbial contamination is to bioinformatically map the sequence data to the target species (in this case, human), and discard the non-target species data. Microbial contamination can also be molecularly reduced (which maximizes the sequencing yield of target-species data generated) by hybridization-based enrichment of parts of the genome [21, 22] or of the complete genome [23]. Modern contamination, on the other hand, can be handled bioinformatically, as I describe below.

Post-mortem DNA degradation is highly dependent on climacteric conditions. This process involves both fragmentation and chemical changes to the DNA molecules [24]. DNA fragmentation is induced by hydrolytic depurination, resulting in single-strand breaks [25, 24]. The hydrolysis of the amine group of cytosines to a carbonyl group, forming uracil, is the most typical aDNA degradation signal. Methylated (by epigenetic processes) cytosines can also be degraded into thymine [24]. NGS technology enabled the confirmation of previous results about this very distinctive pattern of aDNA: C to T transitions [26]. During DNA replication, adenines pairs with uracils (the deaminated cytosines) and following that, adenines are paired with thymines. This way, original Cs are sequenced as Ts and these substitutions are incorporated into the sequencing reads, resulting in an excess of Ts instead of Cs in comparison to the human reference genome [12]. This pattern is more prevalent at fragments termini, which are more likely to be single stranded and more exposed to hydrolysis [12].

However, the current understanding of these processes has transformed the potential challenges into opportunities: observing the patterns of degradation is now used to attest DNA authenticity and study epigenetic patterns from ancient genomes. C to T deamination rates at reads ends are usually plotted for a visual inspection of the damage expected for authentic aDNA sequences and average read length is also observed (Figure 3.1). The damaged nature of aDNA becomes advantageous to discern authentic human aDNA from modern contamination: C to T deamination changes are usually not found in modern contaminating DNA [27] and based on those signals, bioinformatic methods have been developed to separate the modern contaminating fraction from the endogenous human aDNA fraction [28, 29]. To minimize the risk of introducing further modern contamination, aDNA laboratory work is performed in specialised labs, where several measures and guidelines ensure that the environment is kept as free from DNA as possible [30, 31].

Most aDNA studies these days use uracil-DNA-glycosylase (UDG) and endonuclease VIII partial or full treatment of libraries to cleave uracils and repair the damaged ends of the fragments, mitigating the effects of deamination on down-stream analyses [32, 33]. This in-vitro method removes abasic and uracil sites prior to DNA library preparation [34, 35]. This can be done only partially, meaning that the first two or three bases in each fragment will preserve the damage signal (which can be used to validate data authenticity) or fully, removing all uracils. This method brings the advantage of using all SNPs during data analysis and maximizing the amount of genetic information that can be retrieved from sequencing or genotyping data [32]. This way, analyses are not restricted to transversion sites but can include transitions (C-T and G-A SNPs); these frequently have to be omitted, or large proportions of the sequencing reads discarded towards the termini, if libraries are not damage repaired, since it is not possible to distinguish variation originating from DNA damage and real polymorphic variation.

The climate of the region where the archaeological remains are retrieved from has great impact on the extension and frequency of the DNA damage. Environmental conditions such as humidity, temperature, salinity, and pH have a strong effect on DNA preservation. However, conditions are likely to vary

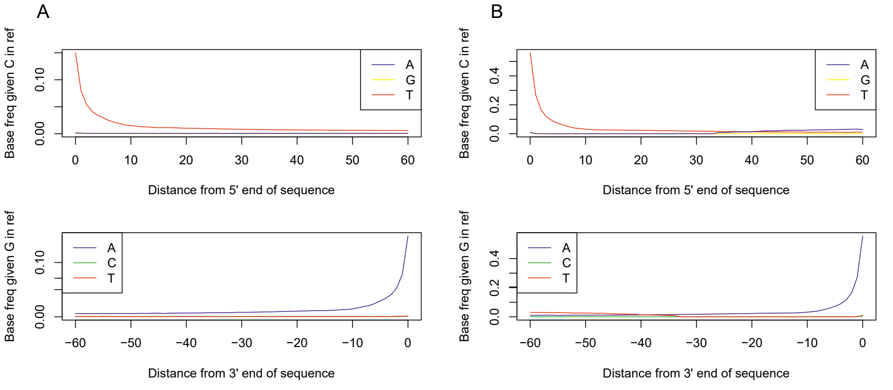

Figure 3.1. Deamination plots characteristic of authentic ancient DNA data, where an increased frequency of C-T substitutions is observed towards the 5’ ends of the DNA fragments, and G-A substitutions on the 3’ ends. The rate of observed substitutions differs between (A) and (B); (A) corresponds to genomic data generated for a sample from Portugal, dated to ca. 400 years ago, mug019/CAM1, and (B) corresponds to data generated for a sample from France, dated to ca. 7.000 years ago, hoe005. In spite of the higher deamination rate, sample (B) yielded a higher proportion of human sequences, 0.45, compared to 0.20 for sample (A).

within a site or even an archaeological stratum, making preservation variable on a microscale [12]. Preserved environments, such as caves that act as a buffer to temperature and other climacteric fluctuations, usually provide aDNA samples with better preservation. There is no evident correlation between fragment length and age [36] and older samples that were frozen or quickly desiccated after death are favoured by the inhibition of degradation processes, presenting much better DNA preservation than younger samples from hotter, more humid environments.

4. Patterns and processes of human genetic variation

Variation is introduced in the genome by mutation. Different variants of a genetic marker (named alleles) can change in frequency due to selection, if they are advantageous for an individual, but most commonly, due to random genetic drift. This process corresponds to the random fluctuation of the frequency of an allele over time in a finite population. When a population size is drastically decreased (a process that is known as genetic bottleneck), the effects of drift are quicker than in the pre-bottleneck population, which would have more variation to draw the next generation from. Drift can bring a certain allele to fixation (when it becomes the only allele in the population) or to loss (when it is no longer observed in any individual in the population). Both of these scenarios result in decreased diversity. Because genetic drift is random, it will have different effects in two populations that split from a common ancestor, resulting in their differentiation. Note that gene flow, the successful movement and transfer of genetic material resulting from the movement of people from one region to another, can occur through large-scale migrations as well as in small increments over many generations when two populations live relatively close together but exchange only a limited number of migrants [37]. If gene flow between two groups is low or absent, the genetic differences accumulate in each of them.

Mutation will introduce new variants in each of the populations differently, and those will be more noticeable if new variants rise in frequency as a consequence of random genetic drift. On the other hand, if a newly generated variant increases the fitness, or survival/reproductive success of its carrier individual, it will more likely increase in frequency in that population. Selective pressures are determined by the environment. If two populations with a common origin are subject to different selective pressures, their genetic background will be refined to better fit the environment, by favouring certain alleles and discarding others [38]. It should be noted that in population genomic analysis it is not easy to distinguish between the impacts of various demographic events, such as a population bottleneck, and selection. These often result in similar alterations in diversity patterns [39]. It is often assumed that demographic events leave traces at a genome-wide scale, thus the importance of population genomics, which studies numerous loci simultaneously to better understand the roles of evolutionary processes such as genetic drift, gene flow or natural selection [40].

Besides mutation, also gene flow and recombination are sources that can maintain genetic diversity. DNA recombination is the process in which genetic material is exchanged between different regions of the same chromosome (or between multiple chromosomes), resulting in the reshuffling of the sequences inherited from both parents before their transmission to the offspring. It reduces Linkage Disequilibrium (LD), which is elevated when long blocks of particular combinations of alleles of neighbouring variants are present and common in a population. Such a block of alleles that is inherited together from a single parent is called a haplotype.

Studies from the fields of genetics, palaeontology and archaeology have consistently shown that anatomically modern humans (AMH) originated in Africa [41]. Genetically, African populations are more diverse than non-Africans. That is evident when comparing for example Africans and non-Africans observed heterozygosity, the proportion of genomic sites that have two different alleles at a genetic locus. While heterozygosity in modern-day Africans is (slightly) above 0.1%, in all non-African populations it is lower than 0.08% [42]. LD is not only reduced in African populations but is also found to increase in other groups as their geographic distance from Africa increases [43]. Worldwide patterns of genetic diversity and differentiation are best explained by a serial founder effect that started in Africa [44, 45, 43], followed by an Out-of-Africa migration event that was the starting point for the peopling of Eurasia, Oceania and Americas. This means that the patterns of population diversity that we observe today are the result of population splits and subsequent differentiation that have happened over ca. 300,000 years, since the first modern human populations diverged in Africa [46].

A serial founder effect such as the described has the consequence of reducing the effective population size (N_e). With reduced N_e, a population undergoes stronger genetic drift, leading to an increased random change of allele frequencies and a loss of genetic diversity over time. N_e represents the size of an idealized population under the Wright-Fisher model (constant finite size, non overlapping generations, random mating). However, human populations cannot meet these conditions, as they are rarely constant in size, generations do overlap and mating is not fully random. The term "inbreeding" is used when offspring is produced from the mating of close relatives. This can happen as a consequence of a small population size, where the background relatedness between all individuals is high, or by consanguinity, if the parents had a close kin relationship. Consanguinity can result from the social reproductive behaviour of endogamy, or within-group marriage and can have deleterious effects on the offspring. Exogamy, on the other hand, happens when mates are exchanged between subpopulations. The mate choice and/or the population's social dynamics and structure have direct consequences on a population's genetic variation patterns, such as observed in Runs of Homozygosity (RoH, section 7.5.5, Figure 4.1). Genomic DNA can be used as a powerful tool to infer ancient populations' social dynamics (Papers II and III). As the field moves

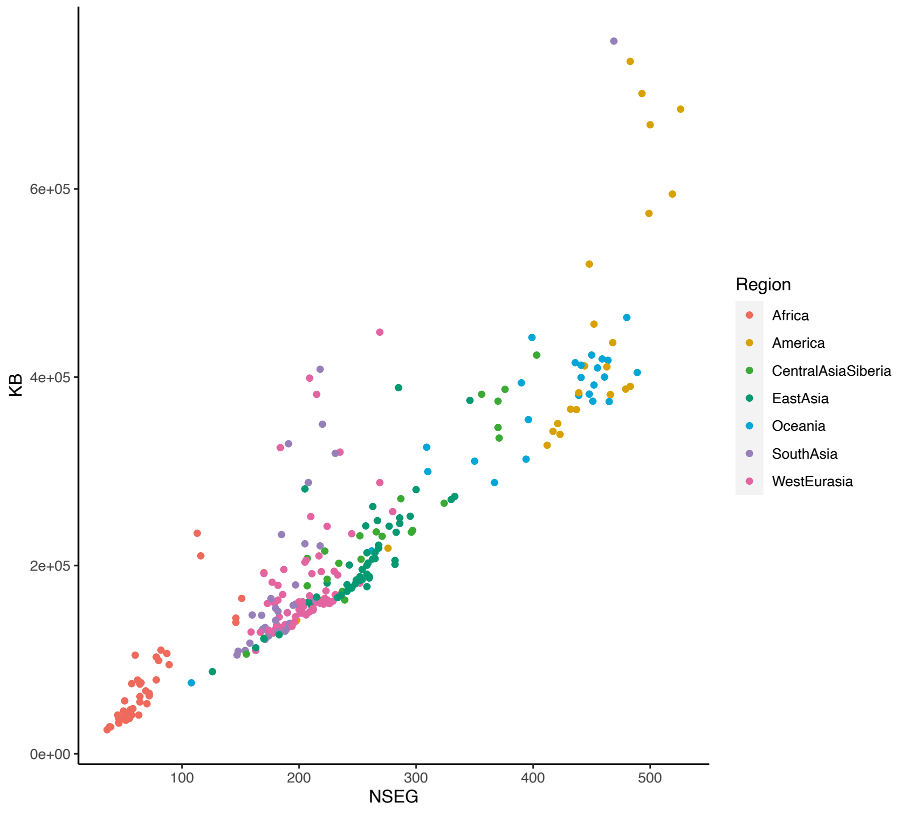

Figure 4.1. RoH for worldwide modern continental populations from the Simons Genomes Diversity Project (SGDP) [48]. The number (NSEG) and cumulative length (in kilobases, KB) of RoH correlates. It increases with distance from Africa (serial founder effect). It is visible that some datapoints deviate from the correlation trend between number and cumulative length of RoH, as a reflect of social and reproductive behaviour, with small populations practicing consanguinity showing the longest and more frequent RoH.

from large-scale studies to a more local perspective under a multidisciplinary approach, ancient mating practices will be better understood, as well as other socio-cultural dynamics [47]. However, datasets that allow such inferences for ancient hunter-gatherers are still rare.

5. Ancient genetic variation in Europe and North Africa

AMH origin and morphological evolution is debated, as mosaics of feature associated with "modern" and "archaic" human groups are found in the African archaeological record [49]. In North Africa, human fossils dating to ca. 300 thousand years ago (kya) from Jebel Irhoud, Morocco, are an example of such mixed morphology [50] and show that the presence of hominid groups in that region is long-standing.

Human history has been greatly influenced by climatic events, which often resulted in alterations to the patterns of genetic diversity. The out-of-Africa migration event was determinant for the peopling of other continents [51, 41] and was likely facilitated by climatic changes [52]. Alternative routes for the dispersal of AMH out of Africa, via Egypt or Arabia have been debated [53]. The Sahara Desert would be a natural obstacle to these dispersal of AMH out of sub-Saharan Africa. Paleoenvironmental and paleoclimatic studies have revealed recurrent humid episodes in North Africa, that resulted in green Sahara periods (GSPs), with savanna expansion throughout most of the desert. As many as 230 GSPs have been identified in the last 8 million years [54]. GSPs take ca. 3 thousand years (ky) to develop and to end, after peaking for 4 to 8 ky. Studies have used different types of data to estimate the dates for GSPs (e.g. [55, 56]). Many of these dates could broadly coincide with major human migration events, including the genetically determined period for the out-of-Africa migration; for example, wet phases were identified around 80,000 years ago (80 kya) and 110/120 kya [54, 56, 57].

After the Out-of-Africa migration ca. 80 kya, AMH colonized Eurasia. The first evidence of AMH in Eurasia is found in the Levant and dates to the Middle to Upper Palaeolithic [58]. By 45 to 40 kya, AMH had dispersed through Europe, with paleontological and archaeological records of occupation in Eastern, Central and Western Europe [59, 60]. As AMH made their way into Eurasia, they met Neanderthals, a hominin group which had lived in Europe since ca. 300 kya. The observation that these two groups met and produced fertile offspring is one of the most important findings made through the analysis of genomic aDNA [61, 62]. A consequence of this admixture event is the presence of ca. 2% of Neanderthal ancestry in all humans living outside sub-Saharan Africa today [61, 62]. Human populations ancestral to all non-sub-Saharan Africans have been estimated to have mixed with Neanderthals between 37 and 86 kya [63], and thus shortly after the dispersal of AMH from Africa.

The LGM, between 26.5-19 kya, is another example of how climatic changes resulted in dramatic changes in the demography of human populations. Ice sheets reached their maximum extents and sea levels were at lowest point [64]. Then, in the Upper Pleistocene, average global temperatures were around 5 or 6°C lower than today and overall climatic conditions were drier, colder and windier. The atmospheric dust loading was 10 times heavier than what we currently observe, which was reflected in strong differences in vegetation cover [65]. While Northern Europe was covered with permanent summer ice sheets, steppe and tundra environment dominated in southern Europe [66].

Before 35 kya, a genetic turn-over has likely happened in Europe. Between 34 and 26 kya, a genetic component associated to individuals of the Gravettian culture became predominant in Europe. Pre-LGM European hunter-gatherers, such as Goyet Q116-1 from Belgium [67], PM1 from Romania [10], Sunghir III [68] and Kostenki 14 [69] from Russia, share genetic ancestry and archaeogenomic data has provided no evidence of substantial genetic influx from elsewhere during that period [67, 70]. Sequenced individuals that predated that period seem have made little genetic contribution to later populations [67]. However, the genetic ancestry of the population that preceded that cultural change, represented by GoyetQ116-1, seems to have survived and is detected in individuals associated with the Magdalenian culture in western Europe, ca. 19-12 kya. The LGM resulted in a severe bottleneck for the hunter-gatherer populations [71], which is reflected in reduced genomic and mitochondrial diversity [10, 72, 70].

With the end of the last ice age and the deglaciation of ice callots, sea levels rose abruptly around 14 kya [64] and the climate became milder. This event marked the end of the Palaeolithic, which had been dominated in Europe by archaeological cultures such as the Gravettian, Magdalenian and Azilian cultures. After the Palaeolithic, a new archaeological period started, a period that covers the transition between late glacial hunter gatherers and Neolithic farmers - the Mesolithic. A detailed picture of the patterns of genetic diversity during the Mesolithic in Europe can be found in Paper I.

During the early Holocene, after the end of the LGM, the Sahara was nearly completely vegetated, with abundant water availability [54, 52]. This Green-Sahara phase (11-5 kya) was characterized by high amounts of rainfall that supported a system of linked lakes, rivers, and inland deltas. The region was then occupied by diverse vegetation, which facilitated the dispersal of species, several aquatic, as well as the establishment of human populations [52, 73]. Human demographic patterns in this region changed rapidly between 10 and 5 kya. The archaeological record is rich and documents intensive human occupation, initially by pottery-using hunter-gatherers, and then by increasingly mobile cattle livestock herders [74, 75], with rock engravings of large-game hunting and livestock herding scenes. It has been estimated from the archaeological record that a massive population size increase happened shortly after

11 kya, followed by a temporary decline 7.6 to 6.7 kya and a major collapse between 6.3 and 5.2 kya [76].

Compared with the extensive ancient genomic studies in Eurasia, North African pre-history remains largely unstudied from a genetic perspective. Genetically, modern North African human populations result from a complex human demographic history, greatly due to this region's location, in between sub-Saharan Africa, Europe and the Middle East. Modern [77] and ancient [78] genetic data analysis have provided evidence for a back-to-Africa migration event, which is estimated to predate the onset of the Holocene (about 13 kya). In fact, late Pleistocene hunter-gatherers from Morocco have been shown to have genetic affinity to Holocene hunter-gatherers from the Near East (Natufians) and to sub-Saharan African populations. This finding corroborates the pre-Neolithic gene flow between out-of-Africa populations and north African populations, as well as across the Sahara [78]. However, definite details about the back-to-Africa migration remain unknown.

The genetic make-up of Upper Palaeolithic north Africans is very distinct from contemporary European populations; no quantifiable genetic contribution from European hunter-gatherers has been found in ancient north Africans living ca. 15 kya [78]. A common ancestry with Europe has so far only been identified in ancient north Africa during the Late Neolithic [79].

6. Archaeological and genetic evidence of the Neolithic transition in the Mediterranean and western Europe

One of the most important events in human prehistory was initiated when the warming climate of the Early Holocene favoured the transition from hunting and gathering to producing food, in the Levant. This marked the beginning of a new prehistorical age, the Neolithic, in Western Eurasia. Starting around 10,000 years ago, the hunter-gatherers from the Fertile Crescent gradually shifted subsistence practices from hunting and gathering to sedentary farming [80]. Agricultural practices spread through the European continent and the so-called "Neolithic package" included elements such as: managed cultivation of plants and animals, ceramics, food storage, sedentary settlements, polished stone tools and other portable artefacts [81, 82]. The archaeological record documents highly increased population numbers associated with the new food-producing economies in Europe [83].

This is one of the most intensely studied periods in archaeology and has also been a major focus of archaeogenetics. Prior to aDNA evidence, the hypothesis of a cultural diffusion, with hunter-gatherer populations adopting the new lifestyle and associated practices from neighbors or trade networks was considered [83, 84]. Archaeogenomic studies have played a major role in clarifying that the spread of the Neolithic lifestyle throughout Europe resulted from the migration of people who carried the Neolithic traditions, farming economy and associated material culture [85, 86, 87, 88, 89, 90, 91]. The population discontinuity from the Mesolithic to the Neolithic has been shown with uniparental markers, with the introduction of new mt haplogroups which were absent in Europe before the Neolithic (e.g. K, J, W, X, H, HV, V, T2 and N1a), and also new Y-haplogroups (e.g. G2a, J2, E3b and F). The simultaneous appearance of paternal and maternal lineages suggested that the migration process was not associated with a strong sex-bias. Genomic aDNA data has refined these findings [92].

The Neolithic transition in Europe has followed two main routes [93, 94]. The "Continental route" expanded across Central Europe, from the Balkan Peninsula, following the river valleys of the Danubian and the Dnipro and further west and north; this route is archaeologically associated with the Neolithic Lineal Pottery complex (Linearbandkeramik, LBK). The "Mediterranean route" followed the shores of the Mediterranean sea and is associated with Impressed Ware technology in the eastern and central Mediterranean and with Cardial

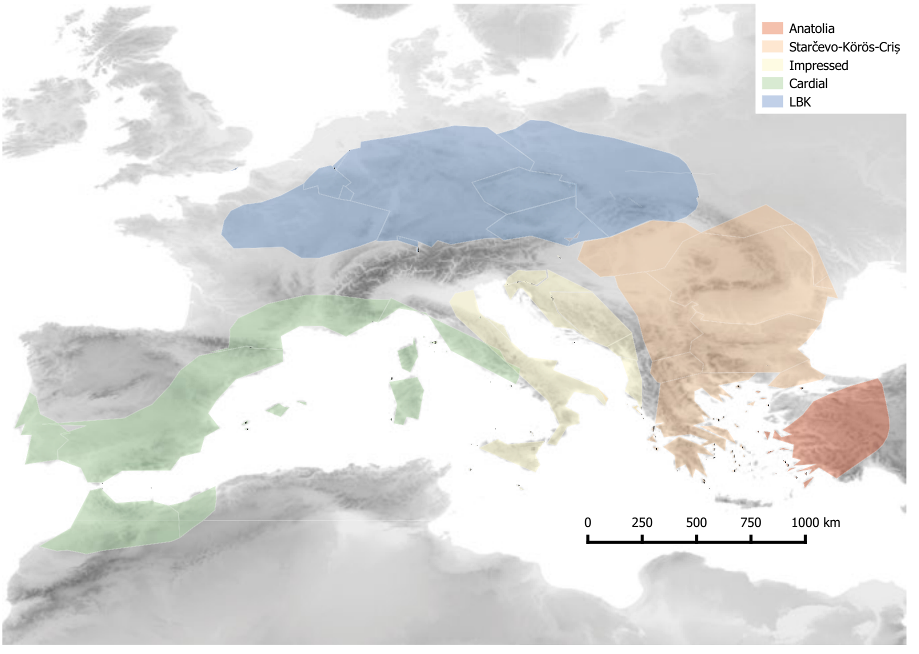

Figure 6.1. Main Early Neolithic cultures of Central and Western Europe approximately 10,000 BP in Anatolia to 7,500 BP in West Iberia and Northwest Africa. Adapted from [89, 95]

Ware in the western Mediterranean (Figure 6.1). Both routes expanded to reach Brittany, modern-day France, after 7 kya.

The first chronological modelling of the Neolithic diffusion through the Mediterranean suggested a regular population spread of ca. 1 km/year [96]. More recent models supported less linear scenarios based on a coastal diffusion of the agro-pastoral system in the Western Mediterranean through seafaring colonists [97, 98]. The spread of the earliest Neolithic settlements from Southeast Italy (ca. 8000-7900 cal BP) to Portugal (ca. 7500 cal BP) was rapid, but discontinuous (leapfrog diffusion). It has been suggested that whenever the progress was slower, technical and cultural blending of the Neolithic traditions with preexisting local technology resulted in new developments and acquisitions to the Neolithic package [99]. However, it remains hard to interpret and evaluate the role of Mesolithic populations in the regional consolidation of the Neolithic [100]. Genetically, we observe that the migrating early farmers admixed to different degrees with local groups of hunter-gatherers along the expansion routes [88, 90, 101, 102]. Early Neolithic populations from Central Europe populations, such as LBK, show less extensive admixture with WHG than Cardial Western European Neolithic populations, following an East to West cline [102, 95]. During the Neolithic period, the proportions of hunter-gatherer ancestry increased [87, 88, 103, 104].

The Neolithic development in northwest Africa happened ca.7400 cal BP, contemporaneously [105] or soon after the Iberian Neolithisation. Evidence points to contacts and cultural exchange across the Strait of Gibraltar [105, 106, 107]. Different models for the Neolithisation of the Maghreb have been proposed and debated [108]: an Iberian origin (preceded by an expansion from Southern France) [109, 110], a north African route of expansion through the Mediterranean Maghreb [111, 107], or even a synchronous arrival of Neolithic in both Southern Iberia and coastal Morocco [105]. The oldest Sicilian pottery, Pre-Sentinellian Impressa Arcaica from ca. 7900 years ago is distinguishable from Cardial, but in the Tingitan peninsula (the African portion of the Gibraltar Strait), the first pottery is characterized as Cardial of archaic aspect, with some features more typical of Impressed ceramics, and clear affinities with the archaic Cardial from Catalonia-Valencia as well as, to a lesser extent, with decorations present in examples of Impressa from other Iberian regions, mainly Valencia (and Andalusia).

Evidence against the North African route has increased since the 1970s, with little documented presence of agriculture in Mediterranean Algeria before 7000 years ago, and no evidence of agriculture in Tunisia at that time. However, evidence of livestock husbandry and pottery can be found there. This has been interpreted as a "failure" of the agricultural expansion in the Eastern Maghreb [112], where the emergence of food-producing societies is connected to native forager groups integrating new technologies and subsistence strategies, adapting them to local traditions and needs [100]. Several studies have highlighted the importance of the development of Epipaleolithic communities to the Neolithisation process of North Africa [74, 113, 100]. Hunter-gatherers exploited the broad range of environments that resulted from the early Holocene climatic fluctuations [114, 115, 116]. Decreasing and fluctuating rainfall and climatic unpredictability may have resulted in the development of strategies to manage resources, such as storage and delayed consumption of wild plants [117] and the use of pottery [118, 113] and more importantly, the development of an economy based on nomadic pastoralism and exploitation of wild plants [119, 74, 118].

The multitude of archaeogenetic studies focusing on the European Neolithic transition strikingly contrasts with the scarcity of analogous studies in North Africa. A single study so far has reported aDNA data from this period in North Africa [79], showing that Early Neolithic farmers at the site of Ifri n'Amr o'Moussa, in Central Morocco, show long-term genetic continuity since the Upper Paleolithic, with no traces of the Anatolian ancestry found in European Early Neolithic farmers. Archaeological excavation campaigns have retrieved Cardial-like ceramics from this site, similar to that found throughout the western Mediterranean, as well as domestic cereals. In spite of some striking similarities with the western Mediterranean European Neolithic package, the genetic evidence of Early Neolithic Morocco so far suggests that the Neolithisation process occurred through the adoption of innovations by local

Epipalaeolithic communities instead of a demic diffusion, as observed in Europe. This is in line with archaeological investigation supporting an important development of the Epipaleolithic communities to the Neolithisation process [105, 100]. With Paper V of this thesis, we present new evidence showing that the first stages of the Neolithisation process in northwestern Africa were started by the migration of Neolithic farmers from Europe across the Mediterranean. The genetic component related to Anatolian farmers had so far only been observed in North Africa in the Late Neolithic, with a 2000 years delay to the arrival of that ancestry to West Mediterranean Europe [79]. Our study shows that gene flow from western Europe to North Africa happened already during the earliest Neolithic phase of northwestern Africa (Paper V).

7. Methods

Unveiling the molecular record carried in the genomes of ancient individuals and the populations they belonged to must comply with some specific methodologies, in order to obtain trustworthy results. While there are variations and alternatives to those described in this chapter, I focus on summarizing the methods applied in the research presented in this dissertation. Some of the most commonly used techniques and methods used in aDNA research are also introduced, even if not applied in this work, mostly as a reference for the state-of-the-art of archaeogenomics.

7.1 Ancient samples collection and preparation

The samples analysed in my thesis were collected from museum collections or directly from excavations in Portugal, Spain, France and Morocco (Figure 7.1). Human remains were inspected and the best element(s) for each of the targeted individuals was evaluated and selected for sampling. Preference was given to the petrous part of the temporal bone and teeth [120, 15, 16]. Depending on museum policies, which often do not allow whole bones to leave the museum premises, for Papers II, III and IV, a majority of the bone/tooth elements were sampled on site, in pre-cleaned museum facilities. If the museum policies allowed, entire bone/tooth elements were directly taken to dedicated aDNA facilities in Uppsala, Sweden, and sampled according to a protocol that broadly follows the steps described next, but under a fully controlled, decontaminated laboratory. In museum facilities, bone elements were decontaminated prior to sampling, by mechanic cleaning and UV exposure. Based on the recent discovery of improved DNA retrieval from bone pieces comparing to bone powder [10], mostly bone/tooth pieces were sampled.

7.2 aDNA extraction

Two methods were used for aDNA extraction in the work of this thesis, published by Yang et al. [121] and Dabney et al. [12], with punctual adaptations described in the Papers. Extraction methods optimized for aDNA usually employ silica-based extraction columns and mostly differ from those used for modern DNA given the need to minimize the co-extraction of potential inhibitors of downstream molecular analyses, such as library building or amplification [122]. Adaptations to the binding buffer composition and volume

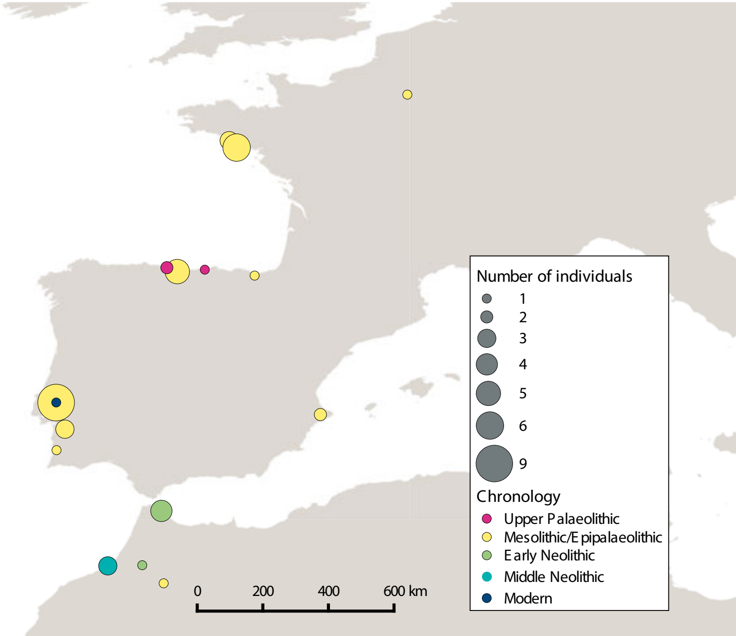

Figure 7.1. Overview of the entire ancient DNA dataset generated for the total of the projects in this thesis.

used resulted in an optimized recovery of DNA fragments as small as 35 bp [12], and therefore I applied this extraction method for particularly challenging, degraded samples. Extraction and library blanks were run alongside the extraction and library building process to monitor possible contamination introduced during the sample processing.

7.3 Library building and sequencing

The prime sequencing approaches for generating aDNA data are, since the introduction of NGS technology, shotgun and target-enriched sequencing. Shotgun approaches involve sequencing of the whole genome (Whole Genome Sequencing, WGS), or as much as available in the DNA extract and then converted into a sequencing library. This ensures an unbiased sequencing of the extracted DNA, including private mutations and rare alleles, which SNP chips or SNP capture sets do not target [123]. However, poor aDNA preservation and microbial contamination can make the shotgun sequencing of the endogenous genomic DNA a strenuous and prohibitively expensive process. Target-enriching methods selectively capture and enrich for genomic regions of interest, such as the mitogenome [21, 124], one chromosome [22], or even the entire genome [23]. A popular capture panel, developed to target 1,2 million autosomal SNPs (the so-called "1240k panel") for ancestry analyses

[90], has been frequently used by laboratories doing aDNA research, and thus a large number of individuals for which aDNA data is available today have been processed using this method. In-solution targeted capture approaches involve the hybridization of magnetic-beaded probes, designed to target genome regions of interest, to complementary DNA fragments [22]. This can substantially increase the relative fraction of endogenous DNA molecules in the total metagenomic pool in the DNA library, maximizing the ratio between data obtained and sequencing efforts, and be particularly efficient and cost-effective for samples with extra low endogenous DNA proportions [90].

However, hybridization enrichment by targeted capture of a user-defined panel of SNPs has disadvantages that can compromise the information obtained from an ancient sample. The main issue arising from this method is possibly the comprehensive comparison of the captured samples with data from other studies, since only overlapping target regions can be co-analysed [125]. Additional challenges include increased clonality and reduced data complexity [23, 126], effective hybridization of targeted sequences for heavily degraded molecules and the inevitably biased nature of the recovered library fragments [127, 23]. Moreover, ascertainment bias, that results from non-random sampling of SNPs [128, 123], may have downstream implications that are less obvious than those listed above, such as systematic deviation of population genetic statistics from theoretical expectations [129, 130]. The ascertained SNPs will tend to be of intermediate frequency in the panel of individuals used to ascertain the captured SNPs. Thus, SNPs of low minor allele frequency will tendentially be missed and those obtained will be biased toward populations related to the discovery panel [131]. Many approaches have been suggested to correct for ascertainment bias [132, 129, 131], but SNP panels ascertainment schemes are difficult to account for in aDNA statistical analyses, as SNPs are usually ascertained from modern groups, with a tendency to European populations, and therefore ancient, non-European genetic variation may not be sequenced and available for analysis [123].

Archaeological remains are not a limitless resource and aDNA research relies on its availability. Concerns about the sustainability and ethics of current bone sampling practices and generated genomic data availability have been raised [133, 134]. Current aDNA practices and methodologies are destructive of the bone tissues; this means that the optimal tissues for extraction (according to the current state-of-the-art), once studied, will not be available for future generations.

Archaeological remains are also often rare, especially for older periods, in regions where archaeological research is patchy and regions with sub-optimal conditions for biological material preservation. Some specimens are unique and unreproducible. Reconstructing the genomic signature of an individual that lived in the past is one way to immortalize the past and its legacy. It is our responsibility, as a scientific community, to ensure that the best possible outcome of our research is reached and the WGS of an ancient individual whenever possible is, as I perceive it, almost an ethical obligation. On one hand, the more loci sequenced, the more independent instances of the same evolutionary history are captured for an individual. On the other hand, the completeness of data may open future possibilities, maybe enabled by future analytical methods, which are likely to bear new insights into human evolutionary history and perspectives that are not within the reach of today's analysis. We are already observing examples of exciting discoveries that rise from the analysis of data generated in the past under new perspective and different ideas [135]. The field evolves as experiments are made, and while I do not believe that the research done today should be restrained waiting for more advanced methods, it is our duty to generate data in the best quality and amount possible, so that the destructive process of archaeological remains results in the best possible outcome for the future.

For NGS sequencing, DNA extracts must be converted into sequencing libraries, that are used for untargeted cluster amplification for massive parallel sequencing of DNA molecules. Even though single-strand library building protocols are more efficient in retrieving short molecules, in part owing to the reduction of purification steps to remove the excess of adaptors [136], their practical optimization and implementation is not always straightforward. Hence, for this thesis, only double stranded libraries were used for sequencing aDNA. Double stranded library building utilizes a modified form of Illumina's library preparation kit and protocol [19]. The process starts with the reparation of fragments' ends, by extension of overhanging single strands at the fragment termini; these present a high incidence of deaminated citosine sites and as a consequence of the end-repair step, adenines are integrated in the sequence, as complementary to uracils, instead of guanines. Adapters are ligated to blunt-end aDNA fragments, and then filling in the adapters creates double stranded molecules. Ligation of adapters allows the sample DNA to attach to a solid flow cell in a sequencer machine, where the strand can then be amplified and sequenced with fluorescent nucleotides [19]. The DNA library is prepared to undergo PCR, with one or both primers indexed with individual "tags", to discriminate between DNA samples. This allows for multiple samples to be pooled and sequenced together, thus saving time and money [19].

Prior to library building, UDG and endonuclease VIII treatment may be used to mitigate the effects of aDNA deamination; we used USER enzyme (a premixed version of the two enzymes) [32]. This treatment excises uracils from DNA sequences from DNA, leaving behind an abasic site and a single-strand break, which compromises the reparation of overhanging single-strand ends. This results in the reduction of the amount of sequence information available for downstream analysis and thus it was only applied on extracts for which good DNA preservation was verified during initial sample screening stages.

7.4 aDNA data processing

Subtle differences between sequencing platforms and bioinformatic pipelines may result in small biases that can mimic fine-scale genetic signatures, and even supersede the proportion of genetic variation resulting from geography [137]. One attainable measure to reduce such biases is to uniformize the bioinformatic pipeline for processing the data, which was done in this thesis, including for comparative aDNA data (whenever fastq files were available for download) as described below. All data generated for this thesis was shotgun sequenced on HiSeq 10X or on NovaSeq 6000. Sequencing data was first demultiplexed according to the indexing primer sequence used for each sample in the sequenced pool. Raw fastq files were trimmed for sequencing adapters.

Forward and reverse paired-end reads were merged when an overlap of at least 11 bp was found and mapped against the human reference genome, build 37 (hs37d5). Most libraries built were sequenced more than once, frequently until depletion (ca. 50% clonality). Resulting bam files were merged per library. To remove PCR duplicates, a consensus of reads with identical start and end positions was generated. Reads shorter than 35 bp, with more than 10% mismatches to the human reference genome or a mapping quality of less than 30 were filtered out. A single bam file per sample was finally created by merging all 90 percent consensus library bam files.

7.5 aDNA Data analysis

While there are several types of polymorphism, or possible variation in the DNA sequence, base substitutions (Single Nucleotide Polymorphism or SNP) are the most common form, and also the form that is primarily studied in human population genomics studies, including aDNA. Recombination creates statistical independence between markers. If two loci are statistically independent, they can be seen as independent stochastic outcomes of the same ancestral process, and the more loci are analysed, the greater is the resolution of population genetic structure. However, certain regions of the genome do not recombine, namely the mitochondrial DNA and the Y chromosome. This results in the inheritance of these genetic markers as a block of alleles, in a combination that is preserved from generation to generation - a haplotype.

7.5.1 Markers of genetic variation in aDNA studies

Haplogroup information for both mitochondrial DNA and Y-chromosome is commonly analyzed in ancient DNA studies. Each haplogroup encloses related patterns of sequences (haplotypes) that are genetically relatively similar. These patterns correlate with the geographic origins of populations, traced through the maternal (mtDNA) and paternal (Y chromosome) lineages. Mitochondrial and Y-chromosome haplogroup frequency and distribution tend to differ between ancient populations, which can be informative of migration history of males and females, for example [138, 139]. For aDNA studies, mtDNA has the advantage of being present in a number of copies per cell, whereas the nuclear DNA has a single copy per cell. Hence, mtDNA is more readily available for extraction and subsequent sequencing than nuclear DNA.

SNPs are genetic markers that present variation at single base pair positions, where there is at least one alternative allele in the population. Since mutations occur rarely in most organisms, it is generally reasonable to assume that there are only two alleles at a single SNP locus (i.e. in most cases, only two of the bases A, T, G or C exist at a certain position in the genome) for the purpose of data analysis.

In studies where the complete genome sequence of an individual is generated, the near complete set of SNPs present is captured. Sequencing errors can be mistaken for rare biological variants [140] which is more relevant in cases where ancient populations are represented by a single or a few individuals, and particularly for low coverage data. In most cases however, it is the patterns of allele sharing between individuals that are informative of population structure, and patterns of rare alleles sharing are particularly powerful for inferring fine-scale population structure [141, 137].

7.5.2 Exploring population structure

A common first step in aDNA studies (after attesting data authenticity and assigning uniparental markers) is to apply exploratory, non-parametric, unsupervised methods, to first investigate population structure. This is possible with methods that subdivide genetic variation resulting from differences in allele frequencies between groups [142]. Principal component analysis and clustering-based approaches are usually applied to obtain an overview of genetic affinities between individuals. These methods usually provide an overview of the genetic dataset and a starting point for further hypothesis testing, but one must beware of the risk of over-interpretation. As any other method, PCA and model-based clustering have limitations, the most relevant for aDNA studies possibly being driven by poor sample quality and consequent scarcity of genetic information. Such samples usually yield a shallow overlap with the SNP set in analysis; the implications are typically not related to wrong assignment or position of these samples but rather related to higher uncertainty in their placement in the PC space (with a tendency to be drawn towards the plot's origin) and in the estimated ancestry proportions [7]. A cut-off of ca. 10.000 SNPs covered for inclusion of a sample in such analyses has become standardized in palaeogenomic studies.

Principal component analysis (PCA)

PCA is a statistical tool that summarizes high-dimensional data (such as the information encoded in thousands of genetic markers) by reducing it to uncorrelated major axes of variation, the principal components (PCs). Usually, the first PCs are used to visualize the genetic relationships between individuals, because these explain most of the variance in the data. The output is usually a two-dimensional plot, corresponding to PC1 and PC2 (which captures variance orthogonally to PC1) (Figure 7.2A). Individuals that are genetically more similar tend to group closer together in the PCA space. PCA is useful to determine study design, identify individuals and populations and draw hypothesis for origins, evolution, dispersion and relatedness. Data from ancient individuals will often show large amounts of missing data, and biases or false patterns introduced by aDNA damage, contamination and pseudohaploidization of the data are also a concern. Often, ancient data will not correspond to the complete set of genetic markers analysed in the PCA due to low coverage, or sometimes due to different processing methods, such as the use of different genotyping arrays [143]. This is why it became common to use a variation of PCA for aDNA studies where the PCs are computed on a set of modern-day reference populations with little to no missing data and the ancient individuals' variation is then projected onto the axes defined by modern populations' genetic variation. This minimizes PCA issues that can be particularly critical for aDNA studies, such as unbalanced sample sizes and the presence of outliers, which can influence the variance of the data. Other limitations that must be accounted for are, for example, that isolation-by-distance gradients can be misleadingly interpreted as specific migration events, which has been shown by integrating spatial covariance data in the analysis [144]. SNPs under strong selection may also influence the layout of the PCA.

Model-based clustering approaches

Alongside PCA, model-based clustering has become a popular approach to visualise and explore genetic ancestry patterns. Models based on Bayesian algorithms (implemented in e.g. STRUCTURE [145] and ADMIXTURE [146]) can use datasets with hundreds of thousands of genetic markers to define clusters and assign individuals to them. Genetic similarity is used to assign individuals under a model of K components. The number of clusters (K) is user-defined. Individuals' ancestry fractions are assigned according to the number of K, and the results are frequently visualised as an intuitive and colorful barplot that paints the proportion of each cluster per individual, corresponding to the individual's ancestry fractions (Figure 7.2B). It is common to identify the number of K that statistically fits the data best; however, it is counterproductive and often misleading to assume that there is a true value or to try to assign the best fitting number of K to true biological populations [147]. The underlying model is not assuming an explicit historical scenario or that all populations have radiated from a single ancestral group and subsequently mixed, which is anyway unrealistic especially when working with individuals/populations from a range of chronologies. In fact, several demographic scenarios, such as recent admixture, ghost (unsampled) population

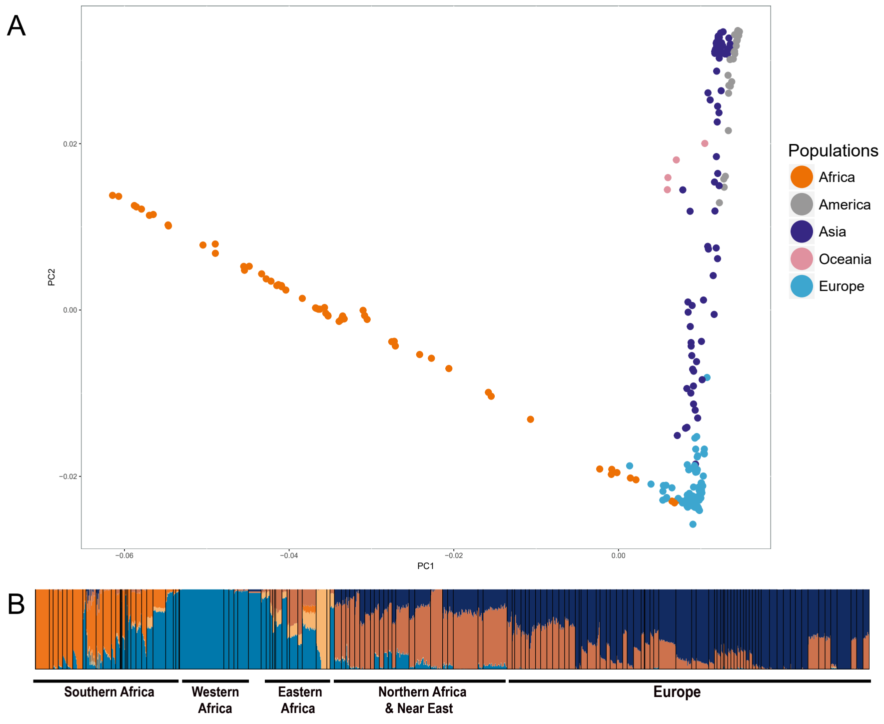

Figure 7.2. Examples of the possibly most commonly used data analyses to visualize population structure in population genetics studies and in archaeogenomics. (A) A PCA of a global dataset of modern populations. (B) An ADMIXTURE plot of that same dataset, with continental and sub-continental population labels, and an ancient individual of African descent showing the same component proportion as found in modern western Africans (unlabeled). Adapted from Paper IV.

admixture and recent bottlenecks can produce indistinguishable results. Like PCA, model based clustering is highly sensitive to sample sizes. Other factors, such as strong recent population-specific drift can lead to over-estimating genetic differentiation [147].

7.5.3 Ancestry inferences

f-statistics

The suite of f-statistics tests has become increasingly used for the ancestry characterization of ancient individuals and populations, and provides sensitive statistical inference for admixture [148, 61, 128]. f-statistics allow formal tests of population mixture by fitting tree-based models to genetic data, namely by measuring allele frequency correlations between populations. The shared genetic drift between sets of (most commonly) three (f_3) and four (f_4)

populations is measured. A block-jackknife resampling is used for estimating standard errors.

The (f_3)-statistics test has typically two major applications: to test if a target population is admixed and to measure shared drift between two populations after their split from a common outgroup. In the first case, an unrooted tree topology of the form (A,B;C) is tested with two source populations (A and B) and a tentative admixed population (C). In the second case, the same topology is applied but with an outgroup (C) and two more closely related populations (A and B). The test statistics is computed as the product of allele frequency differences between population C to A and B, respectively:

$$f_3(A,B;C) = \{(c-a)(c-b)\}$$

Where {} denote the average over all markers and *a,b* and *c* denote the allele frequency for a given site in the three populations A, B and C.

For admixture testing, f_3-statistics can be negative if population C is indeed admixed because the the allele frequency of C is on average (and in many positions) intermediate between the allele frequencies of A and B. However, if population C has undergone stronh genetic drift after the admixture event (such as a severe founder event), it can mask the signal and the test result might not be negative [128]. In *f_3-outgroup*, the test becomes a similarity measure of how closely related A and B are in relation to an outgroup. It focuses the shared branch length between the outgroup C and the internal population node (A,B) [149].

The f_4-statistics test is also used for admixture testing. Similarly to the *D*-statistics (or ABBA-BABA) tests [61, 128], a four population tree topology is tested, with the topology (A,B;C,D). Instead of counting instances of ABBA and BABA topologies, (f_4)-statistics is defined in terms of allele frequency correlations. The statistics is defined:

$$f_4(A,B;C,D) = \{(a-b)(c-d)\}$$

Where {} denote the average over all markers and *a, b, c* and *d* denote the allele frequency for a given locus in the four populations A, B, C and D.

Under an unrooted population model without any admixture, the allele frequency difference between A and B should be independent from the allele frequency difference between C and D, resulting in a f_4 that approaches zero. However, and assuming that the general tree topology is correct, the statistics will significantly deviate from zero in case of gene flow from C or D into A or B. Significantly negative results imply gene flow between either C and B, or D and A; significantly positive results indicate gene flow between A and C, or B and D. To refine the test results and sift the interpretation, population A is usually set as a divergent outgroup with no history of admixture with the remaining populations. This way, a negative result is interpreted as admixture between B and C, whereas a positive statistics will be the result of admixture between B and D.

qpAdm* and *qpGraph

qpAdm has been used in aDNA studies to quantify admixture proportions, after gene flow has been tested with f-statistics. This method has the advantage of not requiring explicit modelling of previous historical relationships between all populations included in the analysis [88]; qpAdm exploits the patterns of shared genetic drift between various populations in order to statistically infer if a population can be modelled as descending from a common ancestor of one or more source populations, and estimates the corresponding proportions of each source. Models are deemed implausible if their estimated admixture proportions fall outside the biologically relevant range (0-1) or if they are rejected statistically by having a small P-value. Based on f-statistics, qpAdm relies on a list of user-specified target and source populations and a list of additional reference populations which provide information about the relationships among the target and source populations [88, 150]. The method is robust to low coverage, missingness and pseudo-haploid data, but less effective when co-analyzing ancient and present-day data, using an extremely large number of reference populations or analyzing population histories involving extended periods of gene flow [150].

Admixture graphs are a way of representing the correlations between allele frequencies of several individuals or populations. They can be fitted once the general genetic structure between the studied individuals, groups or populations has been observed, to investigate their underlying historical relationships. qpGraph is possibly the most implemented admixture graph software in recent aDNA studies. It relies on an user-defined topology of the admixture graph to reconstruct the genetic relationships between different groups. The best-fitting admixture proportions and branch lengths are calculated based on estimated f_2, f_3 and f_4-statistics, providing mixture proportions and branch lengths [128]. The proposed models can thus be favoured or rejected.

The recently developed function find_graphs() (implemented in *ADMIXTOOLS2*) systematically searches for admixture graphs comprising a given set of individuals or populations (including an outgroup) that are compatible with the observed f-statistics. This has the major advantage of exploring a range of graphs beyond those based on previous indications of splits and admixture, instead of leaving a vast majority of alternative models unexplored [151]. It starts with a random admixture graph, and continuously improves it by making random and targeted changes to the graph. The number of admixture events is user defined and prior knowledge can be incorporated by constraining the graph search or by defining the starting graph.

7.5.4 Admixture dating

Admixture between two genetically differentiated populations generates LD patterns that are characterized by previously unshared polymorphism in LD with each of the diverged source ancestries. Such admixture LD patterns extend longer than LD in non recently admixed populations and decay with genetic distance and over time due to recombination: ancestry-specific segments will become shorter as the number of generations since admixture increases. Methods such as ALDER [152] use simple, single-pulse admixture model between two source populations to infer time since admixture on a targeted population based on the fitting of exponential LD decay curves. A more recent method (DATES [153, 154]), suited for small sample sized, low coverage ancient genomic data, has been shown to outperform LD-decay based methods, which in turn rely on measuring the extent of the allelic correlation across markers to infer the time of admixture. DATES models the ancestry covariance patterns of an admixed individual as the combined likelihood of ancestries from populations A and B, even if these are closely related to each other or if highly divergent surrogates of the true source populations are used [154].

7.5.5 Describing genetic diversity with diploid ancient data

The methods described thus far have a wide application in aDNA studies as they are suited for pseudo-haploid data. Pseudo-haploid genotype calling requires a minimum coverage of a single sequencing read at a given locus. If there is higher coverage at a certain SNP, one read is usually randomly drawn and that locus is assumed to be hemizygous, or haploid, for that allele. This can be strongly affected by reference bias and result in underestimation of the genomic variation on the individual- and population-level (which may subsequently bias the measurements of shared genetic drift in, e.g. f-statistics). However, this effect is diluted over the large number of loci analysed and by allowing the analysis of low coverage data, this method results in the inclusion of more individuals per population, conferring statistical power to the analyses.

When high coverage data is available, confident diploid calls can be made. In this case, a high number of reads covering a given locus provides confidence to identify it as heterozygous (carrying different alleles), homozygous reference or homozygous alternative (carrying the same allele on both maternal and paternal chromosomes, comparing to the reference genome sequence). This type of data allows for the application of analyses - particularly at the individual level - that are not possible with pseudo-haploid calls. Haplotype-based methods perform chromosome painting or local ancestry inference at each locus in the genome. This requires accurate phasing and inference of local ancestry, which is harder the older the samples are: molecular phasing is impossible given the fragmented nature of aDNA molecules and statistical phasing relies on a good reference dataset for haplotype inference on the ancient individuals. When studying ancient populations whose ancestry is not well represented in modern populations, or for which admixture has occurred long ago so haplotypes were effectively broken down by recombination, there is a higher risk of introducing biases. This is the case for Stone Age hunter-gatherers (the main study object of this thesis) and the reason why such methods were not attempted.

Heterozygosity and Homozygosity

Heterozygosity can be determined as the fraction of observed heterozygote sites in an individual or sample. Assuming that the parents of the individual in study were not related (and belonged to the same population), heterozygosity is often informative about past demographic events, as these impact a population genetic diversity: for example, bottlenecks tend to reduce genetic variation while admixture tends to increase it.

Homozygosity also holds information about a population's history. Runs of Homozygosity (RoH) are regions of the genome where an individual is homozygous across all sites, showing no variation between both copies of that part of the genome. Such stretches appear when an individual receives two identical copies of an ancestral haplotype from each parent. Combined, the number and the length of RoH are indicators of genetic diversity [155], and can be particularly informative about inbreeding. The burden of ROH relates to the size of the population: populations with long-standing low effective size having more and longer ROH than larger populations. Bottlenecks increase the number of ROH, but these usually remain relatively short (Figure 4.1). Admixture introduces different haplotypes, and recombination acts at breaking down long RoH stretches and reducing the number of RoH. Recent consanguinity has the effect of adding very long ROH, thus also increasing the variance in the sum of ROH [156].

8. Research aims

The main objective of this thesis is to investigate patterns of genomic diversity of pre-historic human populations from understudied periods and places, with a focus on the Mesolithic and Neolithic periods in the western Mediterranean region. To do so, I generated genomic data from arcaheological human remains originating from that region. This thesis central goals are to improve our understanding of the genetic diversity patterns prior to the advent of the Neolithic both in Western Europe and Northern Africa, and to study the impact of the Neolithic transition on extant human genetic variation in North Africa.

Specific aims included:

I. Generating whole genome sequencing data (including high coverage) to characterize the demographic history of hunter-gatherer and farming populations of western Europe and northern Africa;

II. Improving the understanding of the demographic events that shaped the human genetic variation in underrepresented pre-historic populations, such as those of North Africa;

III. Dating and quantifying the extent and direction of contacts across the Mediterranean Sea;

IV. Integrating genomic, isotopic and archaeological data to generate a more comprehensive knowledge of the demic and socio-cultural dynamics of past populations;

V. Investigating the socio-cultural and behavioural patterns that shaped the genetic diversity patterns of Mesolithic hunter-gatherer communities in western Europe and the impact of the first contacts with incoming farming populations;

VI. Understanding the role of Ice Age refugia, such as Iberia, on the patterns of genetic diversity of Mesolithic hunter-gatherer communities in western Europe;

VII. Investigating the use of pre-historic burial grounds during historic times.

9. Summary of results

9.1 Patterns and processes of West European hunter-gatherers' genetic diversity (Papers I, II and III)

Even considering the bias in archaeogenomics research for European sites, compared to the rest of the world, pre-Neolithic populations are still understudied. This is possibly tied to the higher density and better preservation of archaeological sites (and corresponding human remains) in later periods. Single studies have yielded genome-wide data for dozens or even hundreds of individuals from European contexts [104, 157, 95] and while the vast majority dates to the Neolithic and later periods, usually only a few predate the Neolithic. Until very recently, the Mesolithic Western European hunter-gatherers were thought to belong to a rather homogeneous genetic pool; since 2018 a few studies have highlighted diversified patterns of admixture on geographic and temporal scales.

To address questions surrounding the origins of the pre-Neolithic genetic ancestry in western Europe, we generated whole genome sequencing data for 36 western European hunter-gatherers, from the Iberian Peninsula and northwestern France, with a focus on the late Mesolithic period. We analysed the patterns of admixture between two European Late Pleistocene lineages - one associated with central European Magdalenian hunter-gatherers, such as Goyet Q2 (12 kya from Belgium) and El Mirón (19 kya from Spain), and another with southern European Epigravettian hunter-gatherers, such as Villabruna (14 kya from Italy). To do so, we combined our data with previously published ancient hunter-gatherers from Europe. We find that the proportion of shared ancestry with GoyetQ2 is significantly correlated to both the chronology and the geography of hunter-gatherers from the European Atlantic facade, being more pronounced in older and southernmost individuals. This is consistent with a strong gene flow of Villabruna-related ancestry to western and southwestern Europe during the Mesolithic. Analysing patterns of kinship, as well as heterozygosity and runs of homozygosity in high coverage hunter-gatherers, and integrating those results with stable isotopes data, we find that late Mesolithic hunter-gatherers from coastal regions of western Europe likely had structured social dynamics (Papers II and III). Their social units were not based on familial bonds, as we do not find extensive intra-site kin relationships, but apparently relied on exogamic practices such as intermariage networks to avoid inbreeding. This contributed to the maintenance of relatively high effective population sizes. Such strategies could be rooted in hunter-gatherer practices since the Early Upper Palaeolithic [68] and still prevail in present-day hunter-gatherer populations [158, 159]. We also observe that in spite of confirmed chronological overlap with the arrival of Neolithic farmers, the admixture between these groups was not bidirectional, as we do not find farmer ancestry in any of the late Mesolithic hunter-gatherers analysed. The presence of local hunter-gatherer ancestry in Neolithic farmers from this region [160, 161] shows that the interactions between these groups involved unidirectional admixture, from hunter-gatherer into farmer populations.

Some of the analysed individuals from Iberia dated to an older period, the Upper Palaeolithic, ranging from ca. 22.000 to ca. 18.000 years ago (Paper III). To date, the number of European individuals from this period that have been genetically analysed remains very low, and included a single individual from Iberia (El Mirón). Our results indicate a genetic component that was previously unappreciated in Iberia, present at least in the North of the Peninsula during and after the LGM.

9.2 Shell middens: burial grounds from the Mesolithic to the modern Era (Papers II, III and IV)

Shell middens are conspicuous deposits of residues of marine and terrestrial animals gathering, consumption, and disposal, but most visibly sea shells. Over the past 140 ka, people have created these singular and prominent landscape features as a result of their activity [162]. Shell middens can be found in coastal, lacustrine and river environments worldwide, from arctic regions to the tropics, and have been intensely studied by archaeologists, due to their typically good preservation of organic remains. The iconic shell middens of Hoedic and Téviec in southern Brittany, France, and of the Tagus and Sado valleys in Portugal have revealed astonishing burials of some of Europe's last hunter-gatherers. These sites, at the Atlantic facade of Europe, are known for the unusually well-preserved and rich human burials [163].

We retrieved aDNA sequences from hunter-gatherers buried in these sites and revealed that, against longstanding archaeological interpretations, these burials were not based on kin relationships (Papers II and III). That includes multiple burials, which are common across these sites, and even joint burials of adults (namely females) and children, which were previously thought to correspond to mother and child (Paper II). Biological unrelatedness in graves with adult female-child seems to be relatively common in other contexts and periods [164], suggesting that social bonds were established beyond biological kinship and that even after the death of those individuals, those links retained social importance.

In the context of the previous research, we found a surprising outlier: a young adult individual, whose burial in a Portuguese shell midden in Muge, Portugal, was thought to date to the Mesolithic (Paper IV). This outlier was first spotted through aDNA analysis, which revealed that the mt haplogroup was L3b1a, one of the most common L3 subclades in West and Central Africa [165]. We explored multi-disciplinary data integrating genetics, archaeology, radiocarbon dating, carbon, nitrogen and oxygen isotope analyses, and different sources of historical records to uncover the story behind this burial. Genome-wide analysis indicates that this male buried in Portugal sometime between the sixteenth and the eighteenth centuries, was originally from Senegambia, in line with records of shipments of slaves arriving in Portugal during that time. Isotopic data showed that his diet was based on C4 plants and, to some extent, marine foods, consistent with traditional diets on the coastal semi-arid region of Africa where he probably lived for several years before arriving to Portugal. Oxygen stable isotopes data confirmed a coastal region origin. Interestingly, shell middens are actively used in western Africa up to present day, and in Senegambia their uses include ancient and modern cemeteries [166]. We suggest that the burial of this man in an 8000-years old site could be an example of the maintenance of African cultural beliefs and practices by African people translocated to Europe, even though this practice is not documented in the historical records. Our study shows the power of multidisciplinary approaches to investigate individual histories. The different lines of research converged to the same conclusion, conferring reproducibility to the results, which would not be achievable otherwise. We also discuss the continued use of Stone Age mortuary grounds (such as the Mesolithic shell middens) during modern times.

9.3 The role of the Mediterranean in human pre-history (Papers III and V)

Prehistoric contact across the western Mediterranean have been proposed based on observed similarities in material culture or synchronous developments in subsistence strategies [167, 109]. Archaeogenomics studies, however, have to date found no evidence of the crossing of the Mediterranean Sea prior to the Late Neolithic, when European ancestry is found in Northwestern Africa [79]. Even though long distance connections between northwestern African and Levantine hunter-gatherers is observed prior to 15,000 years ago [78], the Mediterranean Sea acted as a barrier to gene flow for millennia. In papers III and V, we generated genomic data that predates the Neolithic transition from Iberia and Northwestern Africa. In line with previous studies, we find no indication of gene flow between these populations. While Late Mesolithic hunter-gatherers from Iberia (paper III) show shared ancestry with central and eastern European hunter-gatherers, an Epipalaeolithic hunter-gatherer from North Africa (oub002, paper V) exhibits a remarkable 8,000 years population continuity and isolation since the Upper Palaeolithic; a distinct group of humans with a unique 'Maghrebi' genetic make-up. With newly generated high coverage genomic data we confirmed that trans-Gibraltar migration across the Mediterranean did not happen until the Neolithic. We show that northwestern Africans went through a severe population bottleneck, similarly to to Eurasian populations, which is consistent with the Maghrebi lineage being related to out-of-Africa populations.

The lifestyle of the people in northwestern Africa rapidly transitioned from foraging to food-production some 7,400 years ago, but recent archaeogenetic data [79] and previous archaeological studies [100] suggested that farming started with local developments of local Maghrebi foragers. We show that human remains from the earliest Neolithic contexts in northwestern Africa had European Neolithic ancestry, in contrast to other farming groups that display exclusively local Maghrebi ancestry. This result leads to the conclusion that farming was introduced by Iberian migrants. The European newcomers brought new ways of life, agriculture, domesticates and pottery traditions that were subsequently adopted by local populations. This makes North Africa Neolithization uniquely different from most other regions: local hunter-gatherers did persist as a genetically differentiated group instead of being replaced or integrated into incoming farming communities. Furthermore, we identified gene flow across the Mediterranean Sea more than 2,000 years earlier than it was previously observed.

We also generated new genomic data from the Maghrebi Middle Neolithic and identified the introduction of new ancestry, from the Levant. This ancestry, which has not been observed on the European side of the Mediterranean, is identified in association with a new ceramic tradition in northern Morocco and a pastoralist lifestyle, as a consequence of the expansion of cattle pastoralism in the current Sahara territory [168]. Levantine gene flow has been previously identified in pastoralist Neolithic East African individuals living around 4,000 years ago [169].

Interestingly, gene flow from North Africa into Europe has only been detected much later, around 4000 years ago [104, 170]. The first available evidence of gene flow from North Africa into Iberia is the male individual I4246 from the Copper Age site of Camino de las Yeseras (Madrid, Central Iberia) [104]. Dated ca. 4200 Cal BP, this finding fits with the Bell Beaker Horizon in Iberia, when North African importations are documented (e.g. elephant ivory tusk, ostrich egg shells). At that same time, a clear Iberian influence in Morocco is noted (with e.g. Beaker vessels of Andalusian type, copper daggers and arrowheads of Iberian typology). The North African gene flow into Iberia is confirmed in later individuals dated more recently than 4000 cal BP: COV20126 from Sima del Angel (Cordoba, Southern Iberia) [171], and I7162 from Loma del Puerco (Cadiz, Southern Iberia) [104]. A predictable increase in North African ancestry much later, during the Iron Age would, however, be related to Punic influence around ca. 2500 cal BP.

10. Conclusions and future prospects

Prior to the start of this dissertation, ancient genomic data from Late Mesolithic hunter-gatherers from Iberia was scarce and no aDNA study had been done on North African prehistory. During the past years, a few of studies have gathered pre-Neolithic genomic data from the west Mediterranean region [78, 79, 104, 160]. Geographic, preservational, socio-economic and cultural biases are likely to have an influence on the higher amount of aDNA (and as a likely cause, archaeological) studies for later periods and higher latitudes. In practice, this is reflected on a much better and more detailed understanding of the demographic processes underlying the Neolithisation of other parts of Europe, compared to the Neolithic transition and previous periods in Mediterranean regions. I expect that in the future, archaeogenomic studies focusing on underrepresented areas of the world and older periods become abundant, encouraged by technological advances in the field.

With this work, I aimed at contributing to narrowing that gap and showed that, even within that relatively small, regional scale, there is diversity in the demographic processes that shaped human prehistory, but also much of a common evolutionary, adaptive and cultural path. This thesis contributes to understanding the evolutionary history of hunter-gatherers from the European and North African sides of the Mediterranean: with the analysis of high coverage genomic data from individuals who lived more than 7,000 years ago, we were able to demonstrate that a common history is shared until ca. 60.000 years ago. This is broadly coincident with the estimates of the out-of-Africa migration events, which resulted in a major decrease in genetic diversity, compared to sub-Saharan African populations, due to a decrease in population size [44]. However, the population split between Western Mediterranean shores was deep and maintained for millennia by the Mediterranean Sea: in spite of being separated by only ca. 13km, the evidence presented in this thesis verifies that the North African and Iberian Palaeolithic and Mesolithic populations did not mix. This strong barrier to gene flow was only overcome by human resourcefulness in the Neolithic. The transition from foraging to farming is one of the most drastic shifts visible in the archaeological record, and arcaheogenomic research has shown the importance of the migration of human populations to that revolution. Here, I show that the Neolithisation process of northwestern Africa and Europe shows some similarities - that migrants introduced new ways of life - but with unique and outstandingly different characteristics too - that local foragers adopted farming practices and coexisted with incoming farmer communities in North Africa. We were able to elucidate long standing archaeological questions about the Neolithic transition in this region, but a lot more are left to be answered. Did the last hunter-gatherers in Europe and North Africa resist to the incoming farmers? Was the mixing process different in both sides of the Mediterranean? Perhaps future studies will be able to further elucidate the role of human behaviour in such transitions.

A big part of my work in this thesis relied heavily on the integration of archaeological, stable isotopes and genetics data. The importance of multidisciplinary collaborations to understand human history is highlighted in papers II and IV, where different lines of investigation bring in complementary information that only when tied together, allow enhancing and magnifying the studies' outcome. The nature of archaeogenetics is, in its genesis, cross-disciplinary. It is also still a recent field of research, that draws together lines of investigation with different traditions, that "speak different languages", and that are not frequently taught together. The fact is that genetics can only assess biological and demographic processes. Links between genetics and culture can be hypothesised but should not presumed. Given that the first aDNA studies focused on a few individuals representing major population changes which were associated with strong cultural transitions, the usage of cultural terms to identify biological populations became common. This was a prevalent drawback in some of the papers in this thesis and as a team, we find it pertinent to try to abandon ambiguous lingo and instead adopt clearer terminology. For example, population replacement as interpreted by archaeologists or general public may imply a certain level of intention and/or violence, which is not what geneticists mean. The successful development of aDNA as a field implies, in my opinion, the establishment of effective and dynamic collaborations, which I was fortunate to be a part of.

The multidisciplinary work allowed me to shed light on social aspects from very different time points: the Mesolithic and the slave trade period. We show that the social dynamics of European Mesolithic hunter-gatherers had a strong impact on their demographic history and ultimately their survival as a population. Within Neanderthal populations, for example, inbreeding has been genetically demonstrated [62] and pointed as having an important role in the collapse of their lineage [172, 173]. Avoiding inbreeding depression may thus have played a role on the maintenance of hunter-gatherer populations, before the onset of the Neolithic. On a much different scale and time period, we gathered lines of evidence that point to the maintenance of some native socio-cultural practices by Africans who were brutally dislocated from their homelands and families during the slave trade period, while forced to adopt new religions, languages and social conducts.

In recent years, a new focus on finer-scale processes has gained momentum in the aDNA field of research. A more nuanced human evolution history will continue to be revealed in the future, as complex processes are better understood, with resources to ever improving data quality, geographic and chronological coverage. Environmental DNA has uncovered million year-old forest

[174]. The oldest human aDNA from African human remains dates to about 15,000 years ago [78]. In Europe, nuclear DNA from hominins dating to the Pleistoce has been retrieved [175, 176]. Genome-wide data from mammoths more than one million years old is now available [177]. Who knows what the future of aDNA research holds?

11. Svensk sammanfattning

Genetiska analyser av forntida mänskliga kvarlevor har revolutionerat hur vi studerar mänsklighetens historia. Upptäckten att DNA finns bevarat i ben och tänder har öppnat ett direkt fönster till att undersöka demografiska och evolutionära händelser som format vår arts förflutna. Forskningen om forntida DNA har trots flera utmaningar genomgått en snabb utveckling under de senaste två decennierna, där flera av utmaningarna har vänts till möjligheter. I takt med att forskningsfältet mognat har nya molekylära, bioinformatiska, teoretiska och analytiska verktyg utvecklats för att kunna analysera forntida DNA. I min avhandling har jag genererat och analyserat forntida DNA-data från västra medelhavsregionen, som är ett område med stora skiftningar i livsmiljöner under människans förhistoria, men som också är ett område där klimatförhållandena inte är idealiska för DNA-bevarande i arkeologiska lämningar.

Den senaste istiden, med ett maximum mellan 26500-19000 år sedan, resulterade i dramatiska förändringar i den mänsklig demografin. Medelhavsregionen var isfri under den här perioden och människor kunde leva här, medan norra Europa var täckt av inlandsisen. Slutet på istiden förde med sig ett mer fördelaktigt klimat, vilket senare möjliggjorde utvecklingen av jordbruket och fasta bosättningar i Mellanöstern. Den processen är känd som den neolitiska revolutionen (perioden kallas ofta yngre stenåldern).Studier av forntida DNA har löst en långvarig frågeställning: att övergången från samlande och jakt till jordbruk spreds över Europa genom migration från Anatolien, och inte via överföring av idéer och tekniker mellan folk. I Nordafrika finns det dock indikationer på att övergången till Neolitikum skedde genom att lokala jägare-samlare anammade jordbruket. Med hjälp av genomisk data från forntida individer, daterade till både före och efter den neolitiska revolutionen i nordvästra Afrika, visar jag att denna övergång startade med ankomsten av bönder från Iberiska halvön, och att deras levnadssätt spred sig till nordafrikanska grupper. Jag visar också att en annan migrationsvåg några tusen år senare, med ursprung i mellanöstern, spred pastoralismen till nordvästra Afrika.

Den neolitiska revolutionen är en av de mest studerade perioderna i förhistorien, men längre tillbaka i tiden är det svårare att hitta välbevarade arkeologiska lämningar. Detta har fått till följd att det från de perioder som föregick yngre stenåldern – mesolitikum och ännu längre tillbaka paleolitikum – finns mycket begränsat med genomiska data. I min avhandling genererade jag helgenomdata för 36 jägare-samlare från Iberiska halvön och den franska Atlantkusten. Jag fann två genetiskt distinkta grupper med rötter i paleolitikum, som båda bidrog till grupperna som levde under mesolitikum. Baserat på genetiska data från individer begravda i samma grav eller gravplatser kan jag dra slutsatsen att mesolitiska jägare-samlare bildade sociala enheter som inte var baserade på familjeband och som förlitade sig på utbyte mellan grupper för att undvika inavel. Jag kan också visa att de sista jägare-samlarna inte genetiskt beblandade sig med de första bönderna som levde i samma område under några hundra år. Men genflödet gick i motsatt riktningen, då genetiskt material från jägare-samlarna finns hos de första bönderna.

I studien av de iberiska jägare- och samlarpopulationerna hittades en ovanlig individ som visade sig vara av afrikansk härkomst. Genom att använda ett tvärvetenskapligt tillvägagångssätt, som innefattade genetiska analyser, arkeologi, kol-14 datering, analys av stabila isotoper och historiska dokument, kunde vi avslöja historien bakom denna gravsättning. Individen visade sig ursprungligen komma från kustregionen vid Senegal och Gambia och anlände förmodligen till Portugal genom slavhandel någon gång mellan 1600- och 1800-talet. Det är möjligt att begravning av denna man i en 8000 år gammal kökkenmödding (avskrädeshög för snäckskal) skulle kunna vara ett exempel på upprätthållandet av kulturella seder och bruk hos individer som flyttats till Europa då ökkenmöddingar har används som begravningsplats i västra Afrika under flera hundra år och fram till idag.

Min avhandling lyfter inte bara fram potentialen med DNA-analyser av forntida lämningar, utan den belyser också vikten av ett tvärvetenskapligt tillvägagångssätt för att kartlägga människans historia.

12. Acknowledgements

When in 2016 I took the "PhD road", I knew it was going to be long and bumpy, but I did not exactly expect the bumps where they have (literally and figuratively) been. I owe getting to the finishing line to a lot of people. First of all, to my supervisors.

Mattias, I remember our first skype conversation, when I was freshly in love with ancient DNA and told you I would like to do my PhD in this field. I am glad you took me as your student! I am thankful for our insightful discussions, your enthusiasm for new findings, your experienced and sharp text revising skills and I must add, for your talent to write successful grant proposals. Thank you for keeping your door open and for dynamizing a great lab.

Torsten, I don't know where to begin. Thank you for taking this duck under your falcon wing and putting up with my lack of abilities. For stepping in when I most needed. I feel privileged to have seen - and learned from - your dedication and the way you do science. For all the encouragement, all the answers and help, for ALWAYS listening and giving your wise and pondered perspective, for being honest, fair, reliable and brilliant, both as a scientist and as a friend. Your support meant (means!) so much to me.

Emma, thank you for supporting me, for being my scout and bringing me back to aDNA and Sweden when I thought science was over for me.

Cris, long distance relationships are not always easy, but we made it through! Thank you for trusting me with the North African samples (and more). I am grateful to have had the chance to work in our amazing projects.

Rita, even though you were not officially my supervisor, you often took that role, and at the same time we grew together as scientists. It has been great to share these years and these projects with you, working so closely together, observing your archaeological perspective, your methodical and organized ways. I am grateful for being your side kick all this time.

To all members of the Human Evolution lab, old and new - thank you! From all of you I learned about how to be a better scientist and a better (evolving) human. Special thanks to all the students that shared this PhD path with me!! Mario, minha bruxa, I'll love you forever, in spite of your ability to annoy me. Gwenna, I am so grateful we ended up sitting in the same office and became close; thank you for all the support, for loving my baby, for your dear friendship. TJ I miss our hour long conversations late in the afternoon (when we should be working instead)! Nina, I was so lucky to have you as a colleague and a friend during and after your PhD. James and Rickard, the last few months have been quite a ride for the three of us! I am proud and happy to have done this by your side. Ayça, I admire and love you so much. Hanna and Helena, thank you for always caring, for your empathy. Arielle, thank you, you were a guiding light when Uppmax was so hard. Fede, cariño, I am so grateful for your support and guidance. But for your friendship, especially. You were (are!) invaluable to me and I am sure I can say, to the lab as well.

Ernst and Darshan: trying to teach you was a great way to learn! It is wonderful to see you thrive!

To the wonderful collaborators I was lucky to work with, thank you for making me grow as a scientist. Especially: Gracias Rafa and Carlos, more than once your enthusiasm renewed my will. Thank you Grégor, I look forward to the day we get to meet in person! Carro, I am grateful for all your help, with the Meso France and more.

Catarina Ginja, who trusted and trained me, and supported me all the way. Thank you!

To many Uppsala friends, inside and outside of EBC, THANK YOU for making life in Sweden better and fun!

Maja, Shaun and Björn, dear chosen family. You literally saved me in my most difficult times and I just can't express how happy and thankful I am for having you guys in my life. I truly hope our paths never EVER diverge too far and that we get to see our little monkeys grow together. I love you!

Cris, Manau, Tiago, Luis, que estiveram sempre perto e presentes, apesar da distância. Não podem imaginar o quanto significa para nós. Gosto muito de vocês!

Família, que me apoiou sempre, e à minha mãe quando eu nao pude. Obrigada a todos!

João amor, Obrigada. Por vires comigo, por seres a minha casa. Pelo apoio incondicional que me permitiu cumprir este sonho e objectivo sem que a vida o levasse. Por 18 (!) anos de amizade, 11 anos de amor, 2 anos de filha e uns meses de peixinho e tudo o que és e fazes por nós. Nem sei expressar por palavras o que sinto e os olhos enchem-se de lagrimas de felicidade ao escrever isto: é tão bom ser tua amiga mas igualmente bom ser tua amante :)

Clara filha, espero que um dia possas ler isto e sentir-te orgulhosa de mim. Ensinas-me tanto. Foram muitas as vezes que este trabalho me tirou tempo que eu preferia ter passado contigo <3 E Gaspar, que mudaste a perspectiva, minha companhia constante neste esforço final, mal posso esperar por ti!

Mãe querida, tantas vezes ao longo destes anos quis estar perto de ti e não pude. Este livro não compensa esses momentos, mas ainda assim, espero que seja uma fonte de orgulho para ti. Tenho muita pena que o pai não o possa ler, de não lhe poder agradecer o incentivo para que me decidisse a envergar por este caminho. Mas agradeço-te a ti, por isso e por muito mais; por tudo o que fizeste e fazes por mim desde sempre, por seres a minha fã numero 1, a melhor pessoa que eu conheço, pelo teu amor incondicional.

References

[1] M. Slatkin and F. Racimo, “Ancient dna and human history,” *Proceedings of the National Academy of Sciences*, vol. 113, no. 23, pp. 6380–6387, 2016.

[2] P. Skoglund and I. Mathieson, “Ancient genomics of modern humans: the first decade,” *Annu. Rev. Genomics Hum. Genet*, vol. 19, pp. 381–404, 2018.

[3] I. Olalde and C. Posth, “Latest trends in archaeogenetic research of west eurasians,” *Current Opinion in Genetics & Development*, vol. 62, pp. 36–43, 2020.

[4] R. Higuchi, B. Bowman, M. Freiberger, O. A. Ryder, and A. C. Wilson, “Dna sequences from the quagga, an extinct member of the horse family,” *Nature*, vol. 312, no. 5991, p. 282, 1984.

[5] S. Pääbo, “Molecular cloning of ancient egyptian mummy dna,” *nature*, vol. 314, no. 6012, pp. 644–645, 1985.

[6] M. J. Collins, C. M. Nielsen-Marsh, J. Hiller, C. Smith, J. Roberts, R. Prigodich, T. J. Wess, J. Csapo, A. R. Millard, and G. Turner-Walker, “The survival of organic matter in bone: a review,” *Archaeometry*, vol. 44, no. 3, pp. 383–394, 2002.

[7] M. E. Allentoft, M. Collins, D. Harker, J. Haile, C. L. Oskam, M. L. Hale, P. F. Campos, J. A. Samaniego, M. T. P. Gilbert, E. Willerslev *et al.*, “The half-life of dna in bone: measuring decay kinetics in 158 dated fossils,” *Proceedings of the Royal Society B: Biological Sciences*, vol. 279, no. 1748, pp. 4724–4733, 2012.

[8] C. Gamba, K. Hanghøj, C. Gaunitz, A. H. Alfarhan, S. A. Alquraishi, K. A. Al-Rasheid, D. G. Bradley, and L. Orlando, “Comparing the performance of three ancient dna extraction methods for high-throughput sequencing,” *Molecular Ecology Resources*, vol. 16, no. 2, pp. 459–469, 2016.

[9] P. B. Damgaard, A. Margaryan, H. Schroeder, L. Orlando, E. Willerslev, and M. E. Allentoft, “Improving access to endogenous dna in ancient bones and teeth,” *Scientific reports*, vol. 5, p. 11184, 2015.

[10] E. Svensson, T. Günther, A. Hoischen, M. Hervella, A. R. Munters, M. Ioana, F. Ridiche, H. Edlund, R. C. Van Deuren, A. Soficaru *et al.*, “Genome of peştera muierii skull shows high diversity and low mutational load in pre-glacial europe,” *Current Biology*, vol. 31, no. 14, pp. 2973–2983, 2021.

[11] P. Korlević, T. Gerber, M.-T. Gansauge, M. Hajdinjak, S. Nagel, A. Aximu-Petri, and M. Meyer, “Reducing microbial and human contamination in dna extractions from ancient bones and teeth,” *Biotechniques*, vol. 59, no. 2, pp. 87–93, 2015.

[12] J. Dabney, M. Meyer, and S. Pääbo, “Ancient dna damage,” *Cold Spring Harbor perspectives in biology*, vol. 5, no. 7, p. a012567, 2013.

[13] D. Higgins and J. J. Austin, “Teeth as a source of dna for forensic identification of human remains: a review,” *Science & Justice*, vol. 53, no. 4, pp. 433–441, 2013.

[14] C. J. Adler, W. Haak, D. Donlon, A. Cooper, G. Consortium *et al.*, "Survival and recovery of dna from ancient teeth and bones," *Journal of Archaeological Science*, vol. 38, no. 5, pp. 956–964, 2011.

[15] R. Pinhasi, D. Fernandes, K. Sirak, M. Novak, S. Connell, S. Alpaslan-Roodenberg, F. Gerritsen, V. Moiseyev, A. Gromov, P. Raczky *et al.*, "Optimal ancient dna yields from the inner ear part of the human petrous bone," *PloS one*, vol. 10, no. 6, p. e0129102, 2015.

[16] H. B. Hansen, P. B. Damgaard, A. Margaryan, J. Stenderup, N. Lynnerup, E. Willerslev, and M. E. Allentoft, "Comparing ancient dna preservation in petrous bone and tooth cementum," *PloS one*, vol. 12, no. 1, p. e0170940, 2017.

[17] L. Orlando, R. Allaby, P. Skoglund, C. Der Sarkissian, P. W. Stockhammer, M. C. Ávila-Arcos, Q. Fu, J. Krause, E. Willerslev, A. C. Stone *et al.*, "Ancient dna analysis," *Nature Reviews Methods Primers*, vol. 1, no. 1, pp. 1–26, 2021.

[18] A. W. Edwards, "Human genetic diversity: Lewontin's fallacy," *BioEssays*, vol. 25, no. 8, pp. 798–801, 2003.

[19] M. Meyer and M. Kircher, "Illumina sequencing library preparation for highly multiplexed target capture and sequencing," *Cold Spring Harbor Protocols*, vol. 2010, no. 6, pp. pdb–prot5448, 2010.

[20] B. Llamas, G. Valverde, L. Fehren-Schmitz, L. S. Weyrich, A. Cooper, and W. Haak, "From the field to the laboratory: Controlling dna contamination in human ancient dna research in the high-throughput sequencing era," *STAR: Science & Technology of Archaeological Research*, vol. 3, no. 1, pp. 1–14, 2017.

[21] A. W. Briggs, J. M. Good, R. E. Green, J. Krause, T. Maricic, U. Stenzel, C. Lalueza-Fox, P. Rudan, D. Brajković, Ž. Kućan *et al.*, "Targeted retrieval and analysis of five neandertal mtdna genomes," *Science*, vol. 325, no. 5938, pp. 318–321, 2009.

[22] Q. Fu, M. Meyer, X. Gao, U. Stenzel, H. A. Burbano, J. Kelso, and S. Pääbo, "Dna analysis of an early modern human from tianyuan cave, china," *Proceedings of the National Academy of Sciences*, vol. 110, no. 6, pp. 2223–2227, 2013.

[23] M. L. Carpenter, J. D. Buenrostro, C. Valdiosera, H. Schroeder, M. E. Allentoft, M. Sikora, M. Rasmussen, S. Gravel, S. Guillén, G. Nekhrizov *et al.*, "Pulling out the 1%: whole-genome capture for the targeted enrichment of ancient dna sequencing libraries," *The American Journal of Human Genetics*, vol. 93, no. 5, pp. 852–864, 2013.

[24] T. Lindahl, "Instability and decay of the primary structure of dna," *nature*, vol. 362, no. 6422, pp. 709–715, 1993.

[25] A. W. Briggs, U. Stenzel, P. L. Johnson, R. E. Green, J. Kelso, K. Prüfer, M. Meyer, J. Krause, M. T. Ronan, M. Lachmann *et al.*, "Patterns of damage in genomic dna sequences from a neandertal," *Proceedings of the National Academy of Sciences*, vol. 104, no. 37, pp. 14 616–14 621, 2007.

[26] M. Hofreiter, D. Serre, H. N. Poinar, M. Kuch, and S. Pääbo, "ancient dna," *Nature Reviews Genetics*, vol. 2, no. 5, p. 353, 2001.

[27] J. Krause, Q. Fu, J. M. Good, B. Viola, M. V. Shunkov, A. P. Derevianko, and S. Pääbo, "The complete mitochondrial dna genome of an unknown hominin

from southern siberia," *Nature*, vol. 464, no. 7290, pp. 894–897, 2010.

[28] P. Skoglund, B. H. Northoff, M. V. Shunkov, A. P. Derevianko, S. Pääbo, J. Krause, and M. Jakobsson, "Separating endogenous ancient dna from modern day contamination in a siberian neandertal," *Proceedings of the National Academy of Sciences*, vol. 111, no. 6, pp. 2229–2234, 2014.

[29] W. Xu, Y. Lin, K. Zhao, H. Li, Y. Tian, J. N. Ngatia, Y. Ma, S. K. Sahu, H. Guo, X. Guo *et al.*, "An efficient pipeline for ancient dna mapping and recovery of endogenous ancient dna from whole-genome sequencing data," *Ecology and evolution*, vol. 11, no. 1, pp. 390–401, 2021.

[30] H. N. Poinar and A. Cooper, "Ancient dna: do it right or not at all," *Science*, vol. 5482, no. 1139, p. 416, 2000.

[31] T. L. Fulton and B. Shapiro, "Setting up an ancient dna laboratory," in *Ancient DNA*. Springer, 2019, pp. 1–13.

[32] A. W. Briggs and P. Heyn, "Preparation of next-generation sequencing libraries from damaged dna," in *Ancient DNA*. Springer, 2012, pp. 143–154.

[33] L. Orlando, M. T. P. Gilbert, and E. Willerslev, "Reconstructing ancient genomes and epigenomes," *Nature Reviews Genetics*, vol. 16, no. 7, p. 395, 2015.

[34] A. W. Briggs, U. Stenzel, M. Meyer, J. Krause, M. Kircher, and S. Pääbo, "Removal of deaminated cytosines and detection of in vivo methylation in ancient dna," *Nucleic acids research*, vol. 38, no. 6, pp. e87–e87, 2010.

[35] N. Rohland, E. Harney, S. Mallick, S. Nordenfelt, and D. Reich, "Partial uracil–dna–glycosylase treatment for screening of ancient dna," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 370, no. 1660, p. 20130624, 2015.

[36] S. Sawyer, J. Krause, K. Guschanski, V. Savolainen, and S. Pääbo, "Temporal patterns of nucleotide misincorporations and dna fragmentation in ancient dna," *PloS one*, vol. 7, no. 3, p. e34131, 2012.

[37] E. M. Wijsman and L. Cavalti-Sforza, "Migration and genetic population structure with special reference to humans," *Annual Review of Ecology and Systematics*, vol. 15, no. 1, pp. 279–301, 1984.

[38] R. Chakraborty and M. Nei, "Genetic differentiation of quantitative characters between populations or species: I. mutation and random genetic drift," *Genetics Research*, vol. 39, no. 3, pp. 303–314, 1982.

[39] J. Li, H. Li, M. Jakobsson, S. Li, P. SjÖDin, and M. Lascoux, "Joint analysis of demography and selection in population genetics: where do we stand and where could we go?" *Molecular ecology*, vol. 21, no. 1, pp. 28–44, 2012.

[40] G. Luikart, P. R. England, D. Tallmon, S. Jordan, and P. Taberlet, "The power and promise of population genomics: from genotyping to genome typing," *Nature reviews genetics*, vol. 4, no. 12, p. 981, 2003.

[41] C. M. Schlebusch and M. Jakobsson, "Tales of human migration, admixture, and selection in africa," *Annual review of genomics and human genetics*, vol. 19, pp. 405–428, 2018.

[42] J. Prado-Martinez, P. H. Sudmant, J. M. Kidd, H. Li, J. L. Kelley, B. Lorente-Galdos, K. R. Veeramah, A. E. Woerner, T. D. Oâconnor, G. Santpere *et al.*, "Great ape genetic diversity and population history," *Nature*, vol. 499, no. 7459, p. 471, 2013.

[43] M. Jakobsson, S. W. Scholz, P. Scheet, J. R. Gibbs, J. M. VanLiere, H.-C. Fung, Z. A. Szpiech, J. H. Degnan, K. Wang, R. Guerreiro *et al.*, "Genotype, haplotype and copy-number variation in worldwide human populations," *Nature*, vol. 451, no. 7181, p. 998, 2008.

[44] S. Ramachandran, O. Deshpande, C. C. Roseman, N. A. Rosenberg, M. W. Feldman, and L. L. Cavalli-Sforza, "Support from the relationship of genetic and geographic distance in human populations for a serial founder effect originating in africa," *Proceedings of the National Academy of Sciences*, vol. 102, no. 44, pp. 15 942–15 947, 2005.

[45] J. Z. Li, D. M. Absher, H. Tang, A. M. Southwick, A. M. Casto, S. Ramachandran, H. M. Cann, G. S. Barsh, M. Feldman, L. L. Cavalli-Sforza *et al.*, "Worldwide human relationships inferred from genome-wide patterns of variation," *science*, vol. 319, no. 5866, pp. 1100–1104, 2008.

[46] C. M. Schlebusch, H. Malmström, T. Günther, P. Sjödin, A. Coutinho, H. Edlund, A. R. Munters, M. Vicente, M. Steyn, H. Soodyall *et al.*, "Southern african ancient genomes estimate modern human divergence to 350,000 to 260,000 years ago," *Science*, vol. 358, no. 6363, pp. 652–655, 2017.

[47] F. Racimo, M. Sikora, H. Schroeder, and C. Lalueza-Fox, "Beyond broad strokes: sociocultural insights from the study of ancient genomes," *arXiv preprint arXiv:1911.00755*, 2019.

[48] S. Mallick, H. Li, M. Lipson, I. Mathieson, M. Gymrek, F. Racimo, M. Zhao, N. Chennagiri, S. Nordenfelt, A. Tandon *et al.*, "The simons genome diversity project: 300 genomes from 142 diverse populations," *Nature*, vol. 538, no. 7624, pp. 201–206, 2016.

[49] A. Mounier and M. Mirazón Lahr, "Deciphering african late middle pleistocene hominin diversity and the origin of our species," *Nature communications*, vol. 10, no. 1, pp. 1–13, 2019.

[50] J.-J. Hublin, A. Ben-Ncer, S. E. Bailey, S. E. Freidline, S. Neubauer, M. M. Skinner, I. Bergmann, A. Le Cabec, S. Benazzi, K. Harvati *et al.*, "New fossils from jebel irhoud, morocco and the pan-african origin of homo sapiens," *Nature*, vol. 546, no. 7657, pp. 289–292, 2017.

[51] R. Nielsen, J. M. Akey, M. Jakobsson, J. K. Pritchard, S. Tishkoff, and E. Willerslev, "Tracing the peopling of the world through genomics," *Nature*, vol. 541, no. 7637, p. 302, 2017.

[52] N. A. Drake, R. M. Blench, S. J. Armitage, C. S. Bristow, and K. H. White, "Ancient watercourses and biogeography of the sahara explain the peopling of the desert," *Proceedings of the National Academy of Sciences*, vol. 108, no. 2, pp. 458–462, 2011.

[53] L. Pagani, S. Schiffels, D. Gurdasani, P. Danecek, A. Scally, Y. Chen, Y. Xue, M. Haber, R. Ekong, T. Oljira *et al.*, "Tracing the route of modern humans out of africa by using 225 human genome sequences from ethiopians and egyptians," *The American Journal of Human Genetics*, vol. 96, no. 6, pp. 986–991, 2015.

[54] J. C. Larrasoaña, A. P. Roberts, and E. J. Rohling, "Dynamics of green sahara periods and their role in hominin evolution," *PloS one*, vol. 8, no. 10, p. e76514, 2013.

[55] R. Tjallingii, M. Claussen, J.-B. W. Stuut, J. Fohlmeister, A. Jahn, T. Bickert,

F. Lamy, and U. Röhl, "Coherent high-and low-latitude control of the northwest african hydrological balance," *Nature Geoscience*, vol. 1, no. 10, p. 670, 2008.

[56] N. Drake and P. Breeze, "Climate change and modern human occupation of the sahara from mis 6-2," in *Africa from MIS 6-2*. Springer, 2016, pp. 103–122.

[57] I. S. Castañeda, S. Mulitza, E. Schefuß, R. A. L. dos Santos, J. S. S. Damsté, and S. Schouten, "Wet phases in the sahara/sahel region and human migration patterns in north africa," *Proceedings of the National Academy of Sciences*, vol. 106, no. 48, pp. 20 159–20 163, 2009.

[58] N. R. Rebollo, S. Weiner, F. Brock, L. Meignen, P. Goldberg, A. Belfer-Cohen, O. Bar-Yosef, and E. Boaretto, "New radiocarbon dating of the transition from the middle to the upper paleolithic in kebara cave, israel," *Journal of Archaeological Science*, vol. 38, no. 9, pp. 2424–2433, 2011.

[59] A. Gibbons, "Oldest homo sapiens genome pinpoints neandertal input," 2014.

[60] T. Higham, T. Compton, C. Stringer, R. Jacobi, B. Shapiro, E. Trinkaus, B. Chandler, F. Gröning, C. Collins, S. Hillson *et al.*, "The earliest evidence for anatomically modern humans in northwestern Europe," *Nature*, vol. 479, no. 7374, p. 521, 2011.

[61] R. E. Green, J. Krause, A. W. Briggs, T. Maricic, U. Stenzel, M. Kircher, N. Patterson, H. Li, W. Zhai, M. H.-Y. Fritz *et al.*, "A draft sequence of the neandertal genome," *science*, vol. 328, no. 5979, pp. 710–722, 2010.

[62] K. Prüfer, F. Racimo, N. Patterson, F. Jay, S. Sankararaman, S. Sawyer, A. Heinze, G. Renaud, P. H. Sudmant, C. De Filippo *et al.*, "The complete genome sequence of a neanderthal from the altai mountains," *Nature*, vol. 505, no. 7481, p. 43, 2014.

[63] S. Sankararaman, N. Patterson, H. Li, S. Pääbo, and D. Reich, "The date of interbreeding between neandertals and modern humans," *PLoS genetics*, vol. 8, no. 10, p. e1002947, 2012.

[64] P. U. Clark, A. S. Dyke, J. D. Shakun, A. E. Carlson, J. Clark, B. Wohlfarth, J. X. Mitrovica, S. W. Hostetler, and A. M. McCabe, "The last glacial maximum," *science*, vol. 325, no. 5941, pp. 710–714, 2009.

[65] S. P. Harrison, K. E. Kohfeld, C. Roelandt, and T. Claquin, "The role of dust in climate changes today, at the last glacial maximum and in the future," *Earth-Science Reviews*, vol. 54, no. 1-3, pp. 43–80, 2001.

[66] V. Janská, B. Jiménez-Alfaro, M. Chytrỳ, J. Divíšek, O. Anenkhonov, A. Korolyuk, N. Lashchinskyi, and M. Culek, "Palaeodistribution modelling of european vegetation types at the last glacial maximum using modern analogues from siberia: Prospects and limitations," *Quaternary Science Reviews*, vol. 159, pp. 103–115, 2017.

[67] Q. Fu, C. Posth, M. Hajdinjak, M. Petr, S. Mallick, D. Fernandes, A. Furtwängler, W. Haak, M. Meyer, A. Mittnik *et al.*, "The genetic history of ice age europe," *Nature*, vol. 534, no. 7606, p. 200, 2016.

[68] M. Sikora, A. Seguin-Orlando, V. C. Sousa, A. Albrechtsen, T. Korneliussen, A. Ko, S. Rasmussen, I. Dupanloup, P. R. Nigst, M. D. Bosch *et al.*, "Ancient genomes show social and reproductive behavior of early upper paleolithic foragers," *Science*, vol. 358, no. 6363, pp. 659–662, 2017.

[69] A. Seguin-Orlando, T. S. Korneliussen, M. Sikora, A.-S. Malaspinas,

A. Manica, I. Moltke, A. Albrechtsen, A. Ko, A. Margaryan, V. Moiseyev *et al.*, "Genomic structure in europeans dating back at least 36,200 years," *Science*, vol. 346, no. 6213, pp. 1113–1118, 2014.

[70] C. Posth, G. Renaud, A. Mittnik, D. G. Drucker, H. Rougier, C. Cupillard, F. Valentin, C. Thevenet, A. Furtwängler, C. Wißing *et al.*, "Pleistocene mitochondrial genomes suggest a single major dispersal of non-africans and a late glacial population turnover in europe," *Current Biology*, vol. 26, no. 6, pp. 827–833, 2016.

[71] H. Li and R. Durbin, "Inference of human population history from individual whole-genome sequences," *Nature*, vol. 475, no. 7357, p. 493, 2011.

[72] P. Skoglund, H. Malmström, A. Omrak, M. Raghavan, C. Valdiosera, T. Günther, P. Hall, K. Tambets, J. Parik, K.-G. Sjögren *et al.*, "Genomic diversity and admixture differs for stone-age scandinavian foragers and farmers," *Science*, vol. 344, no. 6185, pp. 747–750, 2014.

[73] J. E. Tierney, F. S. Pausata *et al.*, "Rainfall regimes of the green sahara," *Science advances*, vol. 3, no. 1, p. e1601503, 2017.

[74] F. Marshall and E. Hildebrand, "Cattle before crops: the beginnings of food production in africa," *Journal of World Prehistory*, vol. 16, no. 2, pp. 99–143, 2002.

[75] J. Dunne, R. P. Evershed, M. Salque, L. Cramp, S. Bruni, K. Ryan, S. Biagetti, and S. di Lernia, "First dairying in green saharan africa in the fifth millennium bc," *Nature*, vol. 486, no. 7403, p. 390, 2012.

[76] K. Manning and A. Timpson, "The demographic response to holocene climate change in the sahara," *Quaternary Science Reviews*, vol. 101, pp. 28–35, 2014.

[77] B. M. Henn, L. R. Botigué, S. Gravel, W. Wang, A. Brisbin, J. K. Byrnes, K. Fadhlaoui-Zid, P. A. Zalloua, A. Moreno-Estrada, J. Bertranpetit *et al.*, "Genomic ancestry of north africans supports back-to-africa migrations," *PLoS genetics*, vol. 8, no. 1, p. e1002397, 2012.

[78] M. van de Loosdrecht, A. Bouzouggar, L. Humphrey, C. Posth, N. Barton, A. Aximu-Petri, B. Nickel, S. Nagel, E. H. Talbi, M. A. El Hajraoui *et al.*, "Pleistocene north african genomes link near eastern and sub-saharan african human populations," *Science*, vol. 360, no. 6388, pp. 548–552, 2018.

[79] R. Fregel, F. L. Méndez, Y. Bokbot, D. Martin-Socas, M. D. Camalich-Massieu, J. Santana, J. Morales, M. C. Ávila-Arcos, P. A. Underhill, B. Shapiro *et al.*, "Ancient genomes from north africa evidence prehistoric migrations to the maghreb from both the levant and europe," *Proceedings of the National Academy of Sciences*, vol. 115, no. 26, pp. 6774–6779, 2018.

[80] V. G. Childe, "Vii.âwhen did the beaker-folk arrive?" *Archaeologia*, vol. 74, pp. 159–180, 1925.

[81] M. Zvelebil, "On the transition to farming in europe, or what was spreading with the neolithic: a reply to ammerman (1989)," *Antiquity*, vol. 63, no. 239, pp. 379–383, 1989.

[82] M. Rosenberg and A. Erim-Özdoğan, "The neolithic in southeastern anatolia," 2011.

[83] G. Barker, *Prehistoric farming in Europe*. Cambridge University Press, 1985.

[84] T. Champion, C. Gamble, S. Shennan, and A. Whittle, *Prehistoric Europe*. Routledge, 2016.

[85] P. Skoglund, H. Malmström, M. Raghavan, J. Storå, P. Hall, E. Willerslev, M. T. P. Gilbert, A. Götherström, and M. Jakobsson, “Origins and genetic legacy of neolithic farmers and hunter-gatherers in europe,” *Science*, vol. 336, no. 6080, pp. 466–469, 2012.
[86] I. Lazaridis, N. Patterson, A. Mittnik, G. Renaud, S. Mallick, K. Kirsanow, P. H. Sudmant, J. G. Schraiber, S. Castellano, M. Lipson *et al.*, “Ancient human genomes suggest three ancestral populations for present-day europeans,” *Nature*, vol. 513, no. 7518, p. 409, 2014.
[87] T. Günther, C. Valdiosera, H. Malmström, I. Ureña, R. Rodriguez-Varela, Ó. O. Sverrisdóttir, E. A. Daskalaki, P. Skoglund, T. Naidoo, E. M. Svensson *et al.*, “Ancient genomes link early farmers from atapuerca in spain to modern-day basques,” *Proceedings of the National Academy of Sciences*, vol. 112, no. 38, pp. 11 917–11 922, 2015.
[88] W. Haak, I. Lazaridis, N. Patterson, N. Rohland, S. Mallick, B. Llamas, G. Brandt, S. Nordenfelt, E. Harney, K. Stewardson *et al.*, “Massive migration from the steppe was a source for indo-european languages in europe,” *Nature*, vol. 522, no. 7555, p. 207, 2015.
[89] I. Olalde, H. Schroeder, M. Sandoval-Velasco, L. Vinner, I. Lobón, O. Ramirez, S. Civit, P. García Borja, D. C. Salazar-García, S. Talamo *et al.*, “A common genetic origin for early farmers from mediterranean cardial and central european lbk cultures,” *Molecular biology and evolution*, vol. 32, no. 12, pp. 3132–3142, 2015.
[90] I. Mathieson, I. Lazaridis, N. Rohland, S. Mallick, N. Patterson, S. A. Roodenberg, E. Harney, K. Stewardson, D. Fernandes, M. Novak *et al.*, “Genome-wide patterns of selection in 230 ancient eurasians,” *Nature*, vol. 528, no. 7583, p. 499, 2015.
[91] A. Omrak, T. Günther, C. Valdiosera, E. M. Svensson, H. Malmström, H. Kiesewetter, W. Aylward, J. Storå, M. Jakobsson, and A. Götherström, “Genomic evidence establishes anatolia as the source of the european neolithic gene pool,” *Current Biology*, vol. 26, no. 2, pp. 270–275, 2016.
[92] A. Goldberg, T. Günther, N. A. Rosenberg, and M. Jakobsson, “Ancient x chromosomes reveal contrasting sex bias in neolithic and bronze age eurasian migrations,” *Proceedings of the National Academy of Sciences*, vol. 114, no. 10, pp. 2657–2662, 2017.
[93] A. W. Whittle, *Europe in the Neolithic: the creation of new worlds*. Cambridge University Press, 1996.
[94] T. D. Price, “2000. europe’s first farmers.”
[95] M. Rivollat, C. Jeong, S. Schiffels, İ. Küçükkalıpçı, M.-H. Pemonge, A. B. Rohrlach, K. W. Alt, D. Binder, S. Friederich, E. Ghesquière *et al.*, “Ancient genome-wide dna from france highlights the complexity of interactions between mesolithic hunter-gatherers and neolithic farmers,” *Science advances*, vol. 6, no. 22, p. eaaz5344, 2020.
[96] A. J. Ammerman and L. L. Cavalli-Sforza, “Measuring the rate of spread of early farming in europe,” *Man*, pp. 674–688, 1971.
[97] J. Guilaine, “La diffusion de lâagriculture en europe: une hypothèse arythmique. la difusión de lâagricultura en europa: una hipótesis aritmética,” *Zephyrus*, vol. 53, p. 54, 2001.

[98] J. Zilhão, "Radiocarbon evidence for maritime pioneer colonization at the origins of farming in west mediterranean europe," *Proceedings of the national Academy of Sciences*, vol. 98, no. 24, pp. 14 180–14 185, 2001.
[99] C. Manen, "Dynamiques spatio-temporelles et culturelles de la néolithisation ouest-méditerranéenne," in *La transition néolithique en Méditerranée (Actes du colloque de Toulouse, 2011)*, 2014, pp. 405–418.
[100] S. Mulazzani, L. Belhouchet, L. Salanova, N. Aouadi, Y. Dridi, W. Eddargach, J. Morales, O. Tombret, A. Zazzo, and J. Zoughlami, "The emergence of the neolithic in north africa: A new model for the eastern maghreb," *Quaternary International*, vol. 410, pp. 123–143, 2016.
[101] I. Lazaridis, D. Nadel, G. Rollefson, D. C. Merrett, N. Rohland, S. Mallick, D. Fernandes, M. Novak, B. Gamarra, K. Sirak *et al.*, "Genomic insights into the origin of farming in the ancient near east," *Nature*, vol. 536, no. 7617, p. 419, 2016.
[102] M. Lipson, A. Szécsényi-Nagy, S. Mallick, A. Pósa, B. Stégmár, V. Keerl, N. Rohland, K. Stewardson, M. Ferry, M. Michel *et al.*, "Parallel palaeogenomic transects reveal complex genetic history of early european farmers," *Nature*, vol. 551, no. 7680, p. 368, 2017.
[103] C. Valdiosera, T. Günther, J. C. Vera-Rodríguez, I. Ureña, E. Iriarte, R. Rodríguez-Varela, L. G. Simões, R. M. Martínez-Sánchez, E. M. Svensson, H. Malmström *et al.*, "Four millennia of iberian biomolecular prehistory illustrate the impact of prehistoric migrations at the far end of eurasia," *Proceedings of the National Academy of Sciences*, vol. 115, no. 13, pp. 3428–3433, 2018.
[104] I. Olalde, S. Mallick, N. Patterson, N. Rohland, V. Villalba-Mouco, M. Silva, K. Dulias, C. J. Edwards, F. Gandini, M. Pala *et al.*, "The genomic history of the iberian peninsula over the past 8000 years," *Science*, vol. 363, no. 6432, pp. 1230–1234, 2019.
[105] J. Linstädter, I. Medved, M. Solich, and G.-C. Weniger, "Neolithisation process within the alboran territory: Models and possible african impact," *Quaternary International*, vol. 274, pp. 219–232, 2012.
[106] C. Manen, "La production céramique de pont de roque-haute: synthèse et comparaisons," 2007.
[107] M. C. Sánchez, F. J. J. Espejo, M. D. S. Vallejo, J. F. G. Bao, A. F. Carvalho, F. Martinez-Ruiz, M. R. Gamiz, J.-A. Flores, A. Paytan, J. A. L. Sáez *et al.*, "The mesolithic–neolithic transition in southern iberia," *Quaternary Research*, vol. 77, no. 2, pp. 221–234, 2012.
[108] H. Martins, F. X. Oms, L. Pereira, A. W. Pike, K. Rowsell, and J. Zilhão, "Radiocarbon dating the beginning of the neolithic in iberia: new results, new problems," *Journal of Mediterranean Archaeology*, vol. 28, no. 1, pp. 105–131, 2015.
[109] J. Zilhão, "Early prehistoric navigation in the western mediterranean: Implications for the neolithic transition in iberia and the maghreb," *Eurasian Prehistory. Island Archaeology and the Origins of Seafaring in the Eastern Mediterranean*, vol. 11, pp. 185–200, 2014.
[110] R. M. Martínez-Sánchez, J. C. Vera-Rodríguez, G. Pérez-Jordà, L. Pena-Chocarro, and Y. Bokbot, "The beginning of the neolithic in

northwestern morocco," *Quaternary International*, vol. 470, pp. 485–496, 2018.

[111] C. Manen, G. Marchand, and A. Carvalho, "Le néolithique ancien de la péninsule ibérique: vers une nouvelle évaluation du mirage africain?" in *XXVIe Congrès Préhistoriuque de France:" Un siècle de construction du discours scientifique en préhistoire"*, vol. 3. Société préhistorique française, 2004, pp. 133–151.

[112] C. Broodbank and G. Lucarini, "The dynamics of mediterranean africa, ca. 9600–1000 bc: An interpretative synthesis of knowns and unknowns," 2020.

[113] S. di Lernia, "The emergence and spread of herding in northern africa," in *The Oxford handbook of African archaeology*, 2013.

[114] F. A. Hassan, "Holocene palaeoclimates of africa," *African Archaeological Review*, vol. 14, no. 4, pp. 213–230, 1997.

[115] M. Cremaschi and S. Di Lernia, "Holocene climatic changes and cultural dynamics in the libyan sahara," *African Archaeological Review*, vol. 16, no. 4, pp. 211–238, 1999.

[116] R. Kuper and S. Kropelin, "Climate-controlled holocene occupation in the sahara: motor of africa's evolution," *science*, vol. 313, no. 5788, pp. 803–807, 2006.

[117] J. Morales, S. Mulazzani, L. Belhouchet, A. Zazzo, L. Berrio, W. Eddargach, A. Cervi, H. Hamdi, M. Saidi, A. Coppa *et al.*, "First preliminary evidence for basketry and nut consumption in the capsian culture (ca. 10,000–7500 bp): Archaeobotanical data from new excavations at el mekta, tunisia," *Journal of Anthropological Archaeology*, vol. 37, pp. 128–139, 2015.

[118] E. A. Garcea, "Semi-permanent foragers in semi-arid environments of north africa," *World Archaeology*, vol. 38, no. 2, pp. 197–219, 2006.

[119] C. Roubet, *Économie pastorale pré-agricole en Algérie orientale. Le Néolithique de tradition capsienne, lâexemple de lâAurès*. Persée-Portail des revues scientifiques en SHS, 1979, vol. 1, no. 1.

[120] C. Gamba, E. R. Jones, M. D. Teasdale, R. L. McLaughlin, G. Gonzalez-Fortes, V. Mattiangeli, L. Domboróczki, I. Kővári, I. Pap, A. Anders *et al.*, "Genome flux and stasis in a five millennium transect of european prehistory," *Nature communications*, vol. 5, p. 5257, 2014.

[121] D. Y. Yang, B. Eng, J. S. Waye, J. C. Dudar, and S. R. Saunders, "Improved dna extraction from ancient bones using silica-based spin columns," *American Journal of Physical Anthropology: The Official Publication of the American Association of Physical Anthropologists*, vol. 105, no. 4, pp. 539–543, 1998.

[122] N. Rohland and M. Hofreiter, "Ancient dna extraction from bones and teeth," *Nature protocols*, vol. 2, no. 7, pp. 1756–1762, 2007.

[123] J. Lachance and S. A. Tishkoff, "Snp ascertainment bias in population genetic analyses: why it is important, and how to correct it," *Bioessays*, vol. 35, no. 9, pp. 780–786, 2013.

[124] T. Maricic, M. Whitten, and S. Pääbo, "Multiplexed dna sequence capture of mitochondrial genomes using pcr products," *PloS one*, vol. 5, no. 11, p. e14004, 2010.

[125] E. Cappellini, A. Prohaska, F. Racimo, F. Welker, M. W. Pedersen, M. E. Allentoft, P. de Barros Damgaard, P. Gutenbrunner, J. Dunne, S. Hammann

et al., “Ancient biomolecules and evolutionary inference,” *Annual review of biochemistry*, vol. 87, pp. 1029–1060, 2018.

[126] M. C. Ávila-Arcos, E. Cappellini, J. A. Romero-Navarro, N. Wales, J. V. Moreno-Mayar, M. Rasmussen, S. L. Fordyce, R. Montiel, J.-P. Vielle-Calzada, E. Willerslev *et al.*, “Application and comparison of large-scale solution-based dna capture-enrichment methods on ancient dna,” *Scientific reports*, vol. 1, no. 1, pp. 1–5, 2011.

[127] D. I. Cruz-Dávalos, B. Llamas, C. Gaunitz, A. Fages, C. Gamba, J. Soubrier, P. Librado, A. Seguin-Orlando, M. Pruvost, A. H. Alfarhan *et al.*, “Experimental conditions improving in-solution target enrichment for ancient dna,” *Molecular ecology resources*, vol. 17, no. 3, pp. 508–522, 2017.

[128] N. Patterson, P. Moorjani, Y. Luo, S. Mallick, N. Rohland, Y. Zhan, T. Genschoreck, T. Webster, and D. Reich, “Ancient admixture in human history,” *Genetics*, vol. 192, no. 3, pp. 1065–1093, 2012.

[129] A. G. Clark, M. J. Hubisz, C. D. Bustamante, S. H. Williamson, and R. Nielsen, “Ascertainment bias in studies of human genome-wide polymorphism,” *Genome research*, vol. 15, no. 11, pp. 1496–1502, 2005.

[130] R. Nielsen, M. J. Hubisz, and A. G. Clark, “Reconstituting the frequency spectrum of ascertained single-nucleotide polymorphism data,” *Genetics*, vol. 168, no. 4, pp. 2373–2382, 2004.

[131] A. Albrechtsen, F. C. Nielsen, and R. Nielsen, “Ascertainment biases in snp chips affect measures of population divergence,” *Molecular biology and evolution*, vol. 27, no. 11, pp. 2534–2547, 2010.

[132] R. Nielsen, “Population genetic analysis of ascertained snp data,” *Human genomics*, vol. 1, no. 3, p. 218, 2004.

[133] C. Makarewicz, N. Marom, and G. Bar-Oz, “Ensure equal access to ancient dna,” *Nature*, vol. 548, no. 7666, pp. 158–158, 2017.

[134] M. E. Prendergast and E. Sawchuk, “Boots on the ground in africa’s ancient dna ârevolutionâ: archaeological perspectives on ethics and best practices,” *Antiquity*, vol. 92, no. 363, pp. 803–815, 2018.

[135] N. Rascovan, K.-G. Sjögren, K. Kristiansen, R. Nielsen, E. Willerslev, C. Desnues, and S. Rasmussen, “Emergence and spread of basal lineages of yersinia pestis during the neolithic decline,” *Cell*, vol. 176, no. 1-2, pp. 295–305, 2019.

[136] M. Meyer, M. Kircher, M.-T. Gansauge, H. Li, F. Racimo, S. Mallick, J. G. Schraiber, F. Jay, K. Prüfer, C. De Filippo *et al.*, “A high-coverage genome sequence from an archaic denisovan individual,” *Science*, vol. 338, no. 6104, pp. 222–226, 2012.

[137] . G. P. Consortium *et al.*, “An integrated map of genetic variation from 1,092 human genomes,” *Nature*, vol. 491, no. 7422, p. 56, 2012.

[138] B. Pakendorf and M. Stoneking, “Mitochondrial dna and human evolution,” *Annual review of genomics and human genetics*, vol. 6, no. 1, pp. 165–183, 2005.

[139] F. Prugnolle and T. De MeeËus, “Inferring sex-biased dispersal from population genetic tools: a review,” *Heredity*, vol. 88, no. 3, pp. 161–165, 2002.

[140] R. Nielsen, J. S. Paul, A. Albrechtsen, and Y. S. Song, “Genotype and snp

calling from next-generation sequencing data," *Nature Reviews Genetics*, vol. 12, no. 6, pp. 443–451, 2011.
[141] S. Gravel, "Population genetics models of local ancestry," *Genetics*, vol. 191, no. 2, pp. 607–619, 2012.
[142] R. Nielsen and M. Slatkin, "An introduction to population genetics," *Sunderland, MA: Sinauer Associates.[Google Scholar]*, 2013.
[143] C. Wang, Z. A. Szpiech, J. H. Degnan, M. Jakobsson, T. J. Pemberton, J. A. Hardy, A. B. Singleton, and N. A. Rosenberg, "Comparing spatial maps of human population-genetic variation using procrustes analysis," *Statistical applications in genetics and molecular biology*, vol. 9, no. 1, 2010.
[144] J. Novembre and M. Stephens, "Interpreting principal component analyses of spatial population genetic variation," *Nature genetics*, vol. 40, no. 5, pp. 646–649, 2008.
[145] J. K. Pritchard, M. Stephens, and P. Donnelly, "Inference of population structure using multilocus genotype data," *Genetics*, vol. 155, no. 2, pp. 945–959, 2000.
[146] D. H. Alexander, J. Novembre, and K. Lange, "Fast model-based estimation of ancestry in unrelated individuals," *Genome research*, vol. 19, no. 9, pp. 1655–1664, 2009.
[147] D. J. Lawson, L. Van Dorp, and D. Falush, "A tutorial on how not to over-interpret structure and admixture bar plots," *Nature communications*, vol. 9, no. 1, pp. 1–11, 2018.
[148] D. Reich, K. Thangaraj, N. Patterson, A. L. Price, and L. Singh, "Reconstructing indian population history," *Nature*, vol. 461, no. 7263, pp. 489–494, 2009.
[149] M. Raghavan, M. Steinrücken, K. Harris, S. Schiffels, S. Rasmussen, M. DeGiorgio, A. Albrechtsen, C. Valdiosera, M. C. Ávila-Arcos, A.-S. Malaspinas *et al.*, "Genomic evidence for the pleistocene and recent population history of native americans," *Science*, vol. 349, no. 6250, p. aab3884, 2015.
[150] E. Harney, N. Patterson, D. Reich, and J. Wakeley, "Assessing the performance of qpadm: a statistical tool for studying population admixture," *Genetics*, vol. 217, no. 4, p. iyaa045, 2021.
[151] R. Maier, P. Flegontov, O. Flegontova, P. Changmai, and D. Reich, "On the limits of fitting complex models of population history to genetic data," *bioRxiv*, 2022.
[152] P.-R. Loh, M. Lipson, N. Patterson, P. Moorjani, J. K. Pickrell, D. Reich, and B. Berger, "Inferring admixture histories of human populations using linkage disequilibrium," *Genetics*, vol. 193, no. 4, pp. 1233–1254, 2013.
[153] P. Moorjani, S. Sankararaman, Q. Fu, M. Przeworski, N. Patterson, and D. Reich, "A genetic method for dating ancient genomes provides a direct estimate of human generation interval in the last 45,000 years," *Proceedings of the National Academy of Sciences*, vol. 113, no. 20, pp. 5652–5657, 2016.
[154] M. Chintalapati, N. Patterson, and P. Moorjani, "The spatiotemporal patterns of major human admixture events during the european holocene," *Elife*, vol. 11, p. e77625, 2022.
[155] M. Kirin, R. McQuillan, C. S. Franklin, H. Campbell, P. M. McKeigue, and J. F. Wilson, "Genomic runs of homozygosity record population history and

consanguinity," *PloS one*, vol. 5, no. 11, p. e13996, 2010.
[156] F. C. Ceballos, P. K. Joshi, D. W. Clark, M. Ramsay, and J. F. Wilson, "Runs of homozygosity: windows into population history and trait architecture," *Nature Reviews Genetics*, vol. 19, no. 4, pp. 220–234, 2018.
[157] I. Mathieson, S. Alpaslan-Roodenberg, C. Posth, A. Szécsényi-Nagy, N. Rohland, S. Mallick, I. Olalde, N. Broomandkhoshbacht, F. Candilio, O. Cheronet *et al.*, "The genomic history of southeastern europe," *Nature*, vol. 555, no. 7695, p. 197, 2018.
[158] K. R. Hill, R. S. Walker, M. Božičević, J. Eder, T. Headland, B. Hewlett, A. M. Hurtado, F. Marlowe, P. Wiessner, and B. Wood, "Co-residence patterns in hunter-gatherer societies show unique human social structure," *science*, vol. 331, no. 6022, pp. 1286–1289, 2011.
[159] M. Dyble, G. D. Salali, N. Chaudhary, A. Page, D. Smith, J. Thompson, L. Vinicius, R. Mace, and A. B. Migliano, "Sex equality can explain the unique social structure of hunter-gatherer bands," *Science*, vol. 348, no. 6236, pp. 796–798, 2015.
[160] V. Villalba-Mouco, M. S. van de Loosdrecht, C. Posth, R. Mora, J. Martínez-Moreno, M. Rojo-Guerra, D. C. Salazar-García, J. I. Royo-Guillén, M. Kunst, H. Rougier *et al.*, "Survival of late pleistocene hunter-gatherer ancestry in the iberian peninsula," *Current Biology*, 2019.
[161] S. Brunel, E. A. Bennett, L. Cardin, D. Garraud, H. Barrand Emam, A. Beylier, B. Boulestin, F. Chenal, E. Ciesielski, F. Convertini *et al.*, "Ancient genomes from present-day france unveil 7,000 years of its demographic history," *Proceedings of the National Academy of Sciences*, vol. 117, no. 23, pp. 12 791–12 798, 2020.
[162] M. Álvarez, I. B. Godino, A. Balbo, and M. Madella, "Shell middens as archives of past environments, human dispersal and specialized resource management," pp. 1–7, 2011.
[163] M. Péquart, S.-J. Péquart, M. Boule, and H.-V. Vallois, *Téviec, station-nécrolopole mésolithique du Morbihan*. Masson et Cie, 1937.
[164] K. Rebay-Salisbury, "3. personal relationships between co-buried individuals in the central european early bronze age," *AmS-Skrifter*, no. 26, pp. 35–48, 2018.
[165] P. Soares, F. Alshamali, J. B. Pereira, V. Fernandes, N. M. Silva, C. Afonso, M. D. Costa, E. Musilová, V. Macaulay, M. B. Richards *et al.*, "The expansion of mtdna haplogroup l3 within and out of africa," *Molecular biology and evolution*, vol. 29, no. 3, pp. 915–927, 2012.
[166] K. Hardy, A. Camara, R. Piqué, E. Dioh, M. Guèye, H. D. Diadhiou, M. Faye, and M. Carré, "Shellfishing and shell midden construction in the saloum delta, senegal," *Journal of Anthropological Archaeology*, vol. 41, pp. 19–32, 2016.
[167] L. G. Straus, "Africa and iberia in the pleistocene," *Quaternary International*, vol. 75, no. 1, pp. 91–102, 2001.
[168] R. M. Martinez Sanchez, J. C. Vera Rodriguez, L. Peña-Chocarro, Y. Bokbot, G. Perez Jorda, and S. Pardo-Gordó, "The middle neolithic of moroccoâs north-western atlantic strip: New evidence from the el-khil caves (tangier)," *African Archaeological Review*, vol. 35, no. 3, pp. 417–442, 2018.
[169] M. E. Prendergast, M. Lipson, E. A. Sawchuk, I. Olalde, C. A. Ogola,

N. Rohland, K. A. Sirak, N. Adamski, R. Bernardos, N. Broomandkhoshbacht *et al.*, "Ancient dna reveals a multistep spread of the first herders into sub-saharan africa," *Science*, vol. 365, no. 6448, p. eaaw6275, 2019.

[170] J. H. Marcus, C. Posth, H. Ringbauer, L. Lai, R. Skeates, C. Sidore, J. Beckett, A. Furtwängler, A. Olivieri, C. W. Chiang *et al.*, "Genetic history from the middle neolithic to present on the mediterranean island of sardinia," *Nature communications*, vol. 11, no. 1, pp. 1–14, 2020.

[171] G. Gonzalez-Fortes, F. Tassi, E. Trucchi, K. Henneberger, J. L. Paijmans, D. Díez-del Molino, H. Schroeder, R. R. Susca, C. Barroso-Ruíz, F. J. Bermúdez *et al.*, "A western route of prehistoric human migration from africa into the iberian peninsula," *Proceedings of the Royal Society B*, vol. 286, no. 1895, p. 20182288, 2019.

[172] K. Harris and R. Nielsen, "The genetic cost of neanderthal introgression," *Genetics*, vol. 203, no. 2, pp. 881–891, 2016.

[173] L. Ríos, T. L. Kivell, C. Lalueza-Fox, A. Estalrrich, A. García-Tabernero, R. Huguet, Y. Quintino, M. de la Rasilla, and A. Rosas, "Skeletal anomalies in the neandertal family of el sidrón (spain) support a role of inbreeding in neandertal extinction," *Scientific Reports*, vol. 9, no. 1, pp. 1–11, 2019.

[174] K. H. Kjær, M. Winther Pedersen, B. De Sanctis, B. De Cahsan, T. S. Korneliussen, C. S. Michelsen, K. K. Sand, S. Jelavić, A. H. Ruter, A. Schmidt *et al.*, "A 2-million-year-old ecosystem in greenland uncovered by environmental dna," *Nature*, vol. 612, no. 7939, pp. 283–291, 2022.

[175] S. Peyrégne, V. Slon, F. Mafessoni, C. De Filippo, M. Hajdinjak, S. Nagel, B. Nickel, E. Essel, A. Le Cabec, K. Wehrberger *et al.*, "Nuclear dna from two early neandertals reveals 80,000 years of genetic continuity in europe," *Science advances*, vol. 5, no. 6, p. eaaw5873, 2019.

[176] M. Meyer, J.-L. Arsuaga, C. De Filippo, S. Nagel, A. Aximu-Petri, B. Nickel, I. Martínez, A. Gracia, J. M. B. de Castro, E. Carbonell *et al.*, "Nuclear dna sequences from the middle pleistocene sima de los huesos hominins," *Nature*, vol. 531, no. 7595, pp. 504–507, 2016.

[177] T. van der Valk, P. Pečnerová, D. Díez-del Molino, A. Bergström, J. Oppenheimer, S. Hartmann, G. Xenikoudakis, J. A. Thomas, M. Dehasque, E. Sağlıcan *et al.*, "Million-year-old dna sheds light on the genomic history of mammoths," *Nature*, vol. 591, no. 7849, pp. 265–269, 2021.

[178] E. Hagelberg, B. Sykes, and R. Hedges, "Ancient bone dna amplified," *Nature*, vol. 342, no. 6249, pp. 485–485, 1989.

[179] J. A. Armour, T. Anttinen, C. A. May, E. E. Vega, A. Sajantila, J. R. Kidd, K. K. Kidd, J. Bertranpetit, S. Pääbo, and A. J. Jeffreys, "Minisatellite diversity supports a recent african origin for modern humans," *Nature genetics*, vol. 13, no. 2, p. 154, 1996.

[180] S. Wright, "Isolation by distance," *Genetics*, vol. 28, no. 2, p. 114, 1943.

[181] M. Kimura and G. H. Weiss, "The stepping stone model of population structure and the decrease of genetic correlation with distance," *Genetics*, vol. 49, no. 4, p. 561, 1964.

[182] L. L. Cavalli-Sforza, I. Barrai, and A. Edwards, "Analysis of human evolution under random genetic drift," in *Cold Spring Harbor symposia on quantitative biology*, vol. 29. Cold Spring Harbor Laboratory Press, 1964, pp. 9–20.

[183] C. Press, *Human evolutionary genetics*. Garland Science, 2013.
[184] T. Frisch, M. Sørensen, S. Overgaard, M. Lind, and P. Bretlau, "Volume-referent bone turnover estimated from the interlabel area fraction after sequential labeling," *Bone*, vol. 22, no. 6, pp. 677–682, 1998.
[185] O. Bar-Yosef, A. Belfer-Cohen, and D. S. Adler, "The implications of the middle-upper paleolithic chronological boundary in the caucasus to eurasian prehistory," *Anthropologie (1962-)*, vol. 44, no. 1, pp. 49–60, 2006.
[186] M. Pala, A. Olivieri, A. Achilli, M. Accetturo, E. Metspalu, M. Reidla, E. Tamm, M. Karmin, T. Reisberg, B. H. Kashani *et al.*, "Mitochondrial dna signals of late glacial recolonization of europe from near eastern refugia," *The American journal of human genetics*, vol. 90, no. 5, pp. 915–924, 2012.
[187] E. R. Jones, G. Gonzalez-Fortes, S. Connell, V. Siska, A. Eriksson, R. Martiniano, R. L. McLaughlin, M. G. Llorente, L. M. Cassidy, C. Gamba *et al.*, "Upper palaeolithic genomes reveal deep roots of modern eurasians," *Nature communications*, vol. 6, p. 8912, 2015.
[188] I. Olalde, M. E. Allentoft, F. Sánchez-Quinto, G. Santpere, C. W. Chiang, M. DeGiorgio, J. Prado-Martinez, J. A. Rodríguez, S. Rasmussen, J. Quilez *et al.*, "Derived immune and ancestral pigmentation alleles in a 7,000-year-old mesolithic european," *Nature*, vol. 507, no. 7491, p. 225, 2014.
[189] T. Günther and M. Jakobsson, "Genes mirror migrations and cultures in prehistoric europeâa population genomic perspective," *Current Opinion in Genetics & Development*, vol. 41, pp. 115–123, 2016.
[190] T. Günther, H. Malmström, E. M. Svensson, A. Omrak, F. Sánchez-Quinto, G. M. Kılınç, M. Krzewińska, G. Eriksson, M. Fraser, H. Edlund *et al.*, "Population genomics of mesolithic scandinavia: Investigating early postglacial migration routes and high-latitude adaptation," *PLoS biology*, vol. 16, no. 1, p. e2003703, 2018.
[191] G. González-Fortes, E. R. Jones, E. Lightfoot, C. Bonsall, C. Lazar, A. Grandal-dâAnglade, M. D. Garralda, L. Drak, V. Siska, A. Simalcsik *et al.*, "Paleogenomic evidence for multi-generational mixing between neolithic farmers and mesolithic hunter-gatherers in the lower danube basin," *Current Biology*, vol. 27, no. 12, pp. 1801–1810, 2017.
[192] Q. Fu, H. Li, P. Moorjani, F. Jay, S. M. Slepchenko, A. A. Bondarev, P. L. Johnson, A. Aximu-Petri, K. Prüfer, C. de Filippo *et al.*, "Genome sequence of a 45,000-year-old modern human from western siberia," *Nature*, vol. 514, no. 7523, p. 445, 2014.
[193] I. Lazaridis, "The evolutionary history of human populations in europe," *Current opinion in genetics & development*, vol. 53, pp. 21–27, 2018.
[194] I. Lazaridis, A. Belfer-Cohen, S. Mallick, N. Patterson, O. Cheronet, N. Rohland, G. Bar-Oz, O. Bar-Yosef, N. Jakeli, E. Kvavadze *et al.*, "Paleolithic dna from the caucasus reveals core of west eurasian ancestry," *bioRxiv*, p. 423079, 2018.
[195] C.-C. Wang, S. Reinhold, A. Kalmykov, A. Wissgott, G. Brandt, C. Jeong, O. Cheronet, M. Ferry, E. Harney, D. Keating *et al.*, "Ancient human genome-wide data from a 3000-year interval in the caucasus corresponds with eco-geographic regions," *Nature communications*, vol. 10, no. 1, p. 590, 2019.
[196] E. R. Jones, G. Zarina, V. Moiseyev, E. Lightfoot, P. R. Nigst, A. Manica,

R. Pinhasi, and D. G. Bradley, "The neolithic transition in the baltic was not driven by admixture with early european farmers," *Current Biology*, vol. 27, no. 4, pp. 576–582, 2017.
[197] S. Mathieson and I. Mathieson, "Fads1 and the timing of human adaptation to agriculture," *Molecular biology and evolution*, vol. 35, no. 12, pp. 2957–2970, 2018.
[198] C. Bycroft, C. Freeman, D. Petkova, G. Band, L. T. Elliott, K. Sharp, A. Motyer, D. Vukcevic, O. Delaneau, J. OâConnell *et al.*, "The uk biobank resource with deep phenotyping and genomic data," *Nature*, vol. 562, no. 7726, p. 203, 2018.
[199] M. Sohail, R. M. Maier, A. Ganna, A. Bloemendal, A. R. Martin, M. C. Turchin, C. W. Chiang, J. Hirschhorn, M. J. Daly, N. Patterson *et al.*, "Polygenic adaptation on height is overestimated due to uncorrected stratification in genome-wide association studies," *ELife*, vol. 8, p. e39702, 2019.
[200] J. J. Berg, A. Harpak, N. Sinnott-Armstrong, A. M. Joergensen, H. Mostafavi, Y. Field, E. A. Boyle, X. Zhang, F. Racimo, J. K. Pritchard *et al.*, "Reduced signal for polygenic adaptation of height in uk biobank," *ELife*, vol. 8, p. e39725, 2019.
[201] L. Saag, L. Varul, C. L. Scheib, J. Stenderup, M. E. Allentoft, L. Saag, L. Pagani, M. Reidla, K. Tambets, E. Metspalu *et al.*, "Extensive farming in estonia started through a sex-biased migration from the steppe," *Current biology*, vol. 27, no. 14, pp. 2185–2193, 2017.
[202] A. Mittnik, C.-C. Wang, S. Pfrengle, M. Daubaras, G. Zariņa, F. Hallgren, R. Allmäe, V. Khartanovich, V. Moiseyev, M. Tõrv *et al.*, "The genetic prehistory of the baltic sea region," *Nature communications*, vol. 9, no. 1, p. 442, 2018.
[203] N. Kashuba, E. Kırdök, H. Damlien, M. A. Manninen, B. Nordqvist, P. Persson, and A. Götherström, "Ancient dna from mastics solidifies connection between material culture and genetics of mesolithic hunter–gatherers in scandinavia," *Communications biology*, vol. 2, no. 1, p. 185, 2019.
[204] A. R. Martin, M. Lin, J. M. Granka, J. W. Myrick, X. Liu, A. Sockell, E. G. Atkinson, C. J. Werely, M. Möller, M. S. Sandhu *et al.*, "An unexpectedly complex architecture for skin pigmentation in africans," *Cell*, vol. 171, no. 6, pp. 1340–1353, 2017.
[205] N. G. Crawford, D. E. Kelly, M. E. Hansen, M. H. Beltrame, S. Fan, S. L. Bowman, E. Jewett, A. Ranciaro, S. Thompson, Y. Lo *et al.*, "Loci associated with skin pigmentation identified in african populations," *Science*, vol. 358, no. 6365, p. eaan8433, 2017.
[206] N. G. Jablonski and G. Chaplin, "The colours of humanity: the evolution of pigmentation in the human lineage," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 372, no. 1724, p. 20160349, 2017.
[207] M. Sørensen, T. Rankama, J. Kankaanpää, K. Knutsson, H. Knutsson, S. Melvold, B. V. Eriksen, and H. Glørstad, "The first eastern migrations of people and knowledge into scandinavia: evidence from studies of mesolithic technology, 9th-8th millennium bc," *Norwegian Archaeological Review*,

vol. 46, no. 1, pp. 19–56, 2013.
[208] M. Sørensen, "The arrival and development of pressure blade technology in southern scandinavia," in *The Emergence of Pressure Blade Making*. Springer, 2012, pp. 237–259.
[209] M. A. Manninen and K. Knutsson, "Lithic raw material diversification as an adaptive strategyâtechnology, mobility, and site structure in late mesolithic northernmost europe," *Journal of Anthropological Archaeology*, vol. 33, pp. 84–98, 2014.
[210] M. A. Manninen *et al.*, "Culture, behaviour, and the 8200 cal bp cold event: Organisational change and culture environment dynamics in late mesolithic northern fennoscandia," *Monographs of the Archaeological Society of Finland*, 2014.
[211] F. Sánchez-Quinto, H. Schroeder, O. Ramirez, M. C. Ávila-Arcos, M. Pybus, I. Olalde, A. M. Velazquez, M. E. P. Marcos, J. M. V. Encinas, J. Bertranpetit *et al.*, "Genomic affinities of two 7,000-year-old iberian hunter-gatherers," *Current biology*, vol. 22, no. 16, pp. 1494–1499, 2012.
[212] T. Günther and M. Jakobsson, "Population genomic analyses of dna from ancient remains," *Handbook of Statistical Genomics 4e 2V SET*, pp. 295–40, 2019.
[213] N. Patterson, A. L. Price, and D. Reich, "Population structure and eigenanalysis," *PLoS genetics*, vol. 2, no. 12, p. e190, 2006.
[214] S. Brace, Y. Diekmann, T. J. Booth, L. van Dorp, Z. Faltyskova, N. Rohland, S. Mallick, I. Olalde, M. Ferry, M. Michel *et al.*, "Ancient genomes indicate population replacement in early neolithic britain," *Nature ecology & evolution*, vol. 3, no. 5, p. 765, 2019.
[215] T. J. Booth, "A stranger in a strange land: a perspective on archaeological responses to the palaeogenetic revolution from an archaeologist working amongst palaeogeneticists," *World Archaeology*, pp. 1–16, 2019.
[216] M. Raghavan, P. Skoglund, K. E. Graf, M. Metspalu, A. Albrechtsen, I. Moltke, S. Rasmussen, T. W. Stafford Jr, L. Orlando, E. Metspalu *et al.*, "Upper palaeolithic siberian genome reveals dual ancestry of native americans," *Nature*, vol. 505, no. 7481, p. 87, 2014.
[217] M. Knapp and M. Hofreiter, "Next generation sequencing of ancient dna: requirements, strategies and perspectives," *Genes*, vol. 1, no. 2, pp. 227–243, 2010.
[218] D. Y. Yang and K. Watt, "Contamination controls when preparing archaeological remains for ancient dna analysis," *Journal of Archaeological Science*, vol. 32, no. 3, pp. 331–336, 2005.
[219] F. Riede, "The resettlement of northern europe," in *Oxford Handbook of the Archaeology and Anthropology of Hunter-gatherers*. Oxford University Press, 2014, pp. 556–581.
[220] E. B. Petersen, "The human settlement of southern scandinavia 12,500-8,700 calbc. i: Street, m., rne barton, t. terberger (eds< em>.) humans, environment and chronology of the late glacial of the north european plain. proceedings of workshop 14 (commission xxxii," *Rgzm-tagungen*, pp. 89–130, 2009.
[221] L. Schmitt, S. Larsson, J. Burdukiewicz, J. Ziker, K. Svedhage, J. Zamon, and H. Steffen, "Chronological insights, cultural change, and resource exploitation

on the west coast of sweden during the late palaeolithic/early mesolithic transition," *Oxford Journal of Archaeology*, vol. 28, no. 1, pp. 1–27, 2009.

[222] H. Damlien, "Eastern pioneers in westernmost territories? current perspectives on mesolithic hunter–gatherer large-scale interaction and migration within northern eurasia," *Quaternary International*, vol. 419, pp. 5–16, 2016.

[223] H. Knutsson and K. Knutsson, "The postglacial colonization of humans, fauna and plants in northern sweden," *Arkeologi i norr*, vol. 13, pp. 1–28, 2012.

[224] B. M. Peter, "Admixture, population structure, and f-statistics," *Genetics*, vol. 202, no. 4, pp. 1485–1501, 2016.

[225] S. L. Cox, C. B. Ruff, R. M. Maier, and I. Mathieson, "Genetic contributions to variation in human stature in prehistoric europe," *Proceedings of the National Academy of Sciences*, vol. 116, no. 43, pp. 21 484–21 492, 2019.

[226] S. Wright, "Evolution in mendelian populations," *Genetics*, vol. 16, no. 2, p. 97, 1931.

[227] V. G. Childe, *The dawn of European civilization*. Routledge, 2013.

[228] A. J. Ammerman and L. L. Cavalli-Sforza, *The Neolithic transition and the genetics of populations in Europe*. Princeton University Press, 2014, vol. 836.

[229] L. L. Cavalli-Sforza, L. Cavalli-Sforza, P. Menozzi, and A. Piazza, *The history and geography of human genes*. Princeton university press, 1994.

[230] P. Menozzi, A. Piazza, and L. Cavalli-Sforza, "Synthetic maps of human gene frequencies in europeans: These maps indicate that early farmers of the near east spread to all of europe in the neolithic." *Science*, vol. 201, no. 4358, pp. 786–792, 1978.

[231] J. B. Aubán, L. M. Balaguer, M. A. E. Bebiá, J. R. O. Pérez, and J. d. D. B. Soler, "La cerámica impresa mediterránea en el origen del neolítico de la península ibérica," in *De Méditerranée et d'ailleurs...: mélanges offerts à Jean Guilaine*, 2009, pp. 83–96.

[232] G. P. Consortium, A. Auton, L. Brooks, R. Durbin, E. Garrison, and H. Kang, "A global reference for human genetic variation," *Nature*, vol. 526, no. 7571, pp. 68–74, 2015.

[233] M. E. Allentoft, M. Sikora, K.-G. Sjögren, S. Rasmussen, M. Rasmussen, J. Stenderup, P. B. Damgaard, H. Schroeder, T. Ahlström, L. Vinner *et al.*, "Population genomics of bronze age eurasia," *Nature*, vol. 522, no. 7555, pp. 167–172, 2015.

Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 2228

Editor: The Dean of the Faculty of Science and Technology

A doctoral dissertation from the Faculty of Science and Technology, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology. (Prior to January, 2005, the series was published under the title "Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology".)

Distribution: publications.uu.se
urn:nbn:se:uu:diva-493382

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2023