Human Mitochondrial DNA from 5th – 3rd Millennia BCE: An Analysis of Iberian Populations

Catherine Robinson

A thesis submitted in partial fulfilment of the requirements of Liverpool John Moores University for the degree of Master of Philosophy

April 2018

Abstract

Well-documented demographic changes accompanied the onset of the Neolithic but subsequent population dynamics are less clear. Ancient DNA data from the Neolithic period of coastal regions of the Iberian Peninsula are relatively scarce. The aim of the present study was to sequence DNA from Neolithic Iberians from the Spanish Levant (Catalonia) and the Portuguese Estremadura, Algarve and Alentejo regions in an attempt to explore the population dynamics in these regions following the adoption of agricultural practices.

Samples were collected from the Late Neolithic site 'La Sagrera' (N=13), situated in modern day Barcelona, and from four sites in Portugal; Castelo Belinho (Early Neolithic, N=8), Algar do Barrão (Middle Neolithic, N=3), and Casais da Mureta (N=3) and Anta da Cabeceira (N=1), both Late Neolithic. Mitochondrial DNA was extracted, amplified and sequenced. Haplogroup assignments were made using diagnostic SNPs. Databases of ancient and modern mitochondrial DNA were searched for haplotypes, and likelihood ratios were calculated to investigate kinship between individuals with shared haplotypes. Obtained sequences were grouped with previously published data and haplogroup frequencies were compared with other ancient populations, and genetic distances were calculated between ancient and modern Iberian populations.

Sequencing was largely unsuccessful for the Portuguese sites due to sub-optimal sample preservation. Regionally, ancient DNA data is relatively scarce, suggesting local climate may promote DNA degradation. Haplotypes were obtained for 10 of the 13 individuals from La Sagrera. Matrilineal kinship was established for two individuals sharing a haplotype. However, it was not possible to ascribe a kinship based burial pattern at this site. F_{ST} analysis indicated a discontinuity between pre- and post-Neolithic populations, though haplogroup analysis indicated a Mesolithic ancestry throughout the Iberian Neolithic. Middle and Late Neolithic populations were genetically divergent from modern Iberian populations, indicating that processes subsequent to the Neolithic period may have influenced the modern mitochondrial gene pool in the region. Haplogroup analysis indicated an increase in haplogroup diversity coinciding with the onset of the Neolithic, and a subsequent reduction in diversity into the Middle and Late Neolithic periods.

Acknowledgements

I would like to thank Professor Richard Brown for his guidance and encouragement, and Dr Eva Fernández for providing training, advice and support throughout my studies. I am grateful to Professor António Carvalho for providing the Portuguese samples and advice regarding the archaeology of the sites, and Professor Miquel Molist for providing samples from La Sagrera. I would also like to thank Professor Terry Brown at the University of Manchester for hosting me at his lab while I completed the independent replication of the samples. Finally, I owe enormous gratitude to John Nicholson for his endless encouragement and patience; I might not have made it without your support.

Contents

List of Tables

List of Figures

Chapter One: Introduction

1.1 Ancient DNA

1.1.1 History of ancient DNA

The field of ancient DNA (aDNA) began with the successful sequencing of the extinct zebra-like quagga (Higuchi et al, 1984). Early studies (e.g. Higuchi et al, 1984; Pääbo, 1985) found that endogenous DNA was restricted to short, highly degraded fragments primarily of multi-copy loci such as mitochondrial DNA (mtDNA), and samples contained large quantities of exogenous fungal or bacterial DNA. The invention of polymerase chain reaction (PCR) (Mullis et al, 1986) suddenly made it possible to amplify DNA from as little as a single surviving copy. This led to a number of publications purporting sequences from millennia old specimens (e.g. Poinar, Cano and Poinar Jr, 1993; Woodward, Weyand and Bunnell, 1994). Many of these studies were later found to result from exogenous contamination (e.g. Hedges and Schweitzer, 1995) or were simply not reproducible (e.g. Austin et al, 1997). Though PCR enabled amplification and subsequent study of minute quantities of surviving endogenous aDNA, it also increased the sensitivity of such studies to contaminant DNA, which was particularly problematic for human aDNA studies (Stoneking, 1995). This led to an examination of ancient DNA protocols, and criteria were laid out for the authentication of human aDNA samples (Cooper and Poinar, 2000), enabling application of aDNA analyses to a broad range of questions in fields including archaeology and anthropology.

1.1.2 Non-recombining DNA

Mitochondrial DNA (mtDNA) and the non-recombining portion of the Y chromosome, inherited maternally and paternally respectively, can be referred to as lineage markers. They undergo uniparental inheritance, passing unchanged from parent to offspring. They are haploid in nature and do not undergo recombination, so their diversity results from mutations which occur during DNA replication. Mutations take the form of single nucleotide polymorphism (SNPs), insertion/deletion polymorphisms (indels) and short tandem repeats (STRs). This makes them useful for tracing historic migrations (e.g. Haak et al, 2015), studying phylogeography (e.g. Richards et al, 1998), and determining kinship (e.g. Haak et al, 2008). They are also used to investigate genetic continuity between

putative ancestral and descendent populations (e.g. Fehren-Schmitz et al, 2010; Fehren-Schmitz et al, 2011).

Mitochondrial DNA has several characteristics that make it useful in aDNA studies. It is contained within the mitochondria of eukaryotic cells. Each cell contains many thousands of copies of mtDNA (Robin and Wong, 1988). Mitochondrial DNA is a circular molecule consisting of typically 16569bp, and contains 37 genes that are necessary for mitochondrial function. The genes are contained within the coding region of the mtDNA molecule. In addition, the non-coding region of the mtDNA, known as the "control region", is organised in three hypervariable regions; hypervariable region I (base positions 16024-16569), hypervariable region II (base positions 1-340) and hypervariable region III (base positions 438-574) (HVRI, HVRII and HVRIII respectively). Mitochondrial DNA has a higher mutation rate than nuclear DNA (Pakendorf and Stoneking, 2005), and in humans these mutations are observed more frequently at these hypervariable regions (Stoneking, 2000) (although rates are quite heterogeneous even within the hypervariable regions: Heyer et al, 2001).

Mutations in an mtDNA sequence are commonly identified by comparison with the revised Cambridge Reference Sequence (rCRS) (Anderson et al, 1981; Andrews et al, 1999). The rCRS was the first human mitochondrial genome to be published and is a member of European haplogroup H. An alternative to the rCRS is the Reconstructed Sapiens Reference Sequence (RSRS) (Behar et al, 2012), though use of this has yet to be widely adopted. The mutations constitute a haplotype, which can be grouped into monophyletic clusters called haplogroups. Haplogroups represent the evolution of mtDNA lineages from a single common matrilineal ancestor (Behar et al, 2008). Van Oven and Kayser (2009) constructed a comprehensive phylogeny of the available mitochondrial genomes: phylotree (www.phylotree.org). A simplified version of this phylogenetic tree can be seen in figure 1. These characteristics – the high mutation rate, non-recombining nature, multicopy, uniparental inheritance – make mtDNA a useful target in aDNA studies.

Figure 1. Simplified version of the human mitochondrial phylogenetic tree showing major haplogroup lineages (van Oven and Keyser, 2008)

1.2.3 Characteristics of ancient DNA

Many ancient specimens contain no endogenous DNA. Those that do characteristically have low numbers of highly fragmented, chemically altered endogenous molecules (Pääbo, 1989). Post-mortem, the protective mechanisms which repair molecular damage in life cease, leading to changes in the strucure and integrity of the DNA molecule. Hydrolytic damage such as cytosine deamination (Figure 2) (Hofreiter et al, 2001), depurination and depyrimidination (Lindahl, 1993), as well as oxidative damage resulting from ionizing radiation (Höss et al, 1996) can all result in strand fragmentation and incorporation of miscoding lesions into a sequence (Figure 3). Other types of damage such as cross-links between molecules also occur (Allentoft et al, 2012). Use of overlapping primer pairs to target short fragment lengths (Rizzi et al, 2012) improves amplification of degraded DNA, while miscoding lesions can be identified through cloning (Lamers, Hayter and Matheson, 2009).

Figure 2. Deamination of cytosine to uracil (Brown and Brown, 2011)

Figure 3. Generation of C to T sequence error during PCR resulting from cytosine deamination (Brown and Brown, 2011)

Theoretically, a relationship between sample age and DNA degradation should be observable. Allentoft et al (2012) estimated the half-life of DNA molecules to be 521 years, meaning that in optimum conditions, DNA up to 1.5 million years old could be sequenced. In practice, however, this is often not the case, with depositional environment being more important (Gamba et al, 2008). Stable, low temperatures are key for DNA survival, and the oldest authenticated DNA sequences tend to be from

permafrost environments (e.g. Orlando et al, 2013). Rapid post-mortem desiccation and high salt concentration are also thought to play a role (Willerslev and Cooper, 2005). Conversely, high or fluctuating temperatures and humidity are all likely to inhibit preservation.

Since the beginning of aDNA studies, exogenous DNA contamination has been a problem (see above). Not only is modern contaminant DNA likely to be more abundant in a sample, but it will be amplified preferentially to the degraded exogenous target if present (Fulton, 2012). Contamination is particularly problematic when manipulating human samples, since the researchers themselves are a contamination source. To overcome problems arising from contamination, many criteria for authenticating aDNA results have been proposed, including physical separation of aDNA workspaces; extraction and PCR controls; independent replication, preferable from an independent sample and in a separate laboratory; molecular cloning; and ensuring phylogenetic sense of sequences (Cooper and Poinar, 2000; Fulton, 2012).

1.2 The Neolithic

1.2.1 Neolithic Transition in Europe

The Neolithic period of human prehistory is characterised by the shift from a nomadic, hunter-gathering lifestyle common during the Mesolithic, to a reliance on agricultural subsistence. Originating in the Fertile Crescent around 10 000 BCE, farming technologies quickly spread westward throughout the rest of Europe (Price, 2000).

Likely arising in response to increasing population size (Soares et al, 2010), the Neolithic brought with it a sharp increase in the scale and complexity of social networks, which led to changes in settlement patterns, population, and social inequalities (Aubán, Lozano and Pardo-Gordó, 2017). The Neolithic transition affected economic, social, and demographic change, revolutionising human societies across Europe (Aubán, Martín and Barton, 2013). These changes were reflected in increased symbolism in cultural practices and funerary rites.

The archaeological record suggests the advance of the Neolithic into Europe occurred predominantly along two routes; the Impressa culture spread along the Mediterranean

coastline starting around 5900 BCE (before current era), and the simultaneous spread of the Linearbandkeramik (LBK) culture into Central Europe along the Danube River (Olalde et al, 2015). There are two contrasting models for how the Neolithic spread through Europe. Ammerman and Cavalli-Sforza (1984) found an east-west cline in the distribution of allele frequencies and used this information alongside radiocarbon dates to propose a human migration from the Near East to Europe. According to the demic diffusion model, migrant farmers spread and replaced autochthonous Mesolithic communities. A complete demic diffusion should be characterised by a discontinuity in the genetic structure between Neolithic and pre-Neolithic populations (Ammerman and Cavalli-Sforza, 1984). Conversely, the cultural diffusion model suggests a sharing of culture and trade between neighbouring hunter-gatherer and farming communities, allowing the spread of farming technologies without population replacement (e.g. Dennell, 1983). Cultural diffusion, therefore, would result in a genetic continuity from the Mesolithic to the Neolithic period.

The reality is likely to lie somewhere between these two extremes, with varying levels of migration and cultural diffusion seen in different areas. It is also unlikely that the Neolithisation of Europe occurred in a single chronological event. Rather multiple migratory events, spanning thousands of years are thought to have taken place out of the Fertile Crescent (Özdoğan, 2014), with various levels of admixture between existing Mesolithic communities and colonising Neolithic migrants. We can begin to better understand these processes by uncovering the genetic structure of prehistoric European communities, which has helped to identify patterns of migration and colonisation (Soares et al, 2010).

While the model proposed by Ammerman and Cavalli-Sforza (1984) suggested a uniform ‘wave of advance’ of Neolithic famers, both molecular and radiocarbon dating have indicated a series of shorter, local migrations punctuated by long pauses as a likely model (Rowley-Conwy, 2009). This model allows for the incorporation of Mesolithic signatures into farming communities before further ‘leapfrog’ colonisation. Pioneer colonisation, involving settlement of a few newcomers and a greater contribution from autochthonous populations has also been suggested (Zvelebil, 1986).

1.2.2 Neolithic Iberia

Within the European landscape, the Iberian Peninsula presents an interesting case study for the Neolithic Transition, due to its relatively segregated position at the southwestern tip of Europe. Much archaeological information has been lost due to inundation of prehistoric shore lines (Olalde et al, 2015), making the arrival and spread of the Neolithic in this region more difficult to understand. Cardial culture – a later incarnation of the Impressa - reached the Iberian coast around 5500BCE (Olalde et al, 2015). The subsequent rapid spread of the Cardial culture along the Iberian coastline is thought to be the result of maritime pioneer colonisation of uninhabited regions (Zilhao, 2001; Gamba et al, 2012). This may have resulted in a patchwork of communities. Zilhao (1998) suggests a coexistence between Mesolithic and Neolithic populations lasting several millennia along the Mediterranean coast, although some sites in Catalonia show an absence of Mesolithic influence (Ribe, Cruells and Molist, 1997), suggesting that co-existence between farmers and autochthonous populations was not uniform. Other archaeological evidence points to terrestrial expansion routes along rivers or across the Pyrenees (García-Martínez de Lagrán, Fernández-Domínguez and Rojo-Guerra, 2017).

1.3 Archaeogenetics

1.3.1 Pre-Neolithic Europe

A genetic bottleneck during the Last Glacial Maximum (LGM) in Europe reduced mitochondrial diversity (García-Martínez de Lagrán, Fernández-Domínguez and Rojo-Guerra, 2017), including the loss of haplogroup M lineages (Fu et al, 2016; Posth et al, 2016). This led to a relatively uniform Mesolithic population characterised by haplogroups U, U2, U4, U5, and U8 (Bramanti et al, 2009; Gamba et al, 2014; Lazaridis et al, 2014). Haplogroups U5, U8 are thought to be the most ancient European haplogroups, and are the only haplogroups - along with V (Torroni et al, 2001; Soares et al, 2010) – that arose in situ. All others were likely to have been imported during migration events (Richards et al, 2000; Achilli et al, 2004; Malyarchuk et al, 2010). Haplogroups H and possibly lineages belonging to K, T, W, and X may have arrived in Europe prior to the LGM and suffered reduced diversity with the LGM, before re-expanding post-LGM (Richards et al, 2000; Achilli et al, 2004).

1.3.2 Neolithic Europe

The arrival of the Neolithic in Central Europe approximately 5500 BCE was characterised by the presence of haplogroups HV, J, K, N1a, T2, V, W and X (Haak et al, 2010; Brandt et al, 2013). This distinct group of haplogroups appears later in southern Scandinavia (Brandt et al, 2013), coinciding with the arrival of the Neolithic in the region (Haak et al, 2015). Haplogroups H, T1, U2, U3, U4 and U5a have also been identified in Central European early farmers (Haak et al, 2005; Bramanti et al, 2009; Brandt et al, 2013; Szécsényi-Nagy et al, 2015). Northern and Central European early farmers were found to be genetically distinct from the region's hunter-gatherers but similar to each other and to subsequent Neolithic populations (Brandt et al, 2013; Szécsényi-Nagy et al, 2015), whereas a greater proportion of hunter-gatherer lineages persist through the Neolithic in Scandinavia (Brandt et al, 2013).

After an initial decline, Mesolithic ancestry increased during the Middle and Late Neolithic in Central and Eastern Europe (Bollongino et al, 2013; Haak et al, 2015), and there is evidence for the existence of distinct hunter-gatherer groups in Late Neolithic Central Europe (Bollongino et al, 2013). Discontinuity has been found between Late Neolithic and modern day Central Europeans (Brandt et al, 2013), and there is evidence for further migrations into Europe from the Caucasus during the Late Neolithic and Early Bronze Age (Allentoft et al, 2015; Haak et al, 2015). Allentoft et al (2015) found that the modern European genetic landscape most closely resembles that of Bronze Age Europe.

1.3.3 Neolithic Iberia

Although data is scarce, Mesolithic Iberian populations appear to have been characterized by haplogroup U (Hervella et al, 2012; Sánchez-Quinto et al, 2012). Additionally, a high frequency of haplogroup H is observed in Mesolithic Iberia (de-la-Rua et al, 2015), probably resulting from post-LGM expansion from Franco-Cantabrian/Iberian Refugia (Achilli et al, 2004). Although haplogroup H was also present in pre-Neolithic Iberia, HVRI is not sufficiently resolved to distinguish between lineages, some of which may have arrived with the Neolithic.

Discontinuity between early farmers and later Neolithic populations has been observed in the Catalonia and Aragon regions of the Iberian Peninsula (Gamba et al, 2012).

Early Neolithic populations in northeastern Iberia contained individuals belonging to haplogroups H, HV, I, J, K, N, N*, T2, U, U5, X1 and V (Lacan et al, 2011; Gamba et al, 2012; Hervella et al, 2012; Haak et al, 2015; Olalde et al, 2015), while haplogroups H3 and H4, and H and K have been found at Early Neolithic sites in southwestern Iberia and southeastern Iberia, respectively (Olalde et al, 2015).

Haplogroups H, I, J1, T2, U, U5 and W have all been confirmed in northeast Iberia (Sampietro et al, 2007; Hervella et al, 2012), and haplogroups H1, J2, K and U in central Iberia (Haak et al, 2015) during the Middle Neolithic. Late Neolithic central Iberian are characterized by H, J, K, T2, U, V and X lineages (Gómez-Sánchez et al, 2014; Günther et al, 2015; Mathieson et al, 2015; Alt et al, 2016)

The transition from the Early to the Middle Neolithic saw an increase in hunter-gatherer ancestry in the farming communities of Iberia (Alt et al, 2016; Szécsényi-Nagy et al, 2017). The presence of hunter-gatherer linages may be the result of admixture between contemporaneous Neolithic and Mesolithic groups on the Iberian Peninsula (Lipson et al, 2017) or it may be that Mesolithic lineages were brought into the region by migrating populations (Martiniano et al, 2017). By the end of the Neolithic, Szécsényi-Nagy et al (2017) suggest a relatively homogenous genetic landscape throughout the Iberian Peninsula, indicating that human mobility had increased and genetic mixing was common.

1.4 Aims and Objectives

The present study aims to sequence mitochondrial DNA from skeletons excavated from Neolithic burial sites across the Iberian Peninsula in order to add to the genetic data available from this region for the Middle and Late Neolithic periods. The data obtained will be used to:

i) Investigate maternal kinship within the burial sites studied, and subsequently make inferences about the social complexities underlying them.
ii) Investigate the population dynamics of Neolithic Iberians by exploring the occurrence and frequency of the mitochondrial haplotypes of the studied populations in earlier and later populations.

iii) Explore genetic divergence between Mesolithic, Early, Middle and Late Neolithic Iberian populations.

iv) Explore genetic divergence between Middle and Late Neolithic, and extant Iberian populations.

Chapter Two: Materials and Methods

2.1 Study Sites

Skeletal and dental samples were obtained for analysis from five archaeological sites on the Iberian Peninsula (Figure 4). The sites ranged from the Early Neolithic to the Late Neolithic (Table 1).

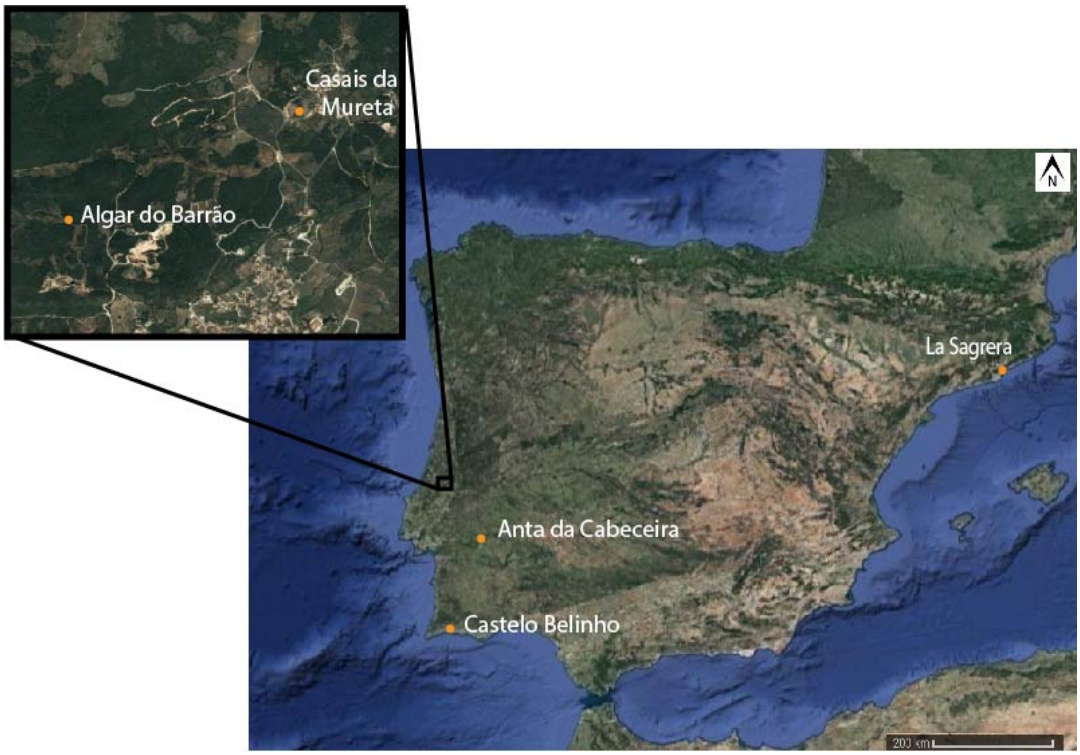


Figure 4. Location of sites from which studied samples were excavated

Table 1. Location, type, and calibrated radiocarbon dates (cal BCE) of the archaeological sites studied.

Site	Region	Type	Date (cal BC)	N
La Sagrera	Barcelona, Catalonia	Hypogeum	2403-2227	13
Castelo Belinho	Algarve, Portugal	Settlement	4500-3900	8
Anta da Cabeceira	Alentejo, Portugal	Dolmen	3600	1
Algar do Barrão	Estremadura, Portugal	Cave	3700-3100	3
Casais da Mureta	Estremadura, Portugal	Cave	3320-2930	3

2.1.1 La Sagrera

Construction works being undertaken at the metro station in the metropolitan borough of La Sagrera in Barcelona in 2010 uncovered a mass grave of human remains (Figure 5).

Upon excavation between 2011 and 2012, the site was found to be a 5.5m x 3m hypogeum, containing the skeletal remains of 207 individuals, which was dated to the Late Neolithic (2403-2227cal BCE) (Balaguer et al, 2013).

Figure 5. Hypogeum of Late Neolithic human remains uncovered in 2010 during construction works at La Sagrera metro station, Barcelona. Balaguer et al. 2013

The hypogeum appears to have three discrete burial stages, with the first stage being characterised as a collective burial, and the third stage being a primary deposition of a single individual (Balaguer, 2016). The samples of interest in the present study are from the second burial stage, which was characterised by a large number of inhumations over a relatively short period. Despite the apparent haste with which these depositions were made, the burials display characteristic funerary rituals, with individuals being positioned in a specific way, either prone with legs bent laterally to the left, or in right lateral decubitus with legs bent, i.e. in the foetal position. The skeletons were also arranged into distinct clusters (Figure 6) (Balaguer, 2016). Thirteen individuals were selected for analysis from three clusters identified as have been almost simultaneously deposited (Table 2). The individuals sampled from the clusters were selected based on macroscopic preservation, and thus the likelihood of successful DNA amplification.

Figure 6. Burial clusters from the second depositional phase at the La Sagrera hypogeum, showing decubitus (e.g. individual SA8 in cluster A) and prone (e.g. individual SA18 in cluster B) positioning of the skeletons. Cluster A contained individuals SA2 (green) (not analysed), SA4 (yellow), and SA8 (orange). Cluster B contained individuals SA18 (blue), SA56 (pink), and SA68 (yellow). Cluster C contained individuals SA43 (blue), SA44 (green), SA46 (pink), SA45 (yellow) SA53 (orange), and SA25, SA26, and SA39 (not shown). Pictures from Balaguer (2016).

Table 2. Age, sex and burial cluster of the skeletons from the La Sagrera hypogeum that were subject to aDNA analysis (Balaguer, 2016)

Cluster	Individual	Age Category	Age	Sex
A	SA4	Child	8-10	Indeterminate
	SA8	Young adult	>25	Female
B	SA18	Mature adult	45	Male
	SA56	Adult	?	Male
	SA68	Young adult	18-23	Female
C	SA25	Child I	18 months	Indeterminate
	SA26	Child II	5-6	Indeterminate
	SA39	Child II	8	Indeterminate
	SA43	Juvenile	17	Female
	SA44	Juvenile/Young adult	18-22	Male
	SA45	Juvenile/Young adult	18-24	Male
	SA46	Young adult	25-34	Male
	SA53	Adult	?	Female

2.1.2 Portugal

The four Portuguese sites date from the Early to the Late Neolithic (Table 1).

Castelo Belinho

Radiocarbon dating places Castelo Belinho at the transition period from the Early Neolithic to the Middle Neolithic (Table 1) (for syntheses of site, see Gomes, 2008; Gomes, 2012). Situated between the foothills of the Monchique Mountain and the estuary of the Arade River, in the Algarve region, Castelo Belinho is an Early/Middle Neolithic settlement site with associated funerary pits. This is unique amongst sites studied here. It represents the later practices of individual burials in southwestern Iberia before the emergence of collective burials typical of the megalithic culture. A rich material culture was excavated at this site, including potsherds, tools, jewellery, and faunal remains. Archaeologists definitively identified 14 graves at the site, and samples from eight individuals excavated from these were provided for analysis (Table 3). Graves varied in size from small dug pits or natural depressions, to larger silo-shaped pits. Bodies were arranged either in the foetal position, or dorsally with legs bent backwards.

Algar do Barrão

Algar do Barrão is a burial cave located in the eastern rim of the Limestone Massif of Estremadura, in central-coastal Portugal, facing the Tagus river valley. It was dated to the Middle Neolithic (3700-3100 cal BCE) (Carvalho and Petchey, 2013). The remains of at least 21 individuals were discovered during excavations, along with a small selection of artefacts, including shell bracelets and pottery. Males, females, adults and juveniles were all represented within the remains. Remains were highly fragmented, though anthropologists found no sign of trauma. It was suggested that the remains found at this site are the result of secondary depositions, with the corpses having been subject to a primary deposition in order to remove the flesh and then transferred along with whitish sediments (thus contrasting in colour with the reddish, clayish sediments of the cave) to their current location. These burial practices and artefacts observed at Algar do Barrão are comparable to those observed at other Middle Neolithic and younger sites (Carvalho, Antunes-Ferreira and Valente, 2003).

Casais da Mureta

Casais da Mureta is a very narrow, vertical cave located around 2 km away from Algar do Barrão. It was uncovered after building works, and only very preliminary salvage works have been carried out. Thus, it remains unpublished. The cave floor was scattered with an abundant faunal assemblage (Valente et al, accepted for publication) and human skeletal remains, including three human mandibles containing more than three teeth, which appeared suitable for genetic analysis (Table 3). This site is estimated to date from the Late Neolithic (3320-3942 cal BCE – Antonio Carvalho: personal communication). The site provides evidence of the same type of funerary practices as those observed at Algar do Barrão, but the particular topography of the cave and the unexpected number of animal remains may reveal a different scenario as work progresses.

Anta da Cabeceira

Anta da (or Dolmen of) Cabeceira is a granite-built dolmen, excavated in the 1930s by a former director of the National Museum of Archaeology (Lisbon), M. Heleno. This dolmen is located in the Mora area of the Alentejo region, where an impressive number of megalithic monuments have been found since the beginning of the 20th century. Given the acidity of the granitic and conglomerate geological substrata, bone preservation conditions are seldom found. This was not the case at Cabeceira, where a minimum number of three individuals were found and recently dated to 3600 cal BCE (Carvalho and Rocha, 2016): one of the earliest dates for a human sample from a megalithic structure in Portugal. Given the poor preservation conditions at the site, only one individual was sampled in the first instance to determine whether successful DNA retrieval might be possible. Since it was not possible to obtain a reproducible sequence from this individual, the decision was taken not to analyse any further samples from this site.

2.2 Sample Selection

The anthropologists responsible for the skeletons selected samples. With the exception of individuals SA43, CBL5, and CBL7, two samples per individual were provided (Table 3), and sent to Liverpool John Moores University (LJMU). Ideally, samples were required to show good macroscopic preservation, with few visible fissures and natural colouration

(Figure 7). Dental samples were preferred, so teeth were provided in all possible cases. Exceptions were samples 1CBL3, 2CBL3, 2CBL6, 1CBL7 and CAB3, where bone fragments were provided due to lack of suitable dental samples (Table 3). In order to rule them out as a possible contamination source, the anthropologists responsible for sample selection were required to provide a DNA sample for genotyping.

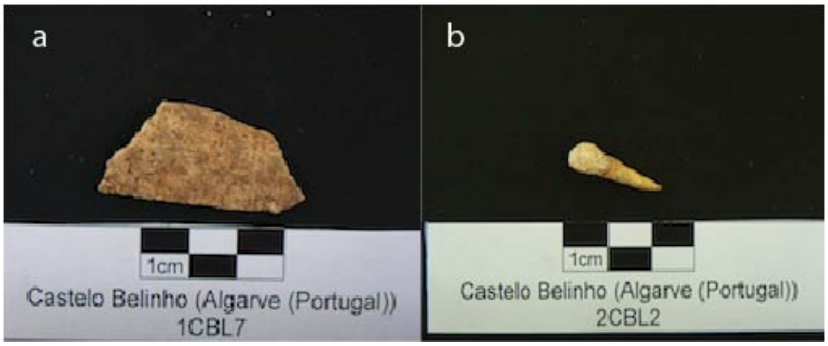


Figure 7. Examples of preferred characteristics of samples selected for ancient DNA analysis. Samples should be of natural colouration, and should have good macroscopic level of preservation with no obvious fissures or damage. The skull fragment (a), though a fragment in itself, has damage to the external surfaces. The incisor (b) is entire with no broken roots or holes in the crown (Robinson, 2015)

All samples were documented on arrival; sample type (i.e. dental sample or bone type) was confirmed, preservation condition and colour noted, and photographs taken (Table 3). Sample pictures can be found in Appendix 1.

Table 3. Samples provided for ancient DNA analysis

Site	Individual	Sample	Sample type	Colour	Preservation
La Sagrera	SA4	1SA4	Tooth	Brown	Roots broken, dirty
		2SA4	Molar	Brown	Root broken
	SA8	1SA8	Tooth	Brown	Surface damaged
		2SA8	Molar	Pale orange	Dirty, enamel broken
	SA18	1SA18	Tooth	Pale brown	Root broken
		2SA18	Molar	Pale orange	Root broken
	SA25	1SA25	Deciduous molar	Natural	Roots destroyed
		2SA25	Deciduous molar	Natural	Roots destroyed
	SA26	1SA26	Tooth	Pale brown	Root broken
		2SA26	Deciduous molar	Pale brown	Root broken
	SA39	1SA39	Tooth	Natural	Roots broken and dirty
		2SA39	Molar	Orange-brown	Roots broken
	SA43	1SA43	Molar	Pale brown	Entire, good
	SA44	1SA44	Tooth	Natural	Entire, good
		2SA44	Molar	Pale orange	Entire, good
	SA45	1SA45	Tooth	Natural	Entire, good
		2SA45	Molar	Pale brown	Roots broken
	SA46	1SA46	Tooth	Natural	Entire, good
		2SA46	Molar	Orange-brown	Root broken
	SA53	1SA53	Tooth	Natural	Root broken
		2SA53	Molar	Brown	Entire, root fissure
	SA56	1SA56	Tooth	Orange-brown	Entire, good
		2SA56	Molar	Orange-brown	Cracked
	SA68	1SA68	Tooth	Brown	Entire, root fissure
		2SA68	Molar	Orange-brown	Entire, good
Algar do Barrão	BAR3	1BAR3	Molar	Orange-brown	Entire, weathered
		2BAR3	Molar	Pale orange	Entire, weathered
	BAR4	1BAR4	Molar	Brown	Root broken, weathered
		2BAR4	Molar	Orange-brown	Entire, weathered
	BAR5	1BAR5	Molar	Natural	Root broken, dirty
		2BAR5	Molar	Brown	Root broken, dirty
Anta da Cabeceira	CAB3	1CAB3	Femur section	Brown	Surface weathered
		2CAB3	Femur section	Orange-brown	Surface weathered
Casais da Mureta	CMR1	1CMR1	Molar	Natural	Chipped enamel
		2CMR1	Molar	Brown	Broken enamel
	CMR2	1CMR2	Canine	Natural	Broken enamel
		2CMR2	Premolar	Brown	Entire, good
	CMR3	1CMR3	Molar	Orange-brown	Cracked
		2CMR3	Premolar	Orange-brown	Entire, good
Castelo Belinho	CBL1	1CBL1	Tooth	Natural	Entire, good
		2CBL1	Tooth	Natural	Entire, good
	CBL2	1CBL2	Incisor	Natural	Cracked
		2CBL2	Tooth	Yellow root	Enamel chipped
	CBL3	1CBL3	Bone fragment	Natural	Dirty
		2CBL3	Skull fragment	Natural	Highly degraded
	CBL4	1CBL4	Tooth	Orange-brown	Highly degraded
		2CBL4	Tooth	Orange-brown	Highly degraded
	CBL5	1CBL5	Tooth	Brown	Highly degraded
	CBL6	1CBL6	Canine	Pale brown	Highly degraded
		2CBL6	Femur fragment	Natural	Highly degraded
	CBL7	1CBL7	Bone fragment	Pale brown	Degraded
	CBL8	1CBL8	Tooth	Orange-brown	Highly degraded
		2CBL8	Tooth	Orange-brown	Highly degraded

2.3 Genetic analyses

Detection of contaminant DNA is an essential component of ancient DNA analyses. This is of particular importance when analysing ancient human samples, since there is a high likelihood of contamination with modern human DNA at all stages of processing, from excavation to PCR. Due to its similarity, modern contaminant DNA may be amplified alongside, or preferentially to, the targeted endogenous DNA, resulting in an inauthentic sequence. Thus, ancient DNA analysis is subject to strict protocols and should be undertaken in isolated facilities (Cooper and Poinar, 2000). Criteria of authenticity must be adhered to in order to guarantee only genuine endogenous sequences are reported. Facilities and authenticity criteria can vary among ancient DNA laboratories, depending upon available resources and preferred methodologies.

2.3.1 Facilities and Criteria of Authenticity

A purpose-designed Low Copy Number (LCN) DNA laboratory was set up at Liverpool John Moores University (LJMU) prior to the start of the current study. Only ancient samples were permitted in the LCN facilities. A restricted-access policy was in place, and researchers wishing to access the facilities were required to provide a DNA sample for genotyping (Table 4).

Table 4. Mitochondrial hypervariable region I haplotypes of researchers associated with samples or aDNA facilities at both Liverpool John Moores University (LJMU) and Manchester Institute of Biotechnology (MIB).

Staff	Details	Extraction lab	HVI
L00	User LCN-LJMU	Barcelona	CRS
L01	User LCN-LJMU	LJMU	16093C 16221T
L02	User LCN-LJMU	LJMU	16263T
L03	User LCN-LJMU	LJMU	16126C 16294T 16304C
L06	External	LJMU	16192T 16274A 16362C
L07	External	Madrid	16145A 16176G 16223T 16290A
M01	User LCN-MIB	Madrid	16037G 16192T 16256T 16270T 16311C 16399G
M02	User LCN-MIB	MIB	16069T 16126C 16519C
M03	User LCN-MIB	MIB	16356C 16519C
M06	User LCN-MIB	MIB	16093C 16224C 16311C 16362C 16400T 16519C
M04	User LCN-MIB	MIB	CRS
M05	User LCN-MIB	MIB	16093C 16223T 16291T 16295T 16337T 16362C 16519C

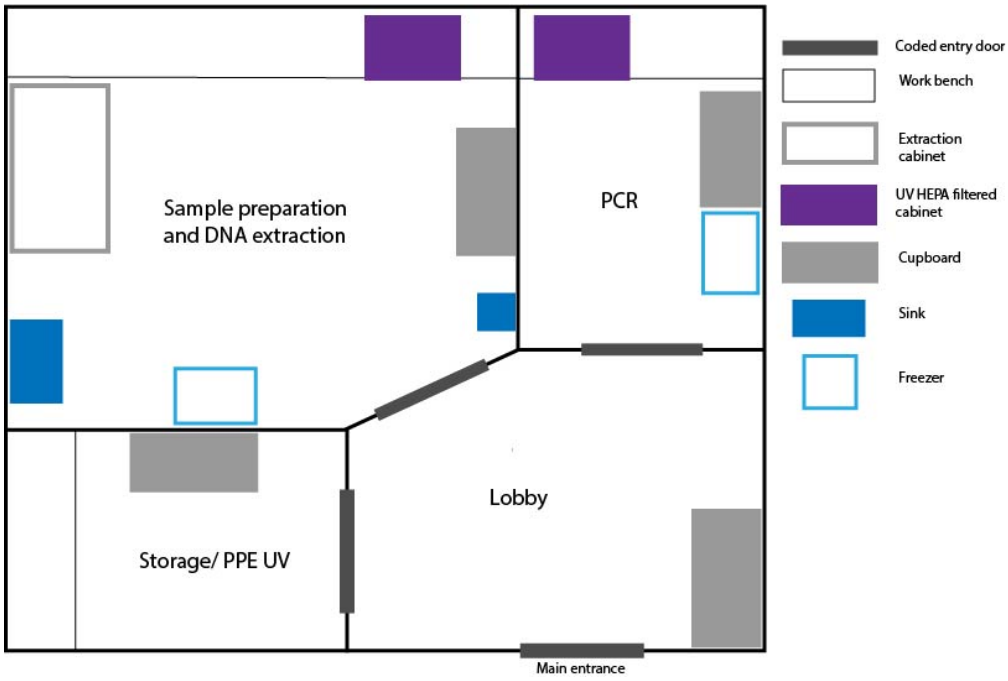


Figure 8. Layout of LCN facilites at Liverpool John Moores University

The LCN facilities, located in the Life Science Building (LSB) at LJMU were separated from the main molecular biology laboratories, and consisted of an extraction laboratory and a PCR laboratory with a connecting lobby area (Figure 8). A unidirectional workflow was maintained between the laboratories meaning that in a single working day, researchers could only move from the extraction room to the PCR room and from there to the post-PCR laboratory.

Both the laboratories were fitted with 254nm ultra-violet lamps, which were switched on overnight prior to and after use of the facilities. All laboratory surfaces and equipment were cleaned with 70% sodium hypochlorite bleach before and after use, as were reagent bottles and other lab ware. Most reagents and all lab ware and consumables were also irradiated in ultraviolet cross linkers for between 15 minutes and 3 hours to ensure surfaces were sterile prior to use. The insides of sample tubes and their lids were also UV irradiated for at least 20 minutes prior to use.

Figure 9. Low copy number facilities situated in the Life Science Building at LJMU: entrance lobby (left) and sample preparation and DNA extraction room (right) (Pictures taken from Robinson et al, 2016)

Sample preparation, including cutting, cleaning, and grinding, was undertaken in an extraction cabinet in the extraction laboratory (Figure 9). Reagent preparation and extractions were carried out in the same room inside a UV HEPA PCR cabinet (UPV-Analytik Jena) in the same laboratory. PCR reactions were prepared in a UV HEPA-filter cabinet in the PCR laboratory and then taken to a thermal cycler outside of the LCN for the reaction to be performed. All post-PCR work was performed in the general laboratories.

Some extractions and PCRs were done at the Manchester Institute of Biotechnology (MIB), The University of Manchester. The ancient DNA facilities at MIB were subject to the same rigorous criteria as those at LJMU. In this case, extraction, PCR and post-PCR laboratories were in separate locations throughout the building.

When working in the laboratory, researchers were required to wear full personal protective equipment in order to prevent DNA contamination. This included wearing a hygiene mask, hairnet, hooded coveralls, shoe covers, protective goggles and two pairs of surgical gloves. Two independent extractions per individual were performed, ideally from two separate skeletal elements. When possible these were conducted in different laboratories (LJMU and MIB).

2.3.2 Modern DNA typing

Genetic profiles were obtained from all researchers involved in handling or selecting samples, and those with access to the LCN facilities. When researchers had not already been genotyped, they were asked to provide a swab of cheek cells from which DNA was extracted in the molecular biology labs at LJMU. Modern DNA samples were extracted using Nucleospin XS Tissue Kit according to the manufacturer's instructions, and PCR reactions were prepared using 45µl vWR Taq DNA Polymerase Master Mix 1.1x, 1µl each of primers forward primer 16069aF (5'-CTCCACCATTAGCACCCAAAGC-3') (Bramanti et al, 2009) and reverse primer H16380 (table 6), and 3µl DNA extract. The thermal cycler was programmed to run for 2 minutes at 95°C, followed by 34 cycles of 30 seconds at 95°C, 30 seconds at 60°C, and 30 seconds at 72°C, and a final extension stage of 1 minute at 72°C. PCR products were purified using NucleoSpin Gel and PCR Clean-up Kit following the manufacturer's instructions. PCR products were sequenced by Macrogen Europe. To comply with the Human Tissue Act 2004 samples were assigned a code, and personal details were separated from genetic profiles. It was a requirement that individuals freely consented to the use and storage of their genetic material within the scope of the project as set out in the participant information sheet provided.

2.3.3 Sample Preparation

Liverpool John Moores University

Both bone and dental samples were processed at LJMU, and were treated in the same way. Samples were cleaned prior to powdering to remove potential external contaminant DNA, and also to remove external dirt which may contain PCR inhibitors. Cleaning and grinding protocols differed between labs. At LJMU, surface dirt was removed by blasting with 90µm aluminium oxide particles using a Base 3 pneumatic sandblaster for approximately one minute, before the sample was transferred to the UV crosslinker to be irradiated for 15 minutes per side in order to break down any remaining contaminant DNA.

The entire sample was placed inside a grinding vial comprising a polycarbonate cylinder, stainless steel end plugs, and a stainless steel impactor bar. The assembly was placed in a SPEX 6775 Freezer/Mill® cryogenic grinder cooled with liquid nitrogen. The sample was

pulverised by magnetically shaking the stainless steel impactor bar against the end plugs for 30 seconds. If samples were not completely powdered, they were pulverised for a further 30 seconds. Powdered sample was then transferred to a sterilised 50ml Falcon Tube in the UV HEPA-filter cabinet and stored at -20°C. The used grinding vials were then disassembled and cleaned rigorously. Polycarbonate tubes were scrubbed with hot water and detergent, rinsed with sterile water and dried. Stainless steel components were scrubbed with hot water and detergent, rinsed with Milli-Q water, flooded with 70% sodium hypochlorite, rinsed again with Milli-Q water, rinsed with 100% ethanol and finally rinsed with sterile water and dried with paper towels. All components were then placed inside the cross linker and UV irradiated for three hours.

During the grinding process, samples 1SA18 and 1SA46, both dental samples, became lodged between the steel impactor bar and the side of the polycarbonate tube. The entire assemblies were transferred to the UV HEPA-PCR cabinet, an end plug removed and sterilised metal forceps used to manipulate the samples free. The end plugs were replaced and grinding recommenced.

Prior to beginning extraction, the required amount of sample (approximately 250mg for Rohland and Hofreiter protocol (Rohland and Hofreiter, 2007; Rohland, Siedel and Hofreiter, 2010), and approximately 50-100mg for Dabney protocol (Dabney et al, 2013) was transferred to new sterilised Falcon tubes in the HEPA-filter cabinet.

Manchester Institute of Biotechnology

Only dental samples were processed at MIB. Samples were cleaned by soaking in 70% sodium hypochlorite bleach for five minute, rinsed with Milli-Q water, coated in 37% phosphoric acid for one minute, and finally rinsed with Milli-Q water, to ensure all surface contamination was removed. An area was prepared for sample powdering by lining a corner of the bench and wall with tinfoil. A tinfoil tray was then created to catch tooth powder for extraction. The foil tray was cleaned with DNA Away ™ Surface Decontaminant and allowed to dry completely before powdering was undertaken. Manual endodontic files were inserted inside the tooth via the root canal and used to

remove approximately 50-100mg of powdered dentine. The powder was transferred to a sterilised 1.5ml Eppendorf tube and stored in the freezer at -20°C.

2.3.4 DNA extraction

All extractions were carried out in HEPA filtered UV cabinets in the dedicated extraction rooms either at LJMU or MIB. Two different extraction protocols were used: a modified version of that proposed by Rohland and Hofreiter (2007) (hereafter referred as RH protocol), and a modification of the protocol used by Dabney et al (2013) (hereafter referred as Dabney protocol). The latter was optimised for the retrieval of short (<40bp) length DNA fragments, and was designed to maximise DNA recovery in whole genome studies. Both methods are silica-based, in which the DNA is bound to silica in the presence of a chaotropic salt. Both protocols consist of an initial digestion step followed by binding of the DNA to the silica, washing of the silica to remove impurities, and elution of the DNA. They differ in the reagents used, and in the silica-binding method employed: RH used a silica suspension while Dabney used silica columns.

Extractions were carried out in batches of seven samples plus one blank in order to minimise and monitor contamination events. Aliquots of reagents were prepared prior to beginning extraction: stock reagents were filtered using 0.22µm Millex-GS syringe filter units into sterilised 50ml Falcon tubes and, with the exception of ethanol, UV irradiated in the cross linker for 30 minutes. Plastic ware and consumables were also UV irradiated for at least 30 minutes before use. All buffers were prepared prior to beginning extractions.

Rohland and Hofreiter protocol

4.8g of silicon dioxide was added to 40ml of sterile water and allowed to settle for one hour. 39ml of the supernatant was transferred to a clean 50ml tube, leaving behind the large silicon particles. This was left to settle for a further four hours before discarding the top 35ml of the supernatant to remove the smallest particles. Hydrochloric acid (48µl of 30%) was added to the remaining 4ml of silica suspension, mixed and 1ml aliquots made.

Extraction buffer, containing 0.375ml of sterile water, 4.5ml of 0.5M EDTA, and 125µl of 10mg/ml proteinase K, was added to 250mg of powdered sample in 15ml tubes, and left to digest overnight at 37°C in a hybridizing oven.

The following day, samples were removed from the hybridisation oven and centrifuged for 2 minutes at 5000 x g in a bench top centrifuge. Supernatant and pellet colour were noted. The supernatant was transferred to a new, sterilised tube. 2.5ml of binding buffer, containing 2.08ml of 6M guanidinium thiocyanate, 0.25ml of 3M sodium acetate (pH 5.2), and 0.17ml of sterile water, was added to each sample, along with 100µl of silica suspension. Tubes were sealed with Parafilm and rotated in a rotary mixer at room temperature in darkness, for three hours. The supernatant was transferred to a new tube and stored in the fridge. A further 1ml of binding buffer was added to the tube and the silica pellet re-suspended. The suspension was then transferred to a 2ml tube, centrifuged for 15 seconds at 16000 x g in a microcentrifuge, and the supernatant discarded. Two washing steps were performed to remove impurities and compounds which may inhibit the PCR by adding 1ml of washing buffer containing 0.5ml of absolute ethanol, 25µl of 5M sodium chloride, 6.65µl of 1.5M Tris, 2µl of 0.5M EDTA and 466.35µl of sterile water. The sample was then mixed using a vortex, centrifuged for 15 seconds at 16000x g and the supernatant discarded. Following this, the silica was left to dry in the PCR cabinet for 15 minutes at room temperature.

The final step was to elute the DNA. This was done by adding 50µl of TE buffer, containing 5µl of 100mM Tris, 1µl of 50mM EDTA, and 44µl of sterile water, to the silica, mixing and incubating at room temperature for 10 minutes. The suspension was then centrifuged for two minutes at 16000 x g and the DNA-containing supernatant transferred to a 1.5ml tube. This step was repeated to obtain two 50µl extracts, labelled as “elution 1” and “elution 2”.

Dabney protocol

Only 50mg of sample was required for this protocol. Powder was weighed into 2ml tubes, and 1ml of extraction buffer containing 0.9ml of 0.5M EDTA, 25µl of 10mg/ml proteinase K and 875µl of sterile water was added to each sample. Tubes were then left overnight in a shaking heat block at 700 RPM and 37°C. The following day, tubes were placed in a microcentrifuge and spun for 5 minutes at 3000 x g. The supernatant was transferred to a 50ml Falcon tube containing 10ml of Qiagen PB buffer.

An assembly of Qiagen MinElute spin column attached to a Cambridge Bioscience Zymo spin column extension reservoir (with Zymo spin column removed), was placed into a new 50ml Falcon tube, and the supernatant-PB buffer mix transferred into the funnel. The spin column apparatus was then spun at 1500 x g in a table top centrifuge for five minutes or more until the liquid had completely passed through the column. The MinElute column was then placed in a clean collection tube and the Zymo extension reservoir removed. Columns were dry spun for 1 minute at maximum speed (16000 x g). Two wash steps were performed by adding 750µl of Qiagen PE wash buffer to the MinElute column and centrifuging for 30 seconds at 3300 x g, discarding the flow-through each time. The column was again dry-spun for 1 minute at maximum speed and transferred to a clean 1.5ml tube. 30µl of Qiagen EB elution buffer was added to the silica membrane and incubated for 5 minutes. The eluted DNA was collected by centrifuging at maximum speed for 45 seconds. This step was repeated so a total of 60µl of DNA extract was collected in the tube. Extracts were stored at -20°C until required. Bone pellets and binding supernatant were also kept at -20°C, as it is possible for further extraction or binding steps to be performed if required.

2.4 Amplification, electrophoresis, purification and sequencing

PCRs were carried out in a dedicated room; the room was UV irradiated for at least two hours prior to use, and all work surfaces cleaned with 70% sodium hypochlorite bleach before and after use. All plastics were UV irradiated for at least 2 hours prior to use. The outer surfaces of sample and reagent tubes were also cleaned with bleach prior to use. All PCR reactions were performed using a commercially available Qiagen Multiplex PCR Kit.

Table 5. Primer pairs used to amplify mitochondrial hypervariable region I target region for haplotype assignment. Primers taken from Fernandez 2005

Primer Name	Sequence (5'-3')	Start position	End position	Fragment amplified	Annealing temp (°C)
L16125	GCC AGC CAC CAT GAA TAT TG	16106	16125	16095-16280	55
H16262	TGG TAT CCT AGT GGG TGA G	16280	16262		
L16251	CAC ACA TCA ACT GCA ACT CC	16232	16251	16232-16399	55
H16380	TCA ACT GCA ACT CCA AAG CC	16399	16380		

Overlapping primer pairs (Table 5) were used to target a 294bp (positions 16106-16399) region of the mitochondrial hypervariable region I (Fernandez, 2005). Aliquots of PCR reagents were prepared to reduce and monitor contamination of reagents, and reactions were set up in rounds of seven samples, one extraction blank plus two PCR blanks to detect reagent contamination and cross contamination. PCR reactions were performed by mixing 7µl of sterile water, 12.5µl of Qiagen Multiplex PCR Master Mix, and 0.25µl of both forward and reverse primer at 20µM concentration (final concentration 0.2µl), per sample. A master solution containing all reagents was mixed and 20µl was transferred to pre-sterilised 0.5ml tubes. 5µl of DNA extract was then added to individual tubes. For the PCR blanks, 5µl of sterile water was added instead of DNA extract. Samples were then placed in a Multigene Mini (Labnet) thermal cycler for the reaction to run. Cycling conditions can be found in Table 6 and annealing temperatures can be found in Tables 5 and 7.

Table 6. PCR cycle conditions for mitochondrial hypervariable region I and coding region SNP amplification

Number of cycles	Step	Temp (°C)	Time
1	Activation	95	15 minutes
40	Denaturation	95	30 seconds
	Annealing	Dependent on primers	90 seconds
	Extension	72	60 seconds
1	Final extension	72	10 minute

Amplification results were visualised using gel electrophoresis; 5µl of PCR product, along with 2µl of loading dye, was loaded into wells in 2% agarose gel containing GelRed and 100V applied for 40-60 minutes. The results were then visualised using Bio-Rad Image Lab™ software by placing the gel under a Bio-Rad Chemidoc Imager.

Samples which had successfully amplified and with a clean set of blanks were then purified using Nucleospin® Gel and PCR Clean-up Kit following the manufacturer's protocol and sent for sequencing at either Macrogen Europe or GATC Biotech. In order to maximise the final concentration of DNA in the sequencing reactions, samples were prepared using 9µl of DNA extract and 1µl of 25µM primer. Samples were initially

sequenced using the forward primer (Table 6), however if sequencing needed repeating for clarification, the reverse primer was used.

2.5 Sequence Analysis

Sequences were read by aligning them to the revised Cambridge Reference Sequence (rCRS) (Anderson et al., 1981; Andrews et al., 1999) using Mutation Surveyor software (Softgenetics, LLC). Mutations from the rCRS were named according to IUPAC nomenclature rules (Cornish-Bowden, 1985).

2.6 Bacterial cloning

Samples that produced reproducible sequences from two independent extractions were then cloned in order to determine the consensus haplotype. Since only a single DNA fragment can be incorporated into a vector, cloning enabled separation of any contaminant DNA from the endogenous sequence and allowed identification of sequence errors resulting from degradation of the template molecule. This work was done in the modern DNA laboratory from the Archaeology Department at Durham University and in the microbiology laboratory at LJMU.

Cloning was carried out using a commercially available TOPO-TA® Cloning Kit with pCR™2.1 Vector and One Shot® TOP10 Chemically Competent *E.coli*, following the protocol provided by the manufacturer. A ligation step was performed by mixing 4µl of PCR product, 1µl of salt solution, and 1µl TOPO® vector, and incubating the mixture at room temperature for 30 minutes. 2µl of the ligation product was added to a vial of One Shot© TOP10 competent cells and incubated for a further 30 minutes. The cells were then heat shocked at 42°C in a thermal block for 30 seconds and immediately transferred to ice. 250µl of room temperature SOC medium was then added to each vial and placed in a shaking incubator at 37°C for one hour. 50µl of each transformation was then spread on pre-warmed LB agar plates containing 0.1mg/ml of carbenicillin. Plates were then incubated overnight at 37°C.

Up to 10 colonies from each plate were cultured for sequencing. When fewer than 10 colonies were available, as many as possible were selected. The colonies were picked

from the plate using a sterile automatic pipette tip and placed in small plastic tubes containing 2ml in liquid Luria-Bertani broth with 0.1mg/ml of carbenicillin. The tubes were then placed in a shaking incubator overnight at 37°C. Cells were harvested by centrifuging at full speed in a microcentrifuge (Eppendorf) for five minutes and removing the supernatant. Plasmid DNA was purified using either a Qiagen Miniprep Kit or a Nucleospin Plasmid Miniprep Kit following manufacturer's instructions. Sequencing reactions were prepared by mixing 3µl of purified plasmid, 2µl of sterile water, and 5µl of 5µM primer, and sent to Macrogen Europe for sequencing. Cloned DNA was sequenced with universal primer SP6 or T7. Sequences were aligned to the rCRS and read using Mutation Surveyor software (Softgenetics, LLC) as described above.

2.7 Consensus HVRI haplotypes

Strict criteria were followed to establish consensus haplotypes. Haplotypes matching those of genotyped researchers were disregarded. Only haplotypes that could be replicated in independent extractions and amplifications, ideally from separate samples were considered authentic endogenous sequences. The complete haplotype should make phylogenetic sense. Miscoding lesions identified by non-replication in the cloned sequences were discounted.

2.8 Haplogroup diagnostic SNP analysis

After confirmation of HVRI haplotypes through cloning, potential mitochondrial haplogroups were determined using Haplogrep (Kloss-Brandstätter et al, 2011). Haplogroup confirmation was then attempted by targeting mitochondrial coding region single nucleotide polymorphism (SNPs) diagnostic of the haplogroup in question. Diagnostic SNPs were determined using Phylotree Build 17 (van Oven and Kayser, 2009), a phylogenetic tree of human mitochondrial DNA haplogroups. PCR reactions were set up following the same procedure used for the HVRI PCRs (section 2.4) but using primers targeting the specific coding region SNP (Table 7). Data was then added to the HVRI haplotype and once again run through Haplogrep.

Table 7. PCR Primers used for diagnostic coding region SNP analysis. SNPs highlighted in **bold** are the required motifs for the haplogroup for the haplogroup in question

Primer name	Primer sequence	Start	End	Fragment length	SNP	Haplogroup	Annealing Temp °C	Reference
L6999	CAAACTCATCACTAGACATCG	6979	6999	108bp	7028 **T**/C	H	55	Fernández et al. 2014
H7066	GAATGAAGCCTCCTATGATGG	7086	7066					
L12227	GAAAGCTCACAAGAACTGC	12209	12227	152bp	12308A/**G**	U	50	Fernández et al. 2014
H12341	GGTTATAGTAGTGTGCATGG	12360	12341					
L14732	AAAACCATCGTTGTATTTCAA	14712	14732	99bp	14766 **C**/T	HV	55	Fernández et al. 2014
H14792	GGAGGTCGATGAATGAGTG	14792	14810					
L10844	AATTTGAATCAACACAACCA	10825	10844	96bp	10873 T/**C**	L3	55	Fernández et al. 2014
H10901	GGGGAACAGCTAAATAGGTT	10901	10920					
L10380	AGTCTGGCCTATGAGTGACTAC	10359	10380	86bp	10398 A/**G**, 10400 C/**T**	M	55	Fernández et al. 2014
H10423	AATGAGTCGAAATCATTCGTTT	10423	10444					
L10550F	GCATTTACCATCTCACTTCTAGG	10500	10522	129bp	10550 A/**G**	K	55	Gamba et al. 2012
H10550R	GGAGTGGGTGTTGAGGGTTA	10609	10628					
3197F	CCTCCCTGTACGAAAGGACA	3116	3135	132bp	3197 T/**C**	U5	55	Gamba et al. 2012
3197R	GGGCTCTGCCATCTTAACAA	3228	3247					
10238F	CCGCGTCCCTTTCTCCATAA	10199	10218	99bp	10238 T/**C**	N1	54	Fernández 2016
10238R	TGTAGGGCTCATGGTAGGGG	10278	10297					
L10014	TTTTAGTATAAATAGTACCG	9995	10014	112bp	10034 T/**C**	I	50	Sampietro et al 2007
H10088	GTAGTAAGGCTAGGAGGGTG	10088	10107					
12705F	ACTCAGACCCAAACATTAATCAGT	12662	12685	112bp	12705 **C/**T	R	50	Fernández 2016
12705R	CCCTCTCAGCCGATGAACAG	12758	12777					

2.9 Population Analysis

2.9.1 Haplotype search

Complete and partial HVRI haplotypes were searched for in the mitochondrial database EMPOP (Parson and Dür, 2007). EMPOP is an accessible forensic database containing 34617 mtDNA haplotypes from all over the world. The sequences stored in the database conform to rCRS nomenclature. The database allows the specification of sequence range, allowing partial haplotypes to be analysed. The database calculates haplotype frequency worldwide and by metapopulation, and gives a breakdown of geographical distribution of haplotypes within the modern population.

Complete haplotypes were also compared with those in a database of 846 published ancient mitochondrial DNA sequences (Reynolds, Bertoncini and Fernandez-Dominguez, in prep.) to determine geographical and temporal distribution in ancient populations.

2.9.2 Genetic distances

In addition to the ancient DNA database, Reynolds, Bertoncini and Fernandez-Dominguez (in prep.) also created a database of modern mitochondrial DNA, comprising 21405 sequences from 139 Eurasian and African populations. Modern populations were constructed around geographical location and ethnicity (Table 9), while ancient sequences were grouped into populations based on period and geographical location (Table 8). Consensus haplotypes from La Sagrera and Barrão were converted to Fasta sequence files using HaploSearch (Fregel and Delgado, 2011) and added to these databases for analysis. Due to the small number of sequences in individual ancient populations, when possible they were further grouped into metapopulations for the purposes of analysis (Table 8).

Owing to the small number of consensus haplotypes obtained in this study, it was necessary to group these sequences with those from other populations with similar geographical and chronological origin in order to perform population genetics analyses. Data from Algar do Barrão were grouped with data from Algar do Bom Santo (Carvalho et al, 2016), another cave site in the Alenquer region of Portugal with samples also dating to the Middle Neolithic. This was done to create a Middle Neolithic Portugal population

group (n=10). Since there were no other similar populations in the database, these sequences could not be further grouped into a metapopulation. Data from La Sagrera were grouped with sequences from Alto de Reinoso megalithic tomb in Burgos (Alt et al, 2016), the El Mirador Neolithic cave site in Burgos (Gómez-Sánchez et al, 2014), and from El Portalón cave in Atapuerca (Günther et al, 2015) to form the population ‘‘Late Neolithic Spain’’ (n=49) (Table 8) .

Pairwise genetic F_{ST} values were calculated using Arlequin, version 3.5 (Excoffier, Laval and Schneider, 2005), using 1000 permutations. This was done for the current study populations against other ancient populations (Table 8) and against modern populations (Table 9).

2.9.3 Haplogroup Frequencies

The ancient database of Reynolds, Bertocini and Fernandez-Dominguez (in prep.) provided a comparative framework in which the haplogroup composition of the studied populations could be put into archaeological context. Thus, haplogroup frequencies were calculated for the ancient metapopulations seen in table 8 in order to compare the current study populations with populations in the region at different archaeological periods.

2.10 Kinship Analysis

Likelihood ratios were calculated for individuals with identical mitochondrial DNA haplotypes, as they potentially shared a matrilineal relationship. LR (Butler, 2010) is a statistic commonly used in forensic genetics that represents the quotient of the probabilities of an event occurring under two different hypotheses (H0 and H1). In this context, LR is the quotient of the probabilities that the occurrence of identical mitochondrial haplotypes in two individuals results from a matrilineal relationship between those individuals (H0), and from random chance (H1). Likelihood ratio is calculated using the formula:

$$LR = \frac{H0}{H1}$$

where H0 = 1, and H1 = frequency of haplotype in the population.

Since the available data for ancient populations is limited, haplotype frequency was calculated using the modern EMPOP database as a proxy.

LR values were converted to probability values using the formula X/(X+Y), where X and Y represent the probabilities of the hypothesis of maternal relationship (H0) and no maternal relationship (H1), respectively. The minimum allele frequency rule was applied to conservatively estimate the frequency of rare alleles. For haploid markers, the minimum allele frequency is calculated as 5/N, where N = number of individuals in the population. In the event of kinship being established, it is not possible to define the type of relationship, as any relationship through the maternal line, such as siblingship or grand-maternity, may apply.

Table 8. Ancient populations in the database prepared by Reynolds, Bertocini and Fernandez-Dominguez (in prep.), showing metapopulation groupings used for statistical analysis

Metapopulation	Number	Population	Number	Reference
Middle Neolithic Portugal	10	Middle Neolithic Portugal	10	Carvalho et al. 2016; present study
Pre-Neolithic Spain & France	13	Palaeolithic Spain and France	10	Hervella et al. 2012; Fu et al. 2013; Posth et al. 2016
		Mesolithic Spain	3	Sanchez-Quinto et al. 2012 ;Hervella et al. 2012
Early Neolithic Spain	65	Cardial-Epicardial Catalonia	21	Gamba et al. 2012; Lacan et al. 2011b;Olalde et al. 2015
		Cardial-Epicardial Aragon	8	Gamba et al. 2012; Haak et al. 2015
		Epicardial Basque Country	36	Hervella et al. 2012
Middle Neolithic Spain	23	Middle Neolithic Catalonia	11	Sampietro et al. 2007
		Middle Neolithic Basque Country	7	Hervella et al. 2012
		Middle Neolithic Castile and León	5	Haak et al. 2015
Late Neolithic Spain	49	Late Neolithic Spain	49	Alt et al. 2016; Gomez Sanchez et al. 2014; Gunther et al. 2015, present study

Table 9. Modern Iberian populations

Metapopulation	Ethnic group	Location	Number	Reference
North Portugal	Portuguese	Braga Bragança Porto Viana do Castelo Vila Real	186	Pereira et al 2004
Central Portugal	Portuguese	Aveiro Coimbra Castelo Branco Guarda Leiria Lisboa Santarém Viseu	236	Pereira et al 2004
Portugal Unlocalised	Portuguese	Portugal	54	Corte-Real et al 1996
South Portugal	Portuguese	Beja Évora Faro Portalegre Setúbal	123	Pereira et al 2004
Spain Andalusia	Spanish	Andalusia	108	Casas et al 2006
Spain Asturias	Spanish	Asturias	76	Garcia et al 2011
Spain Balearic	Spanish	Majorca	67	Falchi et al 2006
Spain Basque	Basques	Basque country	312	Bertranpetit et al 1995 ; Corte-Real et al 1996; Garcia et al 2011
Spain Basque Navarre	Basques	Elizondo Bera Lekaroz Leiza	110	Cardoso et al 2011
Spain Cantabria	Spanish	Santander (Cantabria) Valle de Pas (Cantabria) Liébana (Cantabria) Pasiegos (Cantabria)	196	Alvarez-Iglesias et al 2009 ; Cardoso et al 2010
Spain Castilla Leon	Spanish	Zamora	214	Alvarez et al 2010
Spain Catalonia	Spanish	Catalonia	116	Alvarez-Iglesias et al 2009; Corte-Real et al 1996
Spain Galicia	Spanish	Galicia	374	Alvarez-Iglesias et al 2009; Salas et al 1998
Spain Unlocalised	Spanish	Spain	74	Corte-Real et al 1996; Pinto et al 1996

Chapter Three: Results

3.1 Extraction and amplification

Thirteen extraction rounds of up to seven samples per round were performed on 53 samples, from 28 individuals (Table 10). At least one PCR for each fragment was performed per extract. When amplification failed for a particular extract, PCR was repeated up to three more times before it was discarded. Figure 10 shows an example of an agarose gel containing positive amplifications. Results were disregarded when amplification of the negative control was observed. When amplification was observed only for the PCR blanks, the same DNA extract was used for subsequent amplifications.

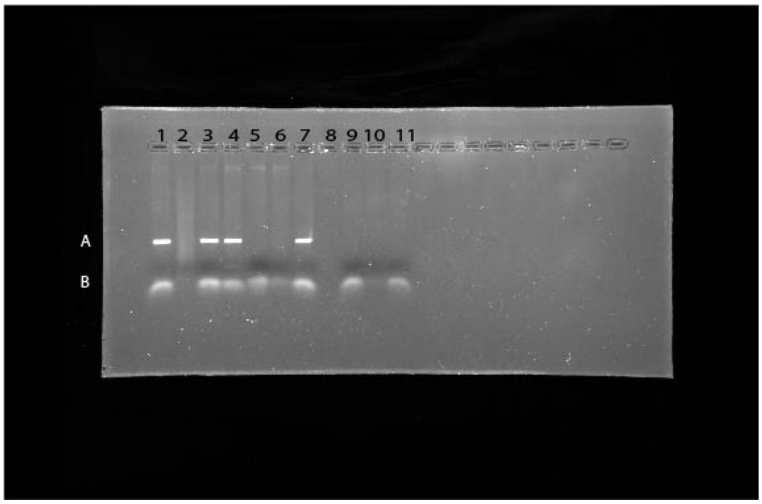


Figure 10. Agarose gel electrophoresis stained with GelRed and visualised under ultraviolet light. Wells 1-7 contain PCR product, well 9 contains the extraction blank, and wells 10 and 11 contain PCR blanks. Bands in row A indicated a positive amplification result, while bands in row B probably result from primer dimers (Robinson 2016)

Only one of the 13 extraction rounds performed was contaminated at the extraction stage (Table 10). Figure 11 show the gel electrophoresis of the amplicons from the extraction round performed on 18/11/14. A band was seen in lane 9, which corresponded to the extraction blank. This indicated that a contamination might have occurred during the extraction stage, rendering the amplification results unreliable. These extracts were discarded and an alternative extraction was performed on the samples affected.

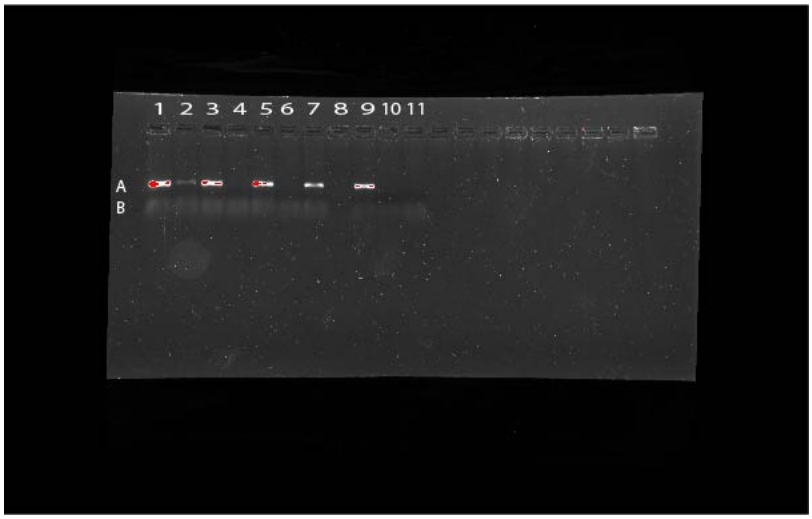


Figure 11. Agarose gel results of PCR performed using primers L16125 and H16262 on the samples extracted on 18/11/2014. Note the presence of a band in column 9 where the extraction blank was loaded (Robinson, 2016)

Table 10. Number of successful extractions performed per sample, total number of PCRs performed per primer pair (L16125, H16262 and L16251, H16380) and number of positive results, and purificate numbers for positive PCRs.

		L16125-H16262			L16251-H16380		
Sample	**Extraction**	**PCR**	**Positive**	**Purificates**	**PCR**	**Positive**	**Purificates**
1SA4	4	5	3	L24, L83, L116	5	3	L34, L89, L127
2SA4	1	1	1	L63	1	0	
1SA8	4	8	2	L30, L117	8	3	L40. L46, L128
2SA8	1	1	1	L64	1	1	L73
1SA18	4	5	5	L25, L84, L94	5	4	L35. L90, L104
2SA18	1	1	1	L65	2	0	
1SA25	1	1	1	L26	2	1	L36
2SA25	1	1	1	L66	2	0	
1SA26	4	6	5	L27, L85, L119	5	4	L37, L91, L130
2SA26	1	1	1	L67	1	1	L74
1SA39	3	8	0		4	1	L41
2SA39	1	2	0		1	1	L75
1SA43	3	3	3	L68, L95, L120	3	3	L76, L111, L131
1SA44	4	7	2	L31, L121	6	2	L42, L132
2SA44	1	1	1	L69	1	1	L77
1SA45	4	6	5	L28, L86, L122	7	3	L38, L48, L133
2SA45	1	1	1	L79	1	1	L78
1SA46	4	5	5	L29, L87, L100	7	5	L39, L92, L105
2SA46	1	1	1	L71	1	1	L79
1SA53	4	6	4	L32, L88, L101	5	3	L96, L106, L135
2SA53	1	1	0		1	1	L80
1SA56	4	7	3	L88, L102, L125	6	2	L107, L136
2SA56	1	1	1	L72	1	1	L81
1SA68	4	9	2	L103, L126	6	4	L43, L108, L137
2SA68	1	2	0		1	1	L82
1CBL1	1	2	1	L4	2	1	L1
2CBL1	1	2	0		2	0	
1CBL2	1	2	1	L5	4	0	
2CBL2	1	3	0		1	1	L44
1CBL3	1	2	1	L6	2	0	
2CBL3	1	2	1	L55	2	0	
1CBL4	1	2	1	L7	4	0	
2CBL4	1	2	1	L33	1	1	L45
1CBL5	1	1	1	L8	1	1	L2
1CBL6	1	1	1	L9	1	1	L3
2CBL6	1	2	0		2	0	
1CBL7	1	2	1	L10	3	0	
1CBL8	1	1	1	L54	1	1	L59
2CBL8	1	1	1	L140	1	1	L149
1BAR3	1	1	1	L56	1	1	L60
2BAR3	2	3	1	L97, L142	2	2	L113
1BAR4	1	1	1	L57	1	1	L61
2BAR4	1	1	1	L98	1	1	L114
1BAR5	1	1	1	L58	1	1	L62
2BAR5	1	1	1	L99	1	1	L115
1CMR1	1	2	0		1	1	L109
2CMR1	1	1	1	L147	1	1	L156
1CMR2	1	2	0		1	1	L110
2CMR2	1	1	0		1	1	L157
1CMR3	1	1	1	L145	1	1	L148
2CMR3	1	1	1	L141	1	1	L151
1CAB3	1	1	1	L138	1	1	L154
2CAB3	2	2	2	L96	2	1	L112

3.2 HVRI amplification and Sanger sequencing

Positive, uncontaminated amplification results were achieved for 43 samples using primers L16125-H16262, and for 43 samples using primers L16251-H16380, indicating successful DNA extraction from all 28 individuals (Table 10).

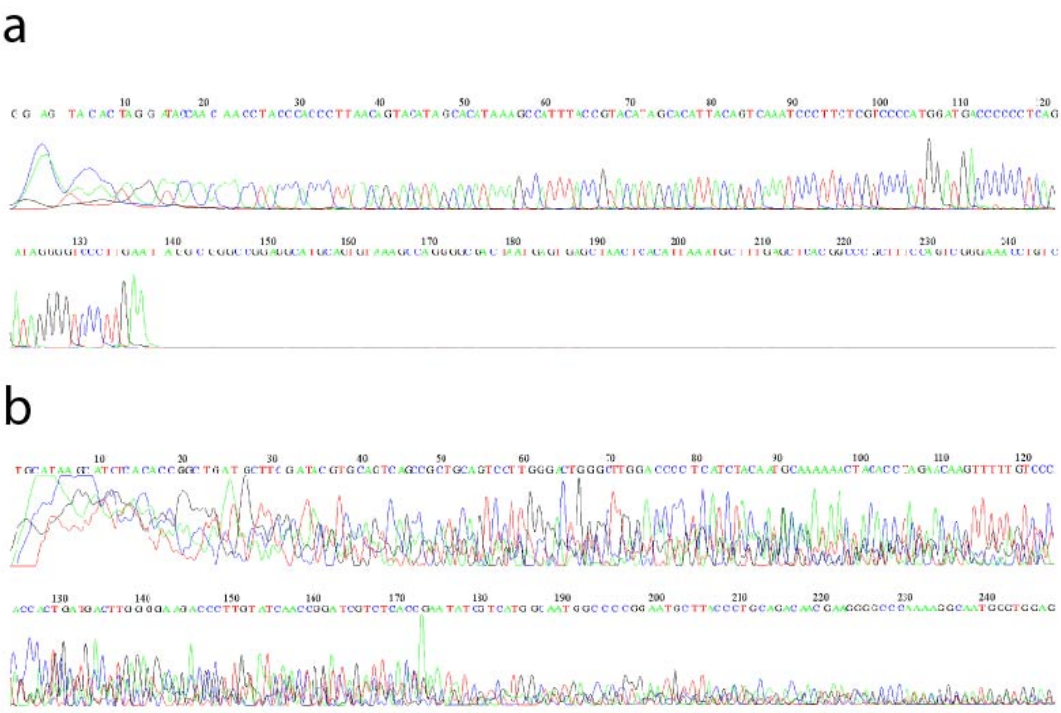


Figure 12. Electropherograms received for purified PCR products sent to Macrogen Europe for sequencing. Purificate L61 from sample 1BAR4 returned a good, clean sequence (a), while purificate L45 from sample 2CBL4 returned a poor quality sequence (b), which could not be aligned to the rCRS and was ultimately unreadable.

All samples with a positive amplification result were purified and sent for sequencing but it was not possible to obtain readable sequences for all purificates. The quality of the sequence data was often too poor to align accurately with the rCRS, and in some cases, despite successful alignment, the sequences contained too many ambiguous base calls (Figure 12). It was possible to obtain complete or partial sequences for 47 samples, representing all 28 individuals (Table 11). Most of the sequences obtained from Castelo Belinho samples showed high levels of noise and had highly saturated regions within the sequence. Often in these samples, multiple peaks were observed sequentially, e.g. in sample 1CBL3, extracted on the 28/1/15, multiple peaks at base positions 16213-16222 were seen, and similarly in samples 1CBL1, 1CBL4, 1CBL6 and 1CBL7. These types of

mutations were generally disregarded as artefacts resulting from either low concentration, poor quality DNA, or a problem with sequencing. In any case, they were not replicated in subsequent extractions. In general, samples from all other sites yielded much cleaner sequences, with lower levels of noise and saturation.

Sequences obtained from purificates L7, L9, L28, L29, L44, L53, L73, L76, L81, L103, L123, L130, L132, and L151 all contained one or more mutations which were in common with researcher profiles (Tables 4 and 11). These mutations were regarded as exogenous contamination and as such were considered not to be a true endogenous mutation.

3.3 Mitochondrial HVRI Haplotype determination

Although sequences were obtained from all 28 individuals, it was only possible to replicate results for one or both HVRI fragments for 13 of them (Table 11). It was not possible to replicate any samples from Castelo Belinho, Anta da Cabeceira, or Casais da Mureta. From Algar do Barrão, a complete HVRI sequence was reproduced for individual BAR4, while partial sequences were reproduced for both BAR3 and BAR5 (Table 11). From La Sagrera, full, validated sequences were obtained for individuals SA4, SA18, SA25, SA26 and SA53. Partial sequences were replicated for individuals SA39, SA43, SA44 and SA45. Initially, it appeared that a partial sequence had been obtained for individual SA46 but was subsequently disregarded as potential exogenous contamination because it was found to match a researcher profile (Table 4). However, this sample has since been re-extracted independently at Durham University, and the same profiles were obtained (Di Renno, 2016). Given that the identified researcher was not associated with this laboratory, it is likely that the sequence represented the individual's endogenous haplotype. For this reason, this sample has been included in haplotype analysis despite not being subject to further SNP analysis.

Table 11. Sequencing results of samples for which a full or partial consensus haplotype could be determined. Actual read positions and mutations from rCRS shown. Mutations in common with staff profiles are listed in red, as potential contamination. Mutations that could be repeated between different extractions of the same individual are listed in green. Mutations in black could not be repeated and were not considered part of the true endogenous haplotype. Purificates highlighted in **bold** were selected for cloning. Also included are details of sample 1SA25 which was not replicated but was cloned.

			Primers	
Sample	**Purificate**	**Read positions**	**L16125, H16262**	**L16251, H16380**
1SA4	**L24**	16152-16281	16223T 16261Y	
	L83	16151-16280	16223T 16261T	
	L34	16285-16399		16320T 16380S
	L127	16280-16399		16320T
1SA18	**L94**	16165-16280	16223Y	
	L35	16272-16399		16311C
	L90	16285-16399		16311Y 16354Y
	L104	16290-16399		16311Y 16315K 16354Y 16362Y
2SA18	**L65**	16125-16280	16223T 16179Y	
1SA25	L26	16157-16281	16223T 16261Y	
	L36	16278-16399		16320T
1SA26	**L85**	16159-16280	CRS	
	L91	16294-16399		16311Y 16355R 16362Y
	L130	16231-16399		16257M 16261Y 16294Y 16304Y 16311T
2SA26	**L67**	16167-16280	CRS	
	L74	16285-16399		16311C
1SA43	**L68**	16171-16280	16279M 16223Y	
	L95	16195-16280	16223T 16241R	
	L120	16105-16280	16223Y	
1SA45	**L53**	16167-16281	16221T 16223Y	
	L122	16105-16280	16158G 16223T	
1SA46	L39	16234-16399		CRS
	L49	16286-16399		CRS
	L92	16282-16399		CRS
1SA53	**L32**	16105-16281	16163del	
	L124	16105-16280	16204Y 16223Y	
	L106	16303-16399		16320Y 16325Y 16355Y 16362Y
	L135	16231-16399		16362C

3.4 Cloning

Replicated samples were cloned to authenticate the presence of endogenous DNA. Individual SA25 was also cloned, despite not having been replicated. There was

insufficient sample to perform a second extraction but the profile was determined to be rare and matched that of another individual from the same site. It was assumed authentic due to the quality of the initial sequences and lack of any match to researcher profiles or any other samples in the same extraction round. Though this was not standard procedure, the decision was taken to continue the analysis of this sample.

Cloning was undertaken at Durham University on purificates from La Sagrera. Of the 20 samples plated, eight of the plates failed to grow colonies, suggesting that the cloning process had been inefficient. Subsequent cloning was performed at LJMU, repeating the failed samples in addition to the remaining samples from La Sagrera, when 13 out of 15 samples were successfully cloned. A final round of cloning was performed at LJMU on samples from Algar do Barrão and samples which had so far failed to produce an adequate number of clones. In this round, all 22 samples were successfully cloned. Altogether, it was possible to obtain at least four clone sequences for all targeted samples (Appendix 2 and 3).

3.5 Consensus haplotype determination

After cloning, full mitochondrial HVRI haplotypes were confirmed for four individuals from La Sagrera and one from Barrão), while partial haplotypes were assigned to a further four from La Sagrera and two from Barrão (Table 12).

Table 12. Mitochondrial hypervariable region I consensus haplotypes for replicated samples as determined relative to rCRS, and haplogroup assignments including quality score made using Haplogrep. *only covers positions 16095-16280, #only covers positions 16232-16369

Individual	Consensus HVRI haplotype	Haplogroup assignment	Haplogrep quality score (%)
SA4	16129A 16223T 16261T 16320T	L3e2	81.3
		G1b	75.5
		I	75.5
SA18	16179T 16223T 16311C	M40a	81.29
		N5	79.9
	16223T 16311C	N5	100
	16179T 16223T 16354T	M73b	74.58
	16223T 16354T	M73b	89.6
SA25	16129A 16223T 16261T 16320T	L3e2	81.3
		G1b	75.5
		I	75.5
SA26	16311C	H	100
SA39	16320T#	H	100
		L3	100
SA43	16223T*	N	100
		I	
SA44	CRS*	H, HV or U	
SA45	16223T*	N	100
		I	
SA46	16221T	H10e	100
SA53	16223T 16362C	L3d5	100
		M3, M9, E, G, D…	100
BAR3	16270T*	H2	50
BAR4	16129A 16224C 16311C	K1a11	100
	16129A 16224C 16311C 16379T	K1a11	100
BAR5	16270T 16311C#	U5b1c	100

Individual SA18 was more difficult to determine. In the first fragment, mutation 16179T appeared in eight out of nine clone sequences for purificate L65; however this mutation did not appear in any of the 10 clones of the replicate extraction, purificate L94. Similarly, mutation 16223T was present in the PCR sequence of both purificates and was present in all nine clones for purificate L65 but only appeared in one of the ten L95 clones. The

second fragment was similarly ambiguous. Direct sequencing suggested a haplotype of 16311C 16354C (Table 11) but cloning revealed that these mutations belonged to different sequences. This suggests that this sample may have become contaminated with exogenous DNA at some point prior to extraction. The mutations did not coincide with the haplotypes of any of the researchers involved in processing the samples (Table 4). Four possible HVRI haplotypes were proposed for this sample (Table 12) and diagnostic SNPs were proposed to attempt haplogroup confirmation, however the ambiguity surrounding this sample precluded its use in population genetic analyses.

Individuals SA4 and SA25 shared a HVRI haplotype for the whole target region, whist individuals SA43 and SA45 appeared to share a haplotype based on the single fragment that was successfully recovered for them. Individual SA53 also has the 16223T mutation in common with SA43 and SA45 for the first HVRI fragment, though the second fragment was also successfully sequenced in this case (Table 12).

BAR3 and BAR5 share a mutation in the overlapping region of the target fragment, suggesting that they may share a haplotype but since only partial sequences were obtained for these samples, it is impossible to confirm this.

3.6 Haplotype search

3.6.1 Modern Population

Table 13 shows the frequency of obtained haplotypes within the EMPOP mtDNA database (Parson and Dür, 2007), both worldwide and within the West Eurasian metapopulation. When more than 1000 matches were returned, geographical distribution was not listed.

The HVRI haplotype shared by individuals SA4 and SA25 and three of the proposed haplotypes for individual SA18 (16179T 16223T 16311C, 16179T 16223T 16354T, and 16223T 16354T) were absent from the modern population according to the EMPOP database (Table 13).

The fourth proposed haplotype for individual SA18 (16223T 16311C) was poorly represented in the modern population, with a worldwide frequency of only 0.0072 and

being most widely observed in the South Asian population. Its current distribution encompasses Central Asia, East Asia, South Asia, South East Asia, and North America.

The partial haplotype of SA39 was observed at a frequency of 0.00026 worldwide, and 0.0004 West Eurasian but was quite widely distributed, being found in Central Asia, southwestern Europe, southeastern Europe, Central, Southern, and Northern Europe.

Individual SA46 has a worldwide frequency of 0.00173, and had a wide global distribution, encompassing North and South America, Europe, and Asia.

SA26 and SA53 were both observed moderately, with worldwide frequencies of 0.00886 and 0.0067 respectively, although SA53 was frequent within the East Asian metapopulation than any other (frequency = 0.04656), whereas SA26 was similarly represented in both West Eurasian (frequency = 0.01688) and South Asian (frequency = 0.01635) metapopulations. SA26 was far more widely distributed than SA53, with both appearing in North America, Central Europe, southeastern Europe, South East Asia, Central Asia, East Asia, North Africa, South Asia, Middle East, while SA26 is also represented in South America, Russian Federation, Northern Europe and southwestern Europe.

The partial haplotypes of SA43 and SA45, and SA44 haplotypes had worldwide frequencies of 0.07502 and 0.15279, respectively. Both haplotypes had frequencies >0.2 in the West Eurasian metapopulation, and were widely geographically distributed over all continents. Interestingly, these were the only haplotypes from this study that are present in Western Europe.

Algar do Barrão haplotypes were poorly represented within the modern population with worldwide frequencies <0.002. The one complete and two partial haplotypes had the greatest representation within the modern West Eurasian metapopulation and were distributed in Southern Europe and southwestern Europe (BAR3), Central Europe and southeastern Europe (BAR4 and BAR5), Central Asia (BAR3 and BAR4), and Cyprus (BAR4).

Table 13. Frequency and geographical distribution of La Sagrera and Algar do Barrão haplotypes in the EMPOP database of modern populations worldwide and in the West Eurasian metapopulation.

Individual	HVRI Haplotype	Range	Worldwide	Worldwide Frequency	Metapopulation	Metapopulation Frequency
SA4	16129A 16223T 16261T 16320T	16105-16399	0/31810			
SA18	16179T 16223T 16311C	16105-16399	0/31810			
	16223T 16311C	16105-16399	22/30577	0.00072	West Eurasian	0.00036377
	16179T 16223T 16354T	16105-16399	0/31810			
	16223T 16354T	16105-16399	0/31810			
SA25	16129A 16223T 16261T 16320T	16105-16399	0/31810			
SA26	16311C	16105-16399	271/30577	0.00886	West Eurasian	0.01687900
SA39	16320T	16105-16280	8/30577	0.00026	West Eurasian	0.00043652
SA43	16223T	16105-16280	2597/34617	0.07502	West Eurasian	0.03278400
					West Eurasian Admixed	0.02912600
SA44	CRS	16105-16280	5289/34617	0.15279	West Eurasian	0.26777000
					West Eurasian Admixed	0.22330000
SA45	16223T	16105-16280	2597/34617	0.07502	West Eurasian	0.03278400
					West Eurasian Admixed	0.02912600
SA46	16221T	16105-16399	53/30577	0.00173	West Eurasian	0.00276460
SA53	16223T 16362C	16105-16399	205/30577	0.00670	West Eurasian	0.00334670
BAR3	16270T	16105-16280	51/34617	0.00147	West Eurasian	0.00259310
BAR4	16129A 16224C 16311C	16105-16399	43/30577	0.00141	West Eurasian	0.00276460
BAR5	16270T 16311C	16232-16399	49/30577	0.00160	West Eurasian	0.00305570

3.6.1 Ancient populations

Haplotypes of individuals SA4 and SA25, SA18, and BAR4 were absent from the ancient database suggesting that they had not previously been recorded in ancient populations. The haplotype from individual SA26 (16311C) is the most widely represented in the ancient database. It has previously been reported from Central, Northern, southeastern, southwestern, and Southern Europe, and the Middle East, and has been found in populations spanning from the Early Neolithic to the Iron Age. Similarly, the haplotype of SA53 was previously reported in Early Neolithic Barcelona (Table 14). Individual SA46 shares a haplotype with two previously reported individuals: one in Iron Age Denmark, and another from the Middle Neolithic cave site Bom Santo in Portugal (Table 14).

Table 14. Location, period, date, haplogroup and reference of sequences in the ancient mtDNA database that share a complete HVRI haplotype with individuals from La Sagrera and Algar do Barrão.

Individual	HVRI haplotype	Location	Period	Dating	Haplo-group	Reference
SA26	16311C	Saxony-Anhalt (Germany)	Early Neolithic	Archaeological context	HV	Brandt et al. 2013
				5078 - 4998 (1σ) BCE cal	HV	Haak et al. 2005, Brandt et al. 2013
				5078 - 4998 (1σ) BCE cal	HV	Haak et al. 2005, Brandt et al. 2013
				5247 ± 45 BCE cal	HV	Haak et al. 2010, Brandt et al. 2013
			Middle Neolithic	4100-3950 BCE cal	HV	Brandt et al. 2013
				Archaeological context	HV	Brandt et al. 2013
				Archaeological context	HV	Brandt et al. 2013
			Early Bronze Age	Archaeological context	H	Brandt et al. 2013
				Archaeological context	H	Brandt et al. 2013
		Olbia-Tempio (Sardinia)	Late Bronze Age	1200–1300 BCE cal	H	Caramelli et al. 2007
		Aleph (Syria)	Early Neolithic	7500-7300 BCE cal	K	Fernández et al. 2014
				7500-7300 BCE cal	K	Fernández et al. 2014
				7500-7300 BCE cal	U*	Fernández et al. 2014
		Navarre (Spain)	Final Early Neolithic	5207-4728 (2σ) BCE cal	HV	Hervella et al. 2012
				5310-4497 (2σ) BCE cal	H	Hervella et al. 2012
				5310-4497 (2σ) BCE cal	H	Hervella et al. 2012
		North-West Romania	Early Neolithic	6500-5500 BCE cal	HV	Hervella et al. 2015
		Lassithi plateau of east-central Crete (Greece)	Bronze Age	4900-3800 BP	HV	Hughey et al. 2013
				4900-3800 BP	HV	Hughey et al. 2013
				4900-3800 BP	HV	Hughey et al. 2013
		Gotland (Sweden)	Middle Neolithic	4800-4000 BP	Others	Malmström et al. 2009
				4800-4000 BP	Others	Malmström et al. 2009
		Southern Zealand (Denmark)	Iron Age	200–400 AD	H	Melchior et al. 2010
		Hungary	Middle Neolithic	Archaeological context	HV	Szécsényi-Nagy et al. 2015
SA46	16221T	Alenquer (Portugal)	Middle Neolithic	3800–3655 (95% prob)		Carvalho et al. 2016
		Roskilde (Denmark)	Iron Age	1000-1250 AD	H	Melchior et al. 2010
SA53	16223T	Kola Peninsula (Russia)	Early Metal Age	3500 BP uncalibrated	D*	Der Sarkissian et al. 2013
				3500 BP uncalibrated	D*	Der Sarkissian et al. 2013
				3500 BP uncalibrated	D*	Der Sarkissian et al. 2013
		Begues, Catalonia (Spain)	Early Neolithic	5475–5305 BCE cal	N*	Gamba et al. 2012
				5475–5305 BCE cal	N*	Gamba et al. 2012

3.7 Haplogroup Analysis

Initial HVRI haplogroup assignments, along with quality scores can be seen in Table 12. Based on these preliminary haplogroup assignments, coding region SNPs were tested to provide greater resolution to the haplogroup assignments. Diagnostic SNP results and subsequent haplogroup assignments can be seen in Table 15.

3.7.1 La Sagrera

Two thirds of the individuals analysed from the La Sagrera site were assigned to mitochondrial macrohaplogroup N.

Individuals SA4 and SA25 share the same mitochondrial haplotype. They were initially considered members of haplogroup L3 based on the HVRI results, with a quality score of 81.3% (Table 12). However, the absence of the 10873C motif of haplogroup L3 (Table 15) ruled out this possibility for both samples. Instead, these individuals are likely to belong to haplogroup N, with a quality score of around 68%. Both samples possess an A at position 10398, excluding assignment to haplogroup N1a1 (including I), as this cluster is characterised by a G in this position. Further SNP typing is required to determine the exact haplogroup.

All potential HVRI haplotypes for SA18 were initially assigned to haplogroup N or M with quality scores of 75% or higher. Testing for the 10873C motif of L3/M branch rules out haplogroup M by its absence. Subsequent testing for haplogroup R – a subclade of haplogroup N – using base position 12705 precluded this option. Individual SA18 was assigned to haplogroup N.

Although individual SA26 was assigned to haplogroup H based on the HVRI haplotype, it was not possible to confirm this. The expected marker of haplogroup H (7028C) was not present for this individual. Further analysis also ruled out haplogroup K by determining base position 10550A. It was concluded that this individual belonged to haplogroups HV1 or R1 with a Haplogrep score of 100%.

Individual SA39 was preliminarily assigned to either haplogroup H or haplogroup L3. The absence of the 7028C SNP ruled out haplogroup H. Haplogroup L3 marker 10873C was

absent in one PCR but sequencing failed in the subsequent replication. HV marker 14766C was observed once but sequencing failed for the second PCR. Taking into account the SNP that failed to replicate, SA39 was assigned to haplogroup HV, with a quality score of 84%.

Based on the HVRI haplotype, individual SA43 was assigned to haplogroup N with a quality score of 100% (Table 15). Two PCRs performed for the 10873C L3/M SNP produced different results and PCRs performed for the 10034C SNP of haplogroup I failed to sequence. If the 10873C is considered accurate, it is likely that this individual is a member of haplogroup L3. Given that the 10873C mutation failed to appear in any other samples, it is unlikely that it results from a contamination event. Thus, this individual may be a member of haplogroup L3, or alternatively may belong within macrohaplogroup N.

Despite only confirming a partial HVRI haplotype for individual SA44, it was possible to assign this individual to a macro haplogroup. The sequenced fragment matched the rCRS, which belongs to haplogroup H2a2a1, so it was possible that this individual was also a member of haplogroup H. Two PCRs from separate extracts from this individual were sequenced and found to contain the 7028C, a marker of haplogroup H. Thus, it was possible to assign individual SA44 to haplogroup H.

Individual SA45 also assigned to haplogroup N due to the presence of the 10873T mutation, while the 16221T mutation of the HVRI haplotype of individual SA46 means it can be assigned to haplogroup H10.

Individual SA53 can be assigned to macrohaplogroup N due to the presence of mutation 10238T. The same haplotype was observed in two early Neolithic samples from the Barcelona region that were assigned to haplogroup N*, so it is possible that SA53 is also a member.

3.7.2 Algar do Barrão

Both individuals BAR3 and BAR5 were assigned to haplogroup U5 – determined by SNP 3197C – while individual BAR4 was assigned to haplogroup K. Haplogroup assignments for all samples from Algar do Barrão achieved a quality score of at least 88%.

Table 15. Coding region SNP results and haplogroup assignment. SNPs highlighted in bold failed to replicate across PCRs.

Sample	HVRI haplotype	7028C/T	12308G/A	10873T/C	10398A/G	10550G/A	14766C/T	12705C/T	3197C/T	10034C/T	10238T/C	Haplogroup
1SA4	16129A 16223T 16261T			10873T	10398					No result		Within N
1SA18	16179T 16223T 16311C			10873T				12705T				N
	16223T 16311C											
	16179T 16223T 16354T											
	16223T 16354T											
1SA25	16129A 16223T 16261T			10873T	10398A					No result		Within N
1SA26	16311C	7028T				10550A						HV1/ R1
1SA39	16320T	7028T		**10873T**			14766C					HV
1SA43	16223T			10873C						No result		L3
1SA44	CRS	7028C	No results									H
1SA45	16223T			10873T						No result		N
1SA46	16221T											H or HV
1SA53	16223T 16362C			10873T							10238T	N
BAR3	16270T	7028T	12308G	10873T			14766C	12705C	3197C			U5
BAR4	16129A 16224C 16311C					10550G						K
BAR5	16270T 16311C					10550A			3197C			U5

3.8 Population Analysis

3.8.1 Ancient populations

Within the grouped metapopulations, the genetic composition of the 'Late Neolithic Spain' population was quite similar to the other ancient Iberian populations, with F_{ST} values <0.1 in all cases. The greatest, significant, pairwise distance was from 'pre-Neolithic Spain and France', which contained populations ' Palaeolithic Spain and France' and 'Mesolithic Spain', (F_{ST} = 0.09, p< 0.005) (Table 16). 'Early Neolithic Spain' and 'Middle Neolithic Spain' were found to be genetically similar to both 'Middle Neolithic Portugal' and 'Late Neolithic Spain' (F_{ST} = 0.03, 0, 0.02, and 0.01, respectively), and to each other (F_{ST} = 0.02). None of these results were significant, reiterating that these populations were genetically similar. The metapopulation 'Pre-Neolithic Spain and France' was more distant from 'Early Neolithic Spain', 'Middle Neolithic Spain' and 'Late Neolithic Spain' Iberian populations (F_{ST}= 0.16, p<0.001; F_{ST} = 0.04, p> 0.05; F_{ST} = 0.08, p< 0.001, respectively), suggesting a shift in the genetic structure from pre-Neolithic to Neolithic populations in this region.

Table 16. Pairwise F_{ST} distances of the 'Late Neolithic Spain' and 'Middle Neolithic Portugal' populations against other ancient populations from Iberia and France found in the database by Reynolds, Bertocini and Fernandez-Dominguez (in prep.). Populations have been grouped into metapopulations, which can be seen in table 8. Darker colour indicates greater distance between populations. Significant results are indicated by red text

	Middle Neolithic Portugal	Pre-Neolithic Spain and France	Early Neolithic Spain	Middle Neolithic Spain	Late Neolithic Spain
Middle Neolithic Portugal	0				
Pre-Neolithic Spain and France	0.04763	0			
Early Neolithic Spain	0.03023	0.16308	0		
Middle Neolithic Spain	0	0.04217	0.01718	0	
Late Neolithic Spain	0.00257	0.08959	0.01555	0.01053	0

3.8.2 Modern populations

The ‘Late Neolithic Spain’ population, which included the sequences from La Sagrera, was found to differ significantly from all modern Iberian populations, with the smallest distance being with Central Portugal (F_{ST} = 0.0144, p=0.006). Both ‘Late Neolithic Spain’ and ‘Middle Neolithic Portugal’ were most distant from Basques (F_{ST} = 0.07099, p < 0.001; F_{ST} = 0.04609, p = 0.036 respectively), although both Basque populations were also significantly distant from all other modern Iberian populations (Table 17). In fact, modern Spanish populations were also generally significantly distant from each other. Modern Portuguese populations seemed to be more genetically similar to other modern Iberian populations. The Middle Neolithic Portugal population was not significantly distant from any other modern Iberian populations.

	Late Neolithic Spain	Middle Neolithic Portugal	North Portugal	Portugal Unlocalised	South Portugal	Spain Andalusia	Spain Asturia	Spain Balearic	Spain Basques	Spain Basques Navarre	Spain Cantabria	Spain Castilla Leon	Spain Catalonia	Spain Galicia	Spain Unlocalised
Late Neolithic Spain	0														
Middle Neolithic Portugal	0.00290	0													
North Portugal	0.02208	0.00765	0												
Portugal Unlocalised	0.02550	0.02419	0	0											
South Portugal	0.02198	0.01712	0.00839	0.01464	0										
Spain Andalusia	0.02442	0.01714	0.00156	0.00626	0.01012	0									
Spain Asturia	0.03128	0.01192	0.0018	0.00145	0.01762	0.0099	0								
Spain Balearic	0.01809	0.00231	0	0	0.01235	0.00349	0.0054	0							
Spain Basques	0.07099	0.04609	0.02185	0.03019	0.04799	0.0343	0.00891	0.03569	0						
Spain Basques Navarre	0.03471	0.01225	0.01911	0.02014	0.04083	0.03411	0.02033	0.01609	0.03298	0					
Spain Cantabria	0.03121	0.00678	0.01132	0.00728	0.0316	0.01928	0.00686	0.01356	0.02279	0.02624	0				
Spain Castilla Leon	0.02344	0	0.00422	0.01127	0.01104	0.01085	0.00159	0.00996	0.01741	0.01873	0.01365	0			
Spain Catalonia	0.01994	0.00927	0.00044	0.00434	0.00316	0	0.00625	0.00548	0.0258	0.02874	0.01541	0.00459	0		
Spain Galicia	0.02314	0.00625	0.00252	0.00149	0.01707	0.00845	0	0.00656	0.01509	0.0187	0.00659	0.0037	0.00407	0	
Spain Unlocalised	0.01804	0.00152	0.00331	0.00692	0.0054	0.00296	0.00817	0.00393	0.03206	0.01949	0.01933	0.00262	0	0.00666	0
Central Portugal	0.01440	0.0083													

Table 17. Pairwise F_{ST} distances of the ancient populations 'Late Neolithic Spain' and 'Middle Neolithic Portugal' from modern Iberian populations found in the Reynolds, Bertocini and Fernandez-Dominguez database. Darker colour indicates greater distance between populations. Significant results are indicated by red text

3.9 Haplogroup Frequencies

Table 19 shows the macrohaplogroup frequencies for the ancient metapopulation groups used for analysis. More than a third of the haplogroups found in 'Middle Neolithic Portugal', which includes the samples from Algar do Barrão, belonged to U (Table 18), with less than 10% of the population belonging to haplogroups H or HV respectively, which are relatively common haplogroups in modern day Europe. Similarly, only 15% of 'Late Neolithic Spain' samples containing the La Sagrera individuals were found to belong to haplogroup H, with none being assigned to haplogroup HV. As expected, pre-Neolithic Iberia had a high frequency of haplogroup U (62%).

Haplogroups J and K are present in all Neolithic populations but absent in pre-Neolithic Iberia (Table 18). All groups contained haplogroups U and H. Cardial/ Epicardial Spain appeared to have the greatest haplogroup diversity, with both earlier and later populations being more homogenous. The pre-Neolithic population displayed the least haplogroup diversity.

Table 18. Haplogroup frequencies of the ancient metapopulations

Haplogroup	Middle Neolithic Portugal	Pre-Neolithic Spain and France	Early Neolithic Spain	Middle Neolithic Spain	Late Neolithic Spain
H	9%	15%	34%	32%	15%
HV	9%		2%		
I			2%	4%	
J	18%		3%	14%	6%
K	18%		20%	9%	22%
M		8%			
N			8%		12%
T	9%	8%	6%	9%	12%
U	36%	69%	17%	32%	18%
V			2%		3%
W			3%		
X			5%		12%

3.10 Kinship Analysis

3.10.1 Algar do Barrão

Only one of the three skeletons sampled from Algar do Barrão yielded a complete haplotype. BAR3 and BAR4 yielded partial haplotypes each from a different fragment. These individuals shared a mutation in the overlapping regions of the two fragments. However, this was not enough information with which to calculate likelihood ratios. Thus, it was not possible to determine kinship for the individuals at Algar do Barrão.

3.10.2 La Sagrera

Of the six individuals with complete haplotypes, individuals SA4 and SA25, from clusters A and C respectively (Figure 11), shared an identical haplotype which was absent from all of the ancient and modern databases searched suggesting it is rare. 'Minimum allele frequency' (5/N) was used to estimate the frequency with which to calculate likelihood ratios. Minimum allele frequency = 1.44×10^{-4}. Using the global population in the modern EMPOP database (N = 34617), a likelihood ratio of 6,925 was calculated, giving the probability of matrilineal relationship between these two individuals of 99.98%. When calculated for the West Eurasian population (N=16197), likelihood ratio was calculated as 3239, giving a probability of 99.96%. Of the partial haplotypes generated, individuals SA43 and SA45, both from cluster C (Figure 13), shared a haplotype for fragment 16105-16280. This haplotype had a worldwide frequency of 0.07502 (Table 13), and a West Eurasian frequency of 0.03278, giving likelihood ratios of 13.33 and 30.51, with probabilities of 93% and 96.8%, respectively. No other individuals at La Sagrera shared a haplotype, indicating a lack of matrilineal relationships among the individuals sampled.

Figure 13. Positions of individuals with shared haplotypes at La Sagrera. Text box colour represents corresponding skeleton within each cluster. Figure modified from Balaguer, 2016.

Chapter Four: Discussion

4.1 DNA Preservation

The general preservation state of the studied samples was poor and as a result, it was only possible to establish full or partial consensus haplotypes for 13 of the individuals sampled. No consensus haplotypes were obtained from the earliest site sampled for this study, Castelo Belinho. Conversely, La Sagrera was the most recent site and yielded the greatest number of sequences. Even under ideal conditions, DNA degrades over time, thus it could be expected that later samples would be better preserved. Damage to the DNA molecule begins almost immediately post-mortem as DNA repair mechanisms cease and nucleases begin to digest DNA molecules. However, in ideal conditions, retrievable DNA molecules should theoretically persist for several millions of years (Allentoft et al, 2012), meaning the difference in time since deposition between these sites (< 2500 years) is unlikely to have been the main cause of the DNA preservation discrepancy between the two. Furthermore, aDNA studies in general display sporadic recovery across temporal scales, indicating that the relationship between sample age and DNA preservation is not easy to define.

The macroscopic preservation of samples from Castelo Belinho, Casais da Mureta and Anta da Cabeceira was generally poor. Obvious damage to the sample is likely to affect DNA preservation, since this could breach the protective quality of the surface. Fissures and openings in the bone or enamel surface will expose the internal DNA to environmental conditions that accelerate and encourage DNA degradation. This could expose the inner surfaces to greater microbial activity, which can degrade DNA (Leney, 2006). Additionally, presence of large quantities of non-target DNA (e.g. bacterial/ fungal) can inhibit amplification of target DNA. Furthermore, damage to the external surfaces of the samples is likely to have exposed the DNA to the surrounding soil. Presence of organic compounds such as humic acids and fulvic acids commonly found in soil and sediments are known to inhibit PCR reactions (Sutlovic et al, 2008).

The genetic data from Neolithic Iberia is relatively scarce, and this is likely the result of the regions environmental conditions. High temperatures associated with the Iberian climate are likely to be sub-optimal for DNA preservation (García-Garcerà et al, 2011; Hofreiter et al, 2015). Samples from Castelo Belinho and Anta da Carbeceira failed to yield

any reproducable sequences; this may be due to greater temperature fluctuations at open sites compared with cave sites (Gamba et al, 2008; Gómez-Sánchez et al, 2014). Szécsényi-Nagy et al (2017) observed higher amplification success rates in northern and easten Iberia than in southern and western regions of the peninsula. This is in keeping with the pattern seen in the current study.

4.2 Burial patterns and kinship in Middle and Late Neolithic Iberia

Most of the individuals at La Sagrera lacked matrilineal relationships. This was somewhat surprising given the careful, clustered arrangement of the individuals. Such clustered burial practices of the Neolithic are often indicative of kinship-based funerary arrangement (e.g. Haak et al, 2008; Alt et al, 2016). The only potential relationship observed on the maternal line was between two juvenile skeletons buried in separate clusters (clusters A and C, respectively). It is likely that these children were siblings or cousins. The likelihood of kinship between these two individuals is further evidenced by $^{87}Sr/^{86}Sr$ ratios, which were within the same range for both skeletons, indicating that they originated from the same population (Robinson et al, 2016).

With the Neolithic, communities began to move away from the use of individual graves to a practice of collective burial (e.g. Lee et al, 2014; Carvalho et al, 2016). The development of permanent settlements, and the associated increase in population size, resulted in individual interment becoming less practical. Thus, communal graves may have been more convenient and more commonplace (Sorensen, 2013). These collective burials may have also reflected an emerging sense of community or collectiveness felt as social complexities increased. Often individuals grouped closely together or placed facing each other are thought to represent kinship groups. Previous ancient DNA analysis has revealed familial relationships between such individuals (e.g. Haak et al, 2008; Lee et al, 2014). It should be noted, however, that only six complete mitochondrial haplotypes were retrieved from the 13 skeletons sampled, thus the opportunity for kinship analysis was somewhat limited.

There were also several males present in the clusters, though the Y chromosome was not analysed here. Patrilocal societies have previously been suggested for Early Neolithic

Europeans (Lacan et al, 2011; Bentley et al, 2012), and both nuclear families and collective burials characterised by the presence of a paternal lineage have been observed in several Late Neolithic contexts (Haak et al, 2005; Lacan et al, 2011). It is therefore possible that the clusters did in fact represent nuclear families but that the preservation status of the samples along with the single lineage marker analysed means that kinships were not detected. Thus, the lack of Y chromosome data presents a problem in determining the kinship of the individuals at La Sagrera.

The simultaneous nature of the deposition suggests that the individuals interred at La Sagrera were victims of some kind of mass fatality. The anthropological analysis showed a lack of perimortem trauma associated with the skeletons (Balaguer, 2016) which, along with the presence of females and adolescents, most likely precludes the mass deposition at La Sagrera resulting from battle or warfare. Moreover, stable isotope analysis conducted on the same skeletons indicated a difference in diet between adults and juveniles (Fernandez-Dominguez et al, 2018). This could indicate an agricultural disaster that may have resulted in a famine event. Multiple fatalities over a short time frame would require mass burials to occur relatively quickly, hindering the application of usual rites associated with burial, which may support the apparent lack of kinship within the clusters. Thus, it is likely that the individuals interred at La Sagrera do not represent a single population.

4.3 The People of Neolithic Iberia

4.3.1 Genetic distance

Genetic distances show that Neolithic Iberian populations were divergent from their pre-Neolithic predecessors, while being similar to one another. This is consistent with findings for Central Europeans (Bramanti et al, 2009), and suggests an influx of migrants to the Iberian Peninsula during the Early Neolithic period (Szécsényi-Nagy et al, 2017). Progressing through the Neolithic, and in concordance with the findings of Szécsényi-Nagy et al (2017), F_{ST} values indicate genetic continuity from the Early Neolithic to the Middle and Late Neolithic. This is in contrast to several regional studies (Gamba et al,

2012; Gómez-Sánchez et al, 2014), where a discontinuity is seen between Early and Middle Neolithic populations.

4.3.2 Haplogroup analysis

An increase in haplogroup diversity coinciding with the start of the Neolithic was found. Pre-Neolithic populations comprised of haplogroups H, M, T and U, while haplogroups HV, I, J, K, N, V, W and X all appear in Early Neolithic populations. The appearance of new lineages coinciding with the Early Neolithic supports a human migration model. Middle Neolithic Iberians, including the two individuals studied here, display an apparent reduction in haplogroup diversity compared to their Early Neolithic predecessors, possibly due to genetic drift. Haplogroups I, N, V, W and X are all absent in Middle Neolithic populations. Similar levels of haplogroup diversity is seen in the Late Neolithic population, although haplogroups N and X are both present, while haplogroup HV is absent. This reduced diversity could indicate a lack of subsequent immigration during the Middle to Late Neolithic period.

Interestingly, Early, Middle and Late Neolithic populations all retain a relatively high frequency of Mesolithic haplogroup U. Archaeological and aDNA evidence suggests coexistence of autochthonous hunter-gatherer groups in Iberia and Central Europe for several millennia after the arrival of the first Neolithic settlers (Bertranpetit and Cavalli-Sforza, 1991; Bollongino et al, 2013; Szécsényi-Nagy et al, 2017). While a resurgence of Mesolithic ancestry is seen across much of Europe during the Middle Neolithic (Haak et al, 2015), the presence of haplogroup U throughout the Neolithic in Iberia suggests admixture between the groups from the outset in this region (Szécsényi-Nagy et al, 2017). The relatively high frequency of haplogroup H in Iberian Neolithic populations compared to those of Central Europe also indicates a higher level of Mesolithic ancestry (Lacan et al, 2011; Brotherton et al, 2013; de-la-Rua et al, 2015). Interestingly, the frequency of haplogroup U increases to approximately one third in Middle Neolithic populations suggesting increased hunter-gatherer ancestry during this time. While it is possible that this Mesolithic signature was incorporated into subsequent Neolithic migrants prior to their arrival in Iberia (García-Martínez de Lagrán, Fernández-Domínguez and Rojo-Guerra,

2017), Lipson et al (2017) suggest that admixture between local hunter-gatherers and farmers is likely to have occurred.

4.4 Contribution of Neolithic Iberian populations to the modern gene pool

The Middle Neolithic Portuguese population was not significantly divergent from any extant Iberian population except the Basques, and had the greatest affinities with North and Central Portuguese populations. Gamba et al (2012) and Sampietro et al (2007) found genetic continuity between Neolithic farmers and extant populations in Aragon and Catalonia. Middle and Late Neolithic populations have been found to have a stronger affinity to modern populations than the Early Neolithic (Gamba et al, 2012). These results may suggest that the demographic changes coinciding with the onset of the Neolithic may not have influenced the modern gene pool to the extent of subsequent migrations. Genetic signatures consistent with both earlier (e.g. Lacan et al, 2011) and later (e.g. Adams et al, 2008; Pardiñas et al, 2012) population movements have been detected in extant populations. Subsequent migrations during the Bronze Age are likely to have reshaped the genetic landscape throughout much of Europe (Allentoft et al, 2015; Haak et al, 2015). Additionally, the Mesolithic contribution may have been underestimated (Lipson et al, 2017). The Iberian Peninsula was subject to numerous small migration events throughout its history (Pardiñas et al, 2012), thus it is likely that the genetic structure of modern day populations will contain signatures from across a range of events.

Genetic distances indicate that the Late Neolithic Spanish population was divergent from all modern Spanish populations, suggesting that demographic changes occurring in the post-Neolithic period had a role in shaping the modern Iberian mitochondrial gene pool. This may highlight the importance of geographical scale when interpreting such data; the Late Neolithic Spanish population here incorporates individuals from various locations across northern Spain. Extant Iberian populations have higher genetic diversity when compared to the rest of Europe (Wang, Zöllner and Rosenberg, 2012), and are quite distinct from one another. This may preclude the determination of accurate genetic affinities with a mixed group such as the Late Neolithic Spain metapopulation.

4.5 Conclusion

Despite multiple attempts, many of the samples initially processed – particularly those from open-air sites in Portugal – were too degraded for DNA sequences to be retrieved using a classical PCR approach. Samples from the La Sagrera site in Spain proved easier to amplify. The hypogeum at La Sagrera did not appear to be a typical burial site often associated with Late Neolithic populations. In fact, it is unlikely that the individuals interred there were a discrete population. A more likely explanation for the presence of the hypogeum is a mass fatality due to famine or disease epidemic.

The obtained results show that there was genetic continuity from the Early Neolithic through to the end of the Late Neolithic in Iberia, and although an apparent discontinuity from the Mesolithic to the Neolithic was observed, Neolithic Iberians retained a higher level of Mesolithic ancestry than their Central European counterparts. The contribution of Neolithic Iberians to Modern populations was unclear, though it is likely that subsequent migrations in the Bronze Age will have shaped the modern genetic landscape considerably.

4.6 Future Work

The poor DNA retrieval rate throughout this study limited the subsequent scope for analyses and hypotheses testing. Pursuing genomic sequencing using next generation high-throughput sequencing techniques may increase success rates. Further DNA extraction attempts might also benefit from better sampling; petrous bone (Gamba, 2014) and tooth cementum (Adler et al, 2011) have been demonstrated to result in superior DNA preservation, and may increase the amplification success rate. Targeting Y chromosome and/or autosomal markers may reveal further insights into the populations at La Sagrera and Barrão. This work should be done in order to to increase the genetic information available for the Iberian Peninsula to further the understanding of the demographic process underpinning the populations of Neolithic Iberia.

References

Achilli, A., Rengo, C., Magri, C., Battaglia, V., Olivieri, A., Scozzari, R., Cruciani, F., Zeviani, M., Briem, E., Carelli, V., Moral, P., Dugoujon, J.-M., Roostalu, U., Loogväli, E.-L., Kivisild, T., Bandelt, H.-J., Richards, M., Villems, R., Santachiara-Benerecetti, A.S., Semino, O. and Torroni, A. (2004) The Molecular Dissection of mtDNA Haplogroup H Confirms That the Franco-Cantabrian Glacial Refuge Was a Major Source for the European Gene Pool. *The American Journal of Human Genetics*, 75 (5), 910-918.

Adams, S.M., Bosch, E., Balaresque, P.L., Ballereau, S.J., Lee, A.C., Arroyo, E., López-Parra, A.M., Aler, M., Grifo, M.S.G., Brion, M., Carracedo, A., Lavinha, J., Martínez-Jarreta, B., Quintana-Murci, L., Picornell, A., Ramon, M., Skorecki, K., Behar, D.M., Calafell, F. and Jobling, M.A. (2008) The Genetic Legacy of Religious Diversity and Intolerance: Paternal Lineages of Christians, Jews, and Muslims in the Iberian Peninsula. *The American Journal of Human Genetics*, 83 (6), 725-736.

Adler, C.J., Haak, W., Donlon, D. and Cooper, A. (2011) Survival and recovery of DNA from ancient teeth and bones. *Journal of Archaeological Science*, 38 (5), 956-964.

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

Allentoft, M.E., Sikora, M., Sjogren, K.-G., Rasmussen, S., Rasmussen, M., Stenderup, J., Damgaard, P.B., Schroeder, H., Ahlstrom, T., Vinner, L., Malaspinas, A.-S., Margaryan, A., Higham, T., Chivall, D., Lynnerup, N., Harvig, L., Baron, J., Casa, P.D., Dabrowski, P., Duffy, P.R., Ebel, A.V., Epimakhov, A., Frei, K., Furmanek, M., Gralak, T., Gromov, A., Gronkiewicz, S., Grupe, G., Hajdu, T., Jarysz, R., Khartanovich, V., Khokhlov, A., Kiss, V., Kolar, J., Kriiska, A., Lasak, I., Longhi, C., McGlynn, G., Merkevicius, A., Merkyte, I., Metspalu, M., Mkrtchyan, R., Moiseyev, V., Paja, L., Palfi, G., Pokutta, D., Pospieszny, L., Price, T.D., Saag, L., Sablin, M., Shishlina, N., Smrcka, V., Soenov, V.I., Szeverenyi, V., Toth, G., Trifanova, S.V., Varul, L., Vicze, M., Yepiskoposyan, L., Zhitenev, V., Orlando, L., Sicheritz-Ponten, T., Brunak, S., Nielsen, R., Kristiansen, K. and Willerslev, E. (2015) Population genomics of Bronze Age Eurasia. *Nature*, 522 (7555), 167-172.

Alt, K.W., Zesch, S., Garrido-Pena, R., Knipper, C., Szécsényi-Nagy, A., Roth, C., Tejedor-Rodríguez, C., Held, P., García-Martínez-de-Lagrán, Í., Navitainuck, D., Arcusa Magallón, H. and Rojo-Guerra, M.A. (2016) A Community in Life and Death: The Late Neolithic Megalithic Tomb at Alto de Reinoso (Burgos, Spain). *PLoS ONE*, 11 (1), e0146176.

Álvarez-Iglesias, V., Mosquera-Miguel, A., Cerezo, M., Quintáns, B., Zarrabeitia, M.T., Cuscó, I., Lareu, M.V., García, O., Pérez-Jurado, L., Carracedo, A. and Salas, A. (2009) New population and phylogenetic features of the internal variation within mitochondrial DNA macro-haplogroup R0. *PLoS ONE*, 4.

Alvarez, L., Santos, C., Ramos, A., Pratdesaba, R., Francalacci, P. and Aluja, M.P. (2010) Mitochondrial DNA patterns in the Iberian Northern plateau: population dynamics and substructure of the Zamora province. *American Journal of Physical Anthropology*, 142, 531-539.

Ammerman, A.J. and Cavalli-Sforza, L.L. (1984) *The Neolithic Transition and the Genetics of Populations in Europe*. Princeton University Press.

Anderson, S., Bankier, A.T., Barrell, B.G., de Bruijn, M.H.L., Coulson, A.R., Drouin, J., Eperon, I.C., Nierlich, D.P., Roe, B.A., Sanger, F., Schreier, P.H., Smith, A.J.H., Staden, R. and Young, I.G. (1981) Sequence and organization of the human mitochondrial genome. *Nature*, 290, 457-465.

Andrews, R.M., Kubacka, I., Chinnery, P.F., Lightowlers, R.N., Turnbull, D.M. and Howell, N. (1999) Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nature Genetics*, 23, 147.

Aubán, J.B., Lozano, S. and Pardo-Gordó, S. (2017) Iberian Neolithic Networks: The Rise and Fall of the Cardial World. *Frontiers in Digital Humanities*, 4 (7).

Aubán, J.B., Martín, A.M. and Barton, C.M. (2013) Complex systems, social networks, and the evolution of social complexity in the east of Spain from the Neolithic to pre-Roman times. In: Berrocal, M. C., Sanjuán, L. G. and Gilman, A. (ed.) *The Prehistory of Iberia: Debating Early Social Stratification and the State.* New York: Taylor and Francis. pp. 53-73.

Austin, J.J., Ross, A.J., Smith, A.B., Fortey, R.A. and Thomas, R.H. (1997) Problems of reproducibility--does geologically ancient DNA survive in amber-preserved insects? *Proceedings of the Royal Society B: Biological Sciences*, 264 (1381), 467-474.

Balaguer, P. (2016) *Les agrupacions poblacionals del segon moment d'us de l'hipogeu funerari de la Sagrera. Estudi antropologic (Technical report). Servei d'Arqueologica. Ajuntament de Barcenlona,*. Barcelona.

Balaguer, P., García, P., Tenza, A. and Antequera, F. (2013) L'hipogeu funerari de la Sagrera (Barcelona). Resultats preliminars. *Rev. Arqueol. Ponent* (23), 77-88.

Behar, Doron M., van Oven, M., Rosset, S., Metspalu, M., Loogväli, E.-L., Silva, Nuno M., Kivisild, T., Torroni, A. and Villems, R. (2012) A “Copernican” Reassessment of the Human Mitochondrial DNA Tree from its Root. *The American Journal of Human Genetics*, 90 (4), 675-684.

Behar, D.M., Villems, R., Soodyall, H., Blue-Smith, J., Pereira, L., Metspalu, E., Scozzari, R., Makkan, H., Tzur, S., Comas, D., Bertranpetit, J., Quintana-Murci, L., Tyler-Smith, C., Wells, R.S. and Rosset, S. (2008) The Dawn of Human Matrilineal Diversity. *The American Journal of Human Genetics*, 82 (5), 1130-1140.

Bentley, R.A., Bickle, P., Fibiger, L., Nowell, G.M., Dale, C.W., Hedges, R.E.M., Hamilton, J., Wahl, J., Francken, M., Grupe, G., Lenneis, E., Teschler-Nicola, M., Arbogast, R.-M., Hofmann, D. and Whittle, A. (2012) Community differentiation and kinship among Europe's first farmers. *Proceedings of the National Academy of Sciences*, 109 (24), 9326-9330.

Bertranpetit, J. and Cavalli-Sforza, L.L. (1991) A genetic reconstruction of the history of the population of the Iberian Peninsula. *Annals of Human Genetics*, 55 (1), 51-67.

Bertranpetit, J., Salaberria, J., Calafell, F., A., U.P., Moral, P. and Comas, D. (1995) Human mitochondrial DNA variation and the origin of the Basques. *Annals of Human Genetics*, 59, 63-81.

Bollongino, R., Nehlich, O., Richards, M.P., Orschiedt, J., Thomas, M.G., Sell, C., Fajkošová, Z., Powell, A. and Burger, J. (2013) 2000 Years of Parallel Societies in Stone Age Central Europe. *Science*, 342, 479-481.

Bramanti, B., Thomas, M.G., Haak, W., Unterlaender, M., Jores, P., Tambets, K., Antanaitis-Jacobs, I., Haidle, M.N., Jankauskas, R., Kind, C.J., Lueth, F., Terberger, T., Hiller, J., Matsumura, S., Forster, P. and Burger, J. (2009) Genetic Discontinuity Between Local Hunter-Gatherers and Central Europe's First Farmers. *Science*, 326 (5949), 137-140.

Brandt, G., Haak, W., Adler, C.J., Roth, C., Szécsényi-Nagy, A., Karimnia, S., Möller-Rieker, S., Meller, H., Ganslmeier, R., Friederich, S., Dresely, V., Nicklisch, N., Pickrell, J.K., Sirocko, F., Reich, D., Cooper, A. and Alt, K.W. (2013) Ancient DNA Reveals Key Stages in the Formation of Central European Mitochondrial Genetic Diversity. *Science*, 342 (6155), 257-261.

Brotherton, P., Haak, W., Templeton, J., Brandt, G., Soubrier, J., Jane Adler, C., Richards, S.M., Sarkissian, C.D., Ganslmeier, R., Friederich, S., Dresely, V., van Oven, M., Kenyon, R., Van der Hoek, M.B., Korlach, J., Luong, K., Ho, S.Y.W., Quintana-Murci, L., Behar, D.M., Meller, H., Alt, K.W., Cooper, A. and The Genographic, C. (2013) Neolithic mitochondrial haplogroup H genomes and the genetic origins of Europeans. 4, 1764.

Brown, T. and Brown, K. (2011) *Biomolecular Archaeology : An Introduction*. United Kingdom: John Wiley & Sons, Incorporated.

Butler, J.M. (2010) *Fundamentals of forensic DNA typing*. Amsterdam, Boston: Academic Press.

Caramelli, D., Vernesi, C., Sanna, S., Sampietro, L., Lari, M., Castrì, L., Vona, G., Floris, R., Francalacci, P., Tykot, R., Casoli, A., Bertranpetit, J., Lalueza-Fox, C., Bertorelle, G. and Barbujani, G. (2007) Genetic variation in prehistoric Sardinia. *Human Genetics*, 122 (3), 327-336.

Cardoso, S., Alfonso-Sanchez, M.A., Valverde, L., Odriozola, A., Perez-Miranda, A.M. and Pena, J.A. (2011) The maternal legacy of Basques in northern Navarre: new insights into the mitochondrial

DNA diversity of the Franco-Cantabrian area. *American Journal of Physical Anthropology*, 145, 480-488.

Carvalho, A.F., Alves-Cardoso, F., Gonçalves, D., Granja, R., Cardoso, J.L., Dean, R.M., Gibaja, J.F., Masucci, M.A., Arroyo-Pardo, E., Fernández-Domínguez, E., Petchey, F., Douglas Price, T., Mateus, J.E., Queiroz, P.F., Callapez, P., Pimenta, C. and Regala, F.T. (2016) The Bom Santo Cave (Lisbon, Portugal): Catchment, Diet, and Patterns of Mobility of a Middle Neolithic Population. *European Journal of Archaeology*, 19 (2), 187-214.

Carvalho, A.F., Antunes-Ferreira, N. and Valente, M.J. (2003) A gruta-necrópole neolíticado Algar do Barrão (Monsanto, Alcanena). *Revista Portuguesa do Arqueologia*, 6 (1), 101-119.

Carvalho, A.F. and Petchey, F. (2013) Stable Isotope Evidence of Neolithic Palaeodiets in the Coastal Regions of Southern Portugal. *The Journal of Island and Coastal Archaeology*, 8 (3), 361-383.

Carvalho, A.F. and Rocha, L. (2016) Datação direta e análise de paleodietas dos indivíduos da Anta de Cabeceira. *digitAR Revista Digital de Arqueologia, Arquitectura e Artes*, 53-61.

Casas, M.J., Hagelberg, E., Fregel, R., Larruga, J.M. and González, A.M. (2006) Human mitochondrial DNA diversity in an archaeological site in Al-Andalus: genetic impact of migrations from North Africa in Medieval Spain. *American Journal of Physical Anthropology*, 131, 539-551.

Cooper, A. and Poinar, H.N. (2000) Ancient DNA: Do It Right or Not at All. *Science*, 289 (5482), 1139.

Cornish-Bowden, A. (1985) Nomenclature for incompletely specified bases in nucleic acid sequences: recommendations 1984. *Nucleic Acids Research*, 13 (9), 3021-3030.

Côrte-Real, H.B.S.M., Macaulay, V.A., Richards, M.B., Hariti, G., Issad, M.S., Cambon-Thomsen, A., Papiha, S., Bertranpetit, J. and Sykes, B.C. (1996) Genetic diversity in the Iberian Peninsula determined from mitochondrial sequence analysis. *Annals of Human Genetics*, 60 (4), 331-350.

Dabney, J., Knapp, M., Glocke, I., Gansauge, M.-T., Weihmann, A., Nickel, B., Valdiosera, C., García, N., Pääbo, S., Arsuaga, J.-L. and Meyer, M. (2013) Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proceedings of the National Academy of Sciences*, 110 (39), 15758-15763.

de-la-Rua, C., Izagirre, N., Alonso, S. and Hervella, M. (2015) Ancient DNA in the Cantabrian fringe populations: A mtDNA study from Prehistory to Late Antiquity. *Quaternary International*, 364 (Supplement C), 306-311.

Dennell, R. (1983) *European economic prehistory: a new approach*. London/ New York: Academic Press.

Der Sarkissian, C., Balanovsky, O., Brandt, G., Khartanovich, V., Buzhilova, A., Koshel, S., Zaporozhchenko, V., Gronenborn, D., Moiseyev, V., Kolpakov, E., Shumkin, V., Alt, K.W., Balanovska, E., Cooper, A., Haak, W. and the Genographic, C. (2013) Ancient DNA Reveals Prehistoric Gene-Flow from Siberia in the Complex Human Population History of North East Europe. *PLOS Genetics*, 9 (2), e1003296.

Di Renno, F. (2016) *Population dynamics in Neolithic Northeastern Iberia: An isotopic and ancient DNA approach*. M.Sc thesis, Durham University.

Excoffier, L., Laval, G. and Schneider, S. (2005) Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, 1, 47-50.

Falchi, A., Giovannoni, L., Calo, C., Piras, I., Moral, P., Paoli, G., Vona, G. and Varesi, L. (2006) Genetic history of some western Mediterranean human isolates through mtDNA HVR1 polymorphisms. *Journal Of Human Genetics*, 51 (1), 9-14.

Fehren-Schmitz, L., Reindel, M., Cagigao, E.T., Hummel, S. and Herrmann, B. (2010) Pre-Columbian population dynamics in coastal southern Peru: A diachronic investigation of mtDNA patterns in the Palpa region by ancient DNA analysis. *American Journal of Physical Anthropology*, 141 (2), 208-221.

Fehren-Schmitz, L., Warnberg, O., Reindel, M., Seidenberg, V., Tomasto-Cagigao, E., Isla-Cuadrado, J., Hummel, S. and Herrmann, B. (2011) Diachronic Investigations of Mitochondrial and Y-Chromosomal Genetic Markers in Pre-Columbian Andean Highlanders from South Peru. *Annals of Human Genetics*, 75 (2), 266-283.

Fernandez-Dominguez, E., Di Renno, F., Robinson, C., Towers, J., Nowell, G.M., Molist, M. and Montgomery, J. (2018) Dying in the Neolithic: isotope analysis of the Late Neolithic mass grave of La Sagrera, Barcelona. *Annual meeting of the European association of archaeologists*,

Fernandez, E. (2005) *Polimorfismos de DNA mitocondrial en poblaciones antiguas de la cuenca mediterranea.* Doctor of Philosophy thesis, Universidad da Barcelona.

Fregel, R. and Delgado, S. (2011) HaploSearch: A tool for haplotype-sequence two-way transformation. *Mitochondrion*, 11 (2), 366-367.

Fu, Q., Posth, C., Hajdinjak, M., Petr, M., Mallick, S., Fernandes, D., Furtwängler, A., Haak, W., Meyer, M., Mittnik, A., Nickel, B., Peltzer, A., Rohland, N., Slon, V., Talamo, S., Lazaridis, I., Lipson, M., Mathieson, I., Schiffels, S., Skoglund, P., Derevianko, A.P., Drozdov, N., Slavinsky, V., Tsybankov, A., Cremonesi, R.G., Mallegni, F., Gély, B., Vacca, E., Morales, M.R.G., Straus, L.G.,

Neugebauer-Maresch, C., Teschler-Nicola, M., Constantin, S., Moldovan, O.T., Benazzi, S., Peresani, M., Coppola, D., Lari, M., Ricci, S., Ronchitelli, A., Valentin, F., Thevenet, C., Wehrberger, K., Grigorescu, D., Rougier, H., Crevecoeur, I., Flas, D., Semal, P., Mannino, M.A., Cupillard, C., Bocherens, H., Conard, N.J., Harvati, K., Moiseyev, V., Drucker, D.G., Svoboda, J., Richards, M.P., Caramelli, D., Pinhasi, R., Kelso, J., Patterson, N., Krause, J., Pääbo, S. and Reich, D. (2016) The genetic history of Ice Age Europe. *Nature*, 534, 200-205.

Fulton, T.L. (2012) Setting Up an Ancient DNA Laboratory. In: Shapiro, B. and Hofreiter, M. (ed.) *Ancient DNA: Methods and Protocols.* Totowa, NJ: Humana Press. pp. 1-11.

Gamba, C., Fernández, E., Oliver, A., Tirado, M., Baeza, C., López-Parra, A.M. and Arroyo-Pardo, E. (2008) Population genetics and DNA preservation in ancient human remains from Eastern Spain. *Forensic Science International: Genetics Supplement Series*, 1 (1), 462-464.

Gamba, C., Fernandez, E., Tirado, M., Deguilloux, M.F., Pemonge, M.H., Utrilla, P., Edo, M., Molist, M., Rasteiro, R., Chikhi, L. and Arroyo-Pardo, E. (2012) Ancient DNA from an Early Neolithic Iberian population supports a pioneer colonization by first farmers. *Molecular Ecology*, 21 (1), 45-56.

Gamba, C., Jones, E.R., Teasdale, M.D., McLaughlin, R.L., Gonzalez-Fortes, G., Mattiangeli, V., Domboróczki, L., Kővári, I., Pap, I., Anders, A., Whittle, A., Dani, J., Raczky, P., Higham, T.F.G., Hofreiter, M., Bradley, D.G. and Pinhasi, R. (2014) Genome flux and stasis in a five millennium transect of European prehistory. *Nature Communications*, 5, 5257.

García-Garcerà, M., Gigli, E., Sanchez-Quinto, F., Ramirez, O., Calafell, F., Civit, S. and Lalueza-Fox, C. (2011) Fragmentation of Contaminant and Endogenous DNA in Ancient Samples Determined by Shotgun Sequencing; Prospects for Human Palaeogenomics. *PLoS ONE*, 6 (8), e24161.

García-Martínez de Lagrán, Í., Fernández-Domínguez, E. and Rojo-Guerra, M.A. (2017) Solutions or illusions? An analysis of the available palaeogenetic evidence from the origins of the Neolithic in the Iberian Peninsula. *Quaternary International*, 83-103.

García, O., Fregel, R., Larruga, J.M., Álvarez, V., Yurrebaso, I., Cabrera, V.M. and González, A.M. (2011) Using mitochondrial DNA to test the hypothesis of a European post-glacial human recolonization from the Franco-Cantabrian refuge. *Heredity*, 106 (1), 37-45.

Gomes, M.V. (2008) Castelo Belinho (Algarve, Portugal) and the first Southwest Iberian villagers. *The Early Neolithic in the Iberian Peninsula. Regional and Transregional Components*, 71-78

Gomes, M.V. (2012) Early Neolithic funerary practices in Castelo Belinho's village. In: Gibaja, J. F., Carvalho, A. F. and Chambon, P. (ed.) *Funerary Practices in the Iberian Peninsula from theMesolithic to the Chalcolithic.* Oxford: Archaeopress. pp. 113-123.

Gómez-Sánchez, D., Olalde, I., Pierini, F., Matas-Lalueza, L., Gigli, E., Lari, M., Civit, S., Lozano, M., Vergès, J.M., Caramelli, D., Ramírez, O. and Lalueza-Fox, C. (2014) Mitochondrial DNA from El Mirador Cave (Atapuerca, Spain) Reveals the Heterogeneity of Chalcolithic Populations. *PLoS ONE*, 9 (8), e105105.

Günther, T., Valdiosera, C., Malmström, H., Ureña, I., Rodriguez-Varela, R., Sverrisdóttir, Ó.O., Daskalaki, E.A., Skoglund, P., Naidoo, T., Svensson, E.M., Bermúdez de Castro, J.M., Carbonell, E., Dunn, M., Storå, J., Iriarte, E., Arsuaga, J.L., Carretero, J.-M., Götherström, A. and Jakobsson, M. (2015) Ancient genomes link early farmers from Atapuerca in Spain to modern-day Basques. *Proceedings of the National Academy of Sciences*, 112 (38), 11917-11922.

Haak, W., Balanovsky, O., Sanchez, J.J., Koshel, S., Zaporozhchenko, V., Adler, C.J., Der Sarkissian, C.S.I., Brandt, G., Schwarz, C., Nicklisch, N., Dresely, V., Fritsch, B., Balanovska, E., Villems, R., Meller, H., Alt, K.W., Cooper, A. and the Genographic, C. (2010) Ancient DNA from European Early Neolithic Farmers Reveals Their Near Eastern Affinities. *PLOS Biology*, 8 (11), e1000536.

Haak, W., Brandt, G., Jong, H.N.d., Meyer, C., Ganslmeier, R., Heyd, V., Hawkesworth, C., Pike, A.W.G., Meller, H. and Alt, K.W. (2008) Ancient DNA, Strontium isotopes, and osteological analyses shed light on social and kinship organization of the Later Stone Age. *Proceedings of the National Academy of Sciences*, 105 (47), 18226-18231.

Haak, W., Forster, P., Bramanti, B., Matsumura, S., Brandt, G., Tänzer, M., Villems, R., Renfrew, C., Gronenborn, D., Alt, K.W. and Burger, J. (2005) Ancient DNA from the First European Farmers in 7500-Year-Old Neolithic Sites. *Science*, 310 (5750), 1016-1018.

Haak, W., Lazaridis, I., Patterson, N., Rohland, N., Mallick, S., Llamas, B., Brandt, G., Nordenfelt, S., Harney, E., Stewardson, K., Fu, Q., Mittnik, A., Bánffy, E., Economou, C., Francken, M., Friederich, S., Pena, R.G., Hallgren, F., Khartanovich, V., Khokhlov, A., Kunst, M., Kuznetsov, P., Meller, H., Mochalov, O., Moiseyev, V., Nicklisch, N., Pichler, S.L., Risch, R., Rojo Guerra, M.A., Roth, C., Szécsényi-Nagy, A., Wahl, J., Meyer, M., Krause, J., Brown, D., Anthony, D., Cooper, A., Alt, K.W. and Reich, D. (2015) Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature*, 522, 207-211.

Hedges, S.B. and Schweitzer, M.H. (1995) Detecting dinosaur DNA. *Science*, 268 (5214), 1191-1192.

Hervella, M., Izagirre, N., Alonso, S., Fregel, R., Alonso, A., Cabrera, V.M. and de la Rúa, C. (2012) Ancient DNA from Hunter-Gatherer and Farmer Groups from Northern Spain Supports a Random Dispersion Model for the Neolithic Expansion into Europe. *PLoS ONE*, 7 (4), e34417.

Heyer, E., Zietkiewicz, E., Rochowski, A., Yotova, V., Puymirat, J. and Labuda, D. (2001) Phylogenetic and Familial Estimates of Mitochondrial Substitution Rates: Study of Control Region Mutations in Deep-Rooting Pedigrees. *The American Journal of Human Genetics*, 69 (5), 1113-1126.

Higuchi, R., Bowman, B., Freiberger, M., Ryder, O.A. and Wilson, A.C. (1984) DNA sequences from the quagga, an extinct member of the horse family. *Nature*, 312, 282-284.

Hofreiter, M., Jaenicke, V., Serre, D., Haeseler, A.v. and Pääbo, S. (2001) DNA sequences from multiple amplifications reveal artifacts induced by cytosine deamination in ancient DNA. *Nucleic Acids Research*, 29 (23), 4793-4799.

Hofreiter, M., Paijmans, J.L.A., Goodchild, H., Speller, C.F., Barlow, A., Fortes, G.G., Thomas, J.A., Ludwig, A. and Collins, M.J. (2015) The future of ancient DNA: Technical advances and conceptual shifts. *BioEssays*, 37 (3), 284-293.

Höss, M., Jaruga, P., Zastawny, T.H., Dizdaroglu, M. and Pääbo, S. (1996) DNA damage and DNA sequence retrieval from ancient tissues. *Nucleic Acids Research*, 24 (7), 1304-1307.

Hughey, J.R., Paschou, P., Drineas, P., Mastropaolo, D., Lotakis, D.M., Navas, P.A., Michalodimitrakis, M., Stamatoyannopoulos, J.A. and Stamatoyannopoulos, G. (2013) A European population in Minoan Bronze Age Crete. *Nature Communications*, 4, 1861.

Kloss-Brandstätter, A., Pacher, D., Schönherr, S., Weissensteiner, H., Binna, R., Specht, G. and Kronenberg, F. (2011) HaploGrep: a fast and reliable algorithm for automatic classification of mitochondrial DNA haplogroups. *Human Mutation*, 32 (1), 25-32.

Lacan, M., Keyser, C., Ricaut, F.-X., Brucato, N., Tarrús, J., Bosch, A., Guilaine, J., Crubézy, E. and Ludes, B. (2011) Ancient DNA suggests the leading role played by men in the Neolithic dissemination. *Proceedings of the National Academy of Sciences*, 108 (45), 18255-18259.

Lamers, R., Hayter, S. and Matheson, C.D. (2009) Postmortem Miscoding Lesions in Sequence Analysis of Human Ancient Mitochondrial DNA. *Journal of Molecular Evolution*, 68 (1), 40-55.

Lazaridis, I., Patterson, N., Mittnik, A., Renaud, G., Mallick, S., Kirsanow, K., Sudmant, P.H., Schraiber, J.G., Castellano, S., Lipson, M., Berger, B., Economou, C., Bollongino, R., Fu, Q., Bos, K.I., Nordenfelt, S., Li, H., de Filippo, C., Prüfer, K., Sawyer, S., Posth, C., Haak, W., Hallgren, F., Fornander, E., Rohland, N., Delsate, D., Francken, M., Guinet, J.-M., Wahl, J., Ayodo, G., Babiker, H.A., Bailliet, G., Balanovska, E., Balanovsky, O., Barrantes, R., Bedoya, G., Ben-Ami, H., Bene, J., Berrada, F., Bravi, C.M., Brisighelli, F., Busby, G.B.J., Cali, F., Churnosov, M., Cole, D.E.C., Corach, D., Damba, L., van Driem, G., Dryomov, S., Dugoujon, J.-M., Fedorova, S.A., Romero, I.G., Gubina, M., Hammer, M., Henn, B.M., Hervig, T., Hodoglugil, U., Jha, A.R., Karachanak-Yankova, S., Khusainova, R., Khusnutdinova, E., Kittles, R., Kivisild, T., Klitz, W., Kučinskas, V., Kushniarevich, A., Laredj, L., Litvinov, S., Loukidis, T., Mahley, R.W., Melegh, B., Metspalu, E., Molina, J., Mountain, J., Näkkäläjärvi, K., Nesheva, D., Nyambo, T., Osipova, L., Parik, J., Platonov, F., Posukh, O., Romano, V., Rothhammer, F., Rudan, I., Ruizbakiev, R., Sahakyan, H., Sajantila, A., Salas, A., Starikovskaya, E.B., Tarekegn, A., Toncheva, D., Turdikulova, S., Uktveryte, I., Utevska, O., Vasquez, R., Villena, M., Voevoda, M., Winkler, C., Yepiskoposyan, L., Zalloua, P., Zemunik, T., Cooper, A., Capelli, C., Thomas, M.G., Ruiz-Linares, A., Tishkoff, S.A., Singh, L., Thangaraj, K., Villems, R., Comas, D.,

Sukernik, R., Metspalu, M., Meyer, M., Eichler, E.E., Burger, J., Slatkin, M., Pääbo, S., Kelso, J., Reich, D. and Krause, J. (2014) Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature*, 513 (7518), 409-413.

Lee, E.J., Renneberg, R., Harder, M., Krause-Kyora, B., Rinne, C., Müller, J., Nebel, A. and von Wurmb-Schwark, N. (2014) Collective burials among agro-pastoral societies in later Neolithic Germany: perspectives from ancient DNA. *Journal of Archaeological Science*, 51 (Supplement C), 174-180.

Leney, M.D. (2006) Sampling Skeletal Remains for Ancient DNA (aDNA): A Measure of Success. *Historical Archaeology*, 40 (3), 31-49.

Lindahl, T. (1993) Instability and decay of the primary structure of DNA. *Nature*, 362, 709-715.

Lipson, M., Szécsényi-Nagy, A., Mallick, S., Pósa, A., Stégmár, B., Keerl, V., Rohland, N., Stewardson, K., Ferry, M., Michel, M., Oppenheimer, J., Broomandkhoshbacht, N., Harney, E., Nordenfelt, S., Llamas, B., Gusztáv Mende, B., Köhler, K., Oross, K., Bondár, M., Marton, T., Osztás, A., Jakucs, J., Paluch, T., Horváth, F., Csengeri, P., Koós, J., Sebők, K., Anders, A., Raczky, P., Regenye, J., Barna, J.P., Fábián, S., Serlegi, G., Toldi, Z., Gyöngyvér Nagy, E., Dani, J., Molnár, E., Pálfi, G., Márk, L., Melegh, B., Bánfai, Z., Domboróczki, L., Fernández-Eraso, J., Antonio Mujika-Alustiza, J., Alonso Fernández, C., Jiménez Echevarría, J., Bollongino, R., Orschiedt, J., Schierhold, K., Meller, H., Cooper, A., Burger, J., Bánffy, E., Alt, K.W., Lalueza-Fox, C., Haak, W. and Reich, D. (2017) Parallel palaeogenomic transects reveal complex genetic history of early European farmers. *Nature*, 551, 368-372.

Malmström, H., Gilbert, M.T.P., Thomas, M.G., Brandström, M., Storå, J., Molnar, P., Andersen, P.K., Bendixen, C., Holmlund, G., Götherström, A. and Willerslev, E. (2009) Ancient DNA Reveals Lack of Continuity between Neolithic Hunter-Gatherers and Contemporary Scandinavians. *Current Biology*, 19 (20), 1758-1762.

Malyarchuk, B., Derenko, M., Grzybowski, T., Perkova, M., Rogalla, U., Vanecek, T. and Tsybovsky, I. (2010) The Peopling of Europe from the Mitochondrial Haplogroup U5 Perspective. *PLoS ONE*, 5 (4), e10285.

Martiniano, R., Cassidy, L.M., Ó'Maoldúin, R., McLaughlin, R., Silva, N.M., Manco, L., Fidalgo, D., Pereira, T., Coelho, M.J., Serra, M., Burger, J., Parreira, R., Moran, E., Valera, A.C., Porfirio, E., Boaventura, R., Silva, A.M. and Bradley, D.G. (2017) The population genomics of archaeological transition in west Iberia: Investigation of ancient substructure using imputation and haplotype-based methods. *PLOS Genetics*, 13 (7), e1006852.

Mathieson, I., Lazaridis, I., Rohland, N., Mallick, S., Patterson, N., Roodenberg, S.A., Harney, E., Stewardson, K., Fernandes, D., Novak, M., Sirak, K., Gamba, C., Jones, E.R., Llamas, B., Dryomov, S., Pickrell, J., Arsuaga, J.L., de Castro, J.M.B., Carbonell, E., Gerritsen, F., Khokhlov, A., Kuznetsov, P., Lozano, M., Meller, H., Mochalov, O., Moiseyev, V., Guerra, M.A.R., Roodenberg, J., Vergès, J.M., Krause, J., Cooper, A., Alt, K.W., Brown, D., Anthony, D., Lalueza-Fox, C., Haak, W., Pinhasi, R.

and Reich, D. (2015) Genome-wide patterns of selection in 230 ancient Eurasians. *Nature*, 528 (7583), 499-503.

Melchior, L., Gilbert, M.T.P., Kivisild, T., Lynnerup, N. and Dissing, J. (2008) Rare mtDNA haplogroups and genetic differences in rich and poor Danish Iron-Age villages. *American Journal of Physical Anthropology*, 135 (2), 206-215.

Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G. and Erlich, H. (1986) Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harbor Symposia Quantitative Biology*, 51 Pt 1, 263-273.

Olalde, I., Schroeder, H., Sandoval-Velasco, M., Vinner, L., Lobón, I., Ramirez, O., Civit, S., García Borja, P., Salazar-García, D.C., Talamo, S., María Fullola, J., Xavier Oms, F., Pedro, M., Martínez, P., Sanz, M., Daura, J., Zilhão, J., Marquès-Bonet, T., Gilbert, M.T.P. and Lalueza-Fox, C. (2015) A Common Genetic Origin for Early Farmers from Mediterranean Cardial and Central European LBK Cultures. *Molecular Biology and Evolution*, 32 (12), 3132-3142.

Orlando, L., Ginolhac, A., Zhang, G., Froese, D., Albrechtsen, A., Stiller, M., Schubert, M., Cappellini, E., Petersen, B., Moltke, I., Johnson, P.L.F., Fumagalli, M., Vilstrup, J.T., Raghavan, M., Korneliussen, T., Malaspinas, A.-S., Vogt, J., Szklarczyk, D., Kelstrup, C.D., Vinther, J., Dolocan, A., Stenderup, J., Velazquez, A.M.V., Cahill, J., Rasmussen, M., Wang, X., Min, J., Zazula, G.D., Seguin-Orlando, A., Mortensen, C., Magnussen, K., Thompson, J.F., Weinstock, J., Gregersen, K., Røed, K.H., Eisenmann, V., Rubin, C.J., Miller, D.C., Antczak, D.F., Bertelsen, M.F., Brunak, S., Al-Rasheid, K.A.S., Ryder, O., Andersson, L., Mundy, J., Krogh, A., Gilbert, M.T.P., Kjær, K., Sicheritz-Ponten, T., Jensen, L.J., Olsen, J.V., Hofreiter, M., Nielsen, R., Shapiro, B., Wang, J. and Willerslev, E. (2013) Recalibrating Equus evolution using the genome sequence of an early Middle Pleistocene horse. *Nature*, 499, 74-78.

Özdoğan, M. (2014) A new look at the introduction of the Neolithic way of life in Southeastern Europe. Changing paradigms of the expansion of the Neolithic way of life. *Documenta Praehistorica*, 41, 33-49.

Pääbo, S. (1985) Molecular cloning of Ancient Egyptian mummy DNA. *Nature*, 314, 644-645.

Pääbo, S. (1989) Ancient DNA: extraction, characterization, molecular cloning, and enzymatic amplification. *Proceedings of the National Academy of Sciences of the United States of America*, 86 (6), 1939-1943.

Pakendorf, B. and Stoneking, M. (2005) Mitochondrial DNA and human evolution. *Annual Review of Genomics and Human Genetics*, 6 (1), 165-183.

Pardiñas, A.F., Roca, A., García-Vazquez, E. and López, B. (2012) Assessing the Genetic Influence of Ancient Sociopolitical Structure: Micro-differentiation Patterns in the Population of Asturias (Northern Spain). *PLoS ONE*, 7 (11), e50206.

Parson, W. and Dür, A. (2007) EMPOP—A forensic mtDNA database. *Forensic Science International: Genetics*, 1 (2), 88-92.

Pereira, L., Cunha, C. and Amorim, A. (2004) Predicting sampling saturation of mtDNA haplotypes: an application to an enlarged Portuguese database. *International Journal of Legal Medicine*, 118 (3), 132-136.

Pinto, F., Gonzalez, A., Hernandez, M., Larruga, J.M. and Cabrera, V.M. (1996) Genetic relationship between the Canary Islanders and their African and Spanish ancestors inferred from mitochondrial DNA sequences. *Annals of Human Genetics*, 60, 321-330.

Poinar, H.N., Cano, R.J. and Poinar Jr, G.O. (1993) DNA from an extinct plant. *Nature*, 363, 677.

Posth, C., Renaud, G., Mittnik, A., Drucker, Dorothée G., Rougier, H., Cupillard, C., Valentin, F., Thevenet, C., Furtwängler, A., Wißing, C., Francken, M., Malina, M., Bolus, M., Lari, M., Gigli, E., Capecchi, G., Crevecoeur, I., Beauval, C., Flas, D., Germonpré, M., van der Plicht, J., Cottiaux, R., Gély, B., Ronchitelli, A., Wehrberger, K., Grigorescu, D., Svoboda, J., Semal, P., Caramelli, D., Bocherens, H., Harvati, K., Conard, Nicholas J., Haak, W., Powell, A. and Krause, J. (2016) Pleistocene Mitochondrial Genomes Suggest a Single Major Dispersal of Non-Africans and a Late Glacial Population Turnover in Europe. *Current Biology*, 26 (6), 827-833.

Price, T.D. (2000) *Europe's first farmers*. Cambridge: Cambridge University Press.

Reynolds, L., Bertoncini, S. and Fernandez-Dominguez, E. (in prep.) Mitopop: A mitochondrial DNA database for forensic and population genetics analyses.

Ribe, G., Cruells, W. and Molist, M. (1997) The Neolithic of the Iberian Peninsula. In: Diaz-Andreu, M. and Keay, S. J. (ed.) *The archaeology of Iberia: the dynamics of change.* London: Routledge. pp. 65-84.

Richards, M., Macaulay, V., Hickey, E., Vega, E., Sykes, B., Guida, V., Rengo, C., Sellitto, D., Cruciani, F., Kivisild, T., Villems, R., Thomas, M., Rychkov, S., Rychkov, O., Rychkov, Y., Gölge, M., Dimitrov, D., Hill, E., Bradley, D., Romano, V., Calì, F., Vona, G., Demaine, A., Papiha, S., Triantaphyllidis, C., Stefanescu, G., Hatina, J., Belledi, M., Di Rienzo, A., Novelletto, A., Oppenheim, A., Nørby, S., Al-Zaheri, N., Santachiara-Benerecetti, S., Scozzari, R., Torroni, A. and Bandelt, H.-J. (2000) Tracing European Founder Lineages in the Near Eastern mtDNA Pool. *The American Journal of Human Genetics*, 67 (5), 1251-1276.

Richards, M.B., Macaulay, V.A., Bandelt, H.J. and Sykes, B.C. (1998) Phylogeography of mitochondrial DNA in western Europe. *Annals of Human Genetics*, 62 (3), 241-260.

Rizzi, E., Lari, M., Gigli, E., De Bellis, G. and Caramelli, D. (2012) Ancient DNA studies: new perspectives on old samples. *Genetics, Selection, Evolution : GSE*, 44 (1), 21-21.

Robin, E.D. and Wong, R. (1988) Mitochondrial DNA molecules and virtual number of mitochondria per cell in mammalian cells. *Journal of Cellular Physiology*, 136 (3), 507-513.

Robinson, C., Di Renno, F., Novell, G., Towers, J., Montgomery, J. and Fernandez, E. (2016) *Ancient DNA analysis and stable isotope analysis of a set of human skeletal remains from La Sagrera Hypogeum: insights into population structure, kinship, diet and mobility during the Late Neolithic*. Unpublished.

Rohland, N. and Hofreiter, M. (2007) Ancient DNA extraction from bones and teeth. *Nature Protocols*, 2, 1756-1762.

Rohland, N., Siedel, H. and Hofreiter, M. (2010) A rapid column-based ancient DNA extraction method for increased sample throughput. *Molecular Ecology Resources*, 10 (4), 677-683.

Rowley-Conwy, P. (2009) Human Prehistory: Hunting for the Earliest Farmers. *Current Biology*, 19 (20), R948-R949.

Salas, A., Comas, D., Lareu, M.V., Bertranpetit, J. and Carracedo, A. (1998) Mitochondrial DNA analysis of the Galician population: a genetic edge of European variation. *European Journal Of Human Genetics*, 6, 365-375.

Sampietro, M.L., Lao, O., Caramelli, D., Lari, M., Pou, R., Martí, M., Bertranpetit, J. and Lalueza-Fox, C. (2007) Palaeogenetic evidence supports a dual model of Neolithic spreading into Europe. *Proceedings of the Royal Society B: Biological Sciences*, 274 (1622), 2161-2167.

Sánchez-Quinto, F., Schroeder, H., Ramirez, O., Ávila-Arcos, María C., Pybus, M., Olalde, I., Velazquez, Amhed M.V., Marcos, María Encina P., Encinas, Julio Manuel V., Bertranpetit, J., Orlando, L., Gilbert, M.Thomas P. and Lalueza-Fox, C. (2012) Genomic Affinities of Two 7,000-Year-Old Iberian Hunter-Gatherers. *Current Biology*, 22 (16), 1494-1499.

Soares, P., Achilli, A., Semino, O., Davies, W., Macaulay, V., Bandelt, H.-J., Torroni, A. and Richards, M.B. (2010) The Archaeogenetics of Europe. *Current Biology*, 20 (4), R174-R183.

Sorensen, B. (2013) *A history of energy: Northern Europe from the Stone Age to the present day*. Routledge.

Stoneking, M. (1995) Ancient DNA: how do you know when you have it and what can you do with it? *American Journal of Human Genetics*, 57 (6), 1259-1262.

Stoneking, M. (2000) Hypervariable Sites in the mtDNA Control Region Are Mutational Hotspots. *American Journal of Human Genetics*, 67 (4), 1029-1032.

Sutlovic, D., Gamulin, S., Definis-Gojanovic, M., Gugic, D. and Andjelinovic, S. (2008) Interaction of humic acids with human DNA: Proposed mechanisms and kinetics. *Electrophoresis*, 29 (7), 1467-1472.

Szécsényi-Nagy, A., Brandt, G., Haak, W., Keerl, V., Jakucs, J., Möller-Rieker, S., Köhler, K., Mende, B.G., Oross, K., Marton, T., Osztás, A., Kiss, V., Fecher, M., Pálfi, G., Molnár, E., Sebők, K., Czene, A., Paluch, T., Šlaus, M., Novak, M., Pećina-Šlaus, N., Ősz, B., Voicsek, V., Somogyi, K., Tóth, G., Kromer, B., Bánffy, E. and Alt, K.W. (2015) Tracing the genetic origin of Europe's first farmers reveals insights into their social organization. *Proceedings of the Royal Society B: Biological Sciences*, 282 (1805), 20150339.

Szécsényi-Nagy, A., Roth, C., Brandt, G., Rihuete-Herrada, C., Tejedor-Rodríguez, C., Held, P., García-Martínez-de-Lagrán, Í., Arcusa Magallón, H., Zesch, S., Knipper, C., Bánffy, E., Friederich, S., Meller, H., Bueno Ramírez, P., Barroso Bermejo, R., de Balbín Behrmann, R., Herrero-Corral, A.M., Flores Fernández, R., Alonso Fernández, C., Jiménez Echevarria, J., Rindlisbacher, L., Oliart, C., Fregeiro, M.-I., Soriano, I., Vicente, O., Micó, R., Lull, V., Soler Díaz, J., López Padilla, J.A., Roca de Togores Muñoz, C., Hernández Pérez, M.S., Jover Maestre, F.J., Lomba Maurandi, J., Avilés Fernández, A., Lillios, K.T., Silva, A.M., Magalhães Ramalho, M., Oosterbeek, L.M., Cunha, C., Waterman, A.J., Roig Buxó, J., Martínez, A., Ponce Martínez, J., Hunt Ortiz, M., Mejías-García, J.C., Pecero Espín, J.C., Cruz-Auñón Briones, R., Tomé, T., Carmona Ballestero, E., Cardoso, J.L., Araújo, A.C., Liesau von Lettow-Vorbeck, C., Blasco Bosqued, C., Ríos Mendoza, P., Pujante, A., Royo-Guillén, J.I., Esquembre Beviá, M.A., Dos Santos Goncalves, V.M., Parreira, R., Morán Hernández, E., Méndez Izquierdo, E., Vega y Miguel, J., Menduiña García, R., Martínez Calvo, V., López Jiménez, O., Krause, J., Pichler, S.L., Garrido-Pena, R., Kunst, M., Risch, R., Rojo-Guerra, M.A., Haak, W. and Alt, K.W. (2017) The maternal genetic make-up of the Iberian Peninsula between the Neolithic and the Early Bronze Age. *Scientific Reports*, 7, 15644.

Torroni, A., Bandelt, H.-J., Macaulay, V., Richards, M., Cruciani, F., Rengo, C., Martinez-Cabrera, V., Villems, R., Kivisild, T., Metspalu, E., Parik, J., Tolk, H.-V., Tambets, K., Forster, P., Karger, B., Francalacci, P., Rudan, P., Janicijevic, B., Rickards, O., Savontaus, M.-L., Huoponen, K., Laitinen, V., Koivumäki, S., Sykes, B., Hickey, E., Novelletto, A., Moral, P., Sellitto, D., Coppa, A., Al-Zaheri, N., Santachiara-Benerecetti, A.S., Semino, O. and Scozzari, R. (2001) A Signal, from Human mtDNA, of Postglacial Recolonization in Europe. *The American Journal of Human Genetics*, 69 (4), 844-852.

Valente, M.J., Dias, R.S., Pereira-Goncalves, R., Catarino, D. and Carvalho, A.F. (accepted for publication) Casais da Mureta (Alcanena, Portugal), a cave necropolis from the Middle/ Late Neolithic: first data on its faunal remains. *Iberian Zooarchaeology Meeting*.

van Oven, M. and Kayser, M. (2009) Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Human Mutation*, 30 (2), E386-E394.

Wang, C., Zöllner, S. and Rosenberg, N.A. (2012) A Quantitative Comparison of the Similarity between Genes and Geography in Worldwide Human Populations. *PLOS Genetics*, 8 (8), e1002886.

Willerslev, E. and Cooper, A. (2005) Ancient DNA. *Proceedings of the Royal Society B: Biological Sciences*, 272 (1558), 3-16.

Woodward, Weyand, N.J. and Bunnell, M. (1994) DNA sequence from Cretaceous period bone fragments. *Science*, 266 (5188), 1229-1232.

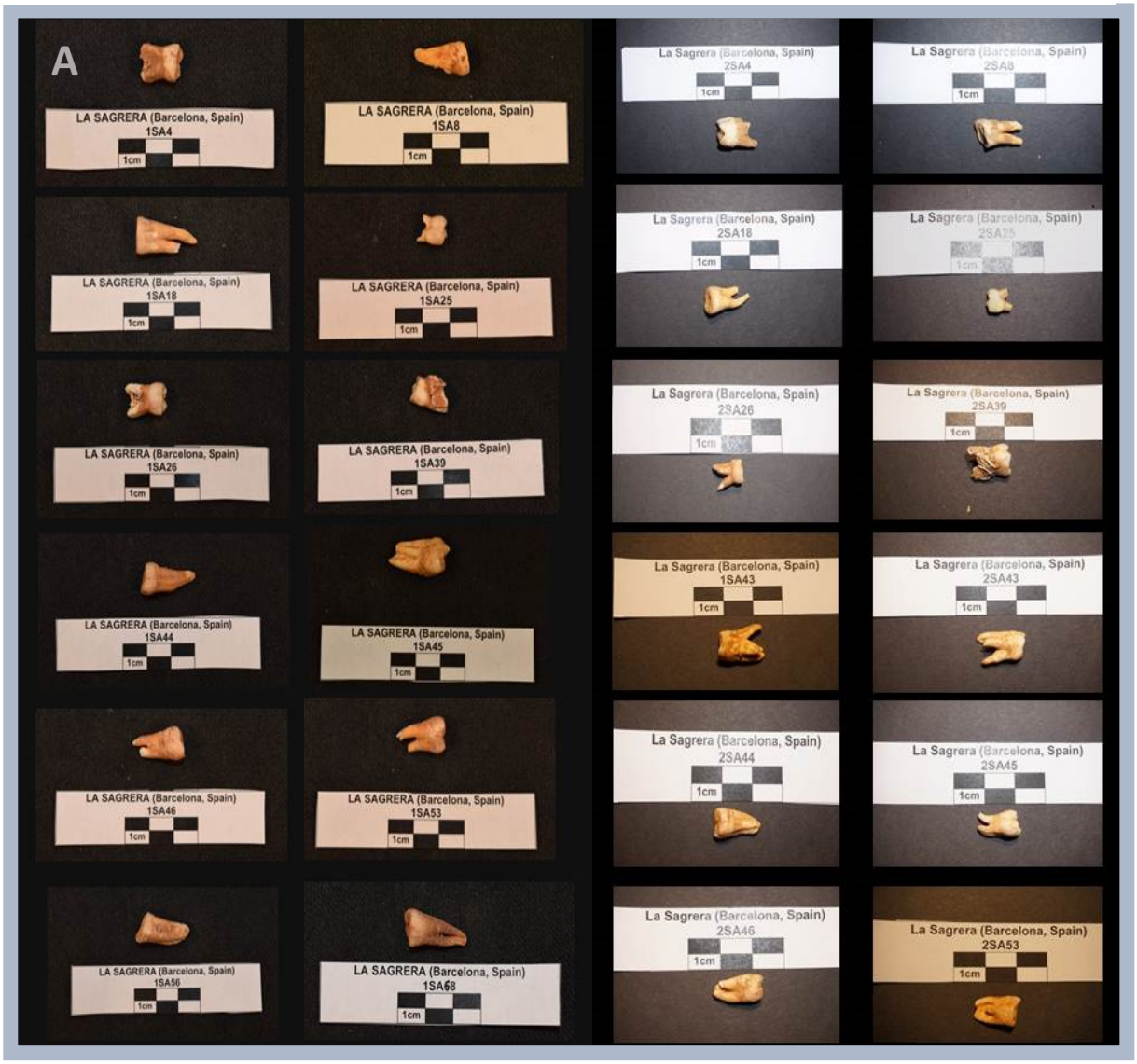
Zilhao, J. (1998) On Logical and Empirical Aspects of the Mesolithic-Neolithic Transition in the Iberian Peninsula. *Current Anthropology*, 39 (5), 690-698.

Zilhao, J. (2001) Radiocarbon evidence for maritime pioneer colonization at the origins of farming in west Mediterranean Europe. *Proceedings of the National Academy of Sciences*, 98 (24), 14180-14185.

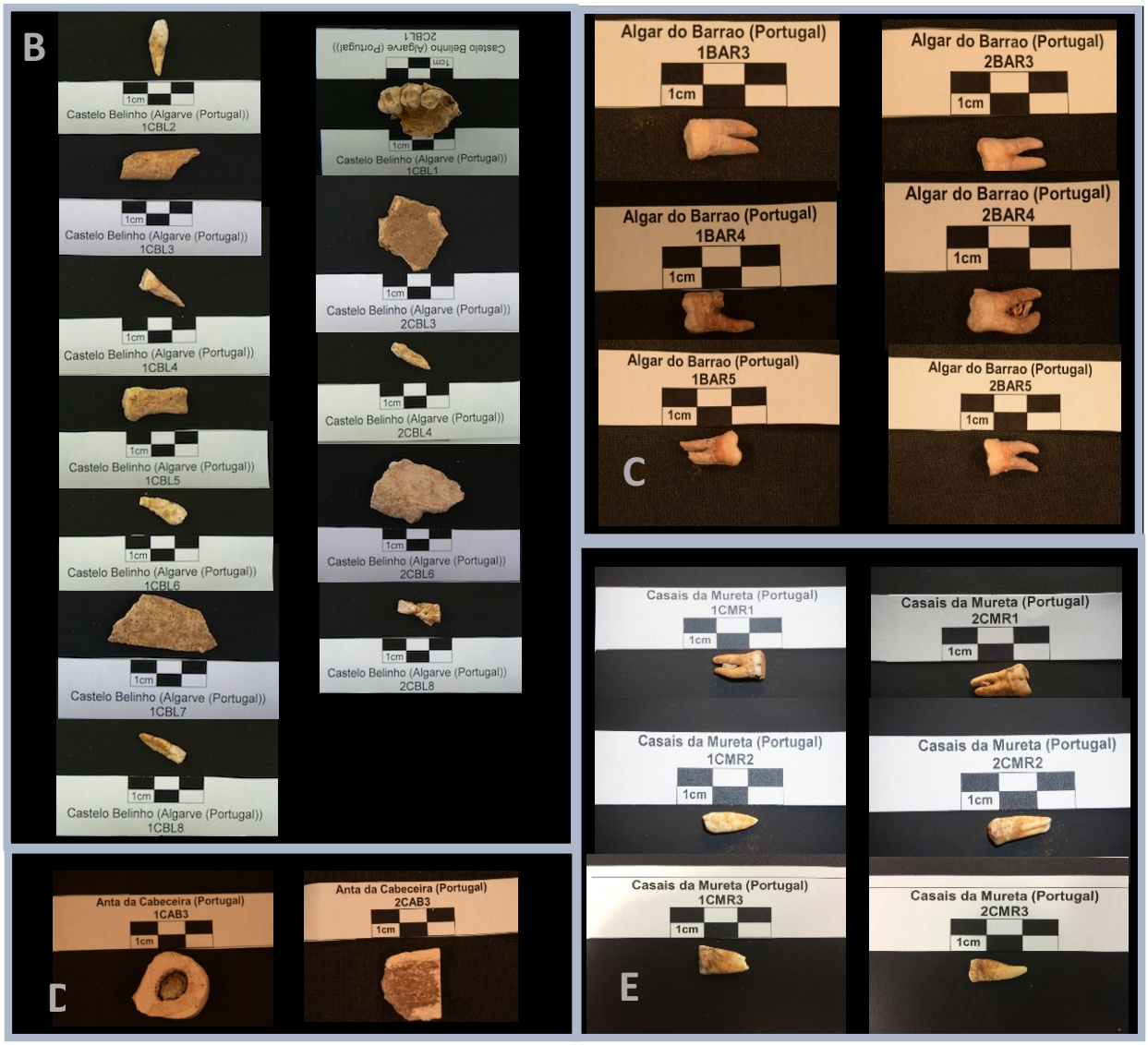
Zvelebil, M. (1986) *Hunters in Transition: Mesolithic Societies of Temperate Eurasia and Their Transition to Farming*. Cambridge: Cambridge University Press.

Appendix

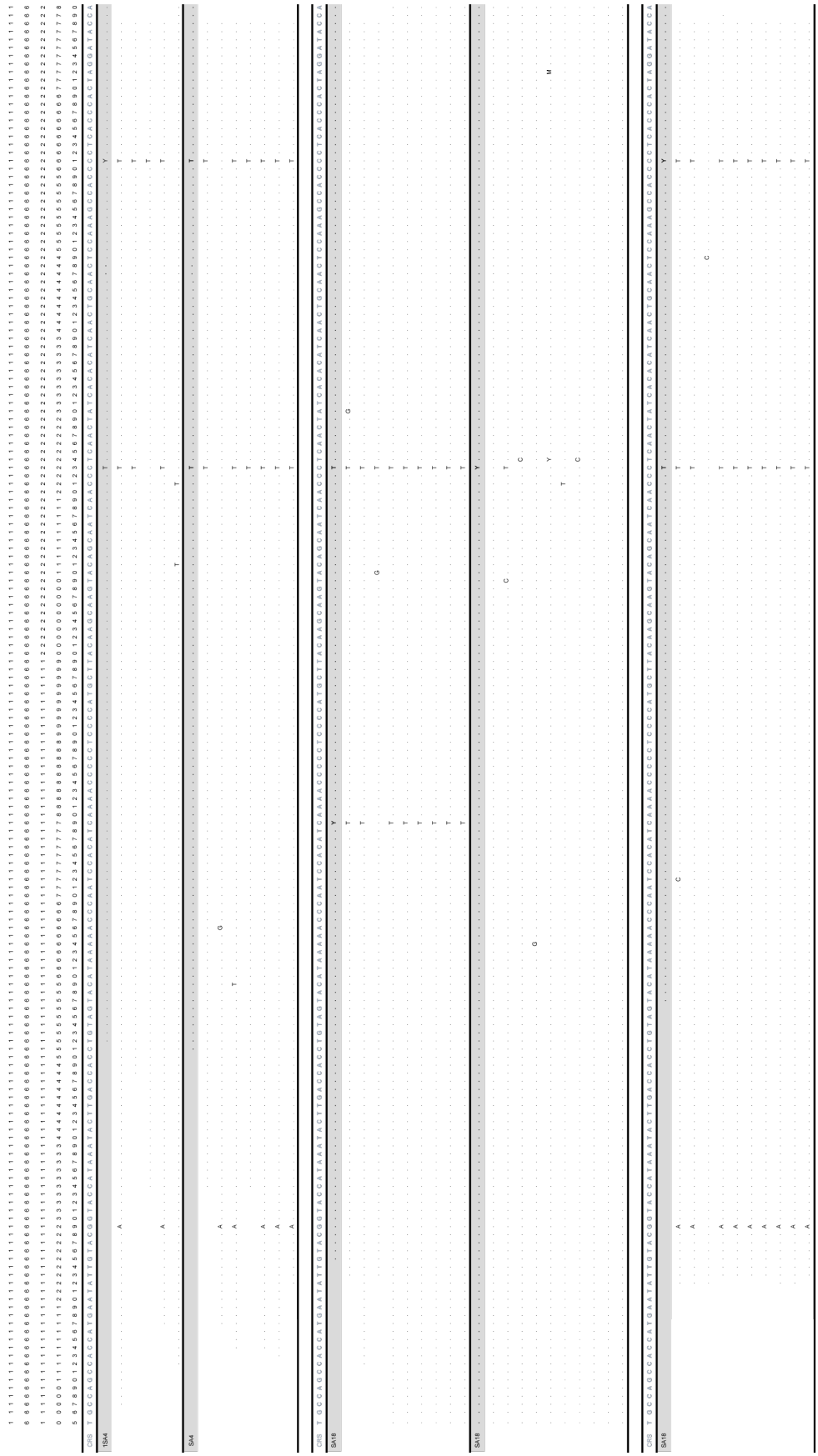
Appendix 1. Samples selected for ancient DNA analysis. A) La Sagrera, B) Castelo Belhino, C) Algar do Barrao, D) Anta da Cabeceira, E) Casais da Mureta



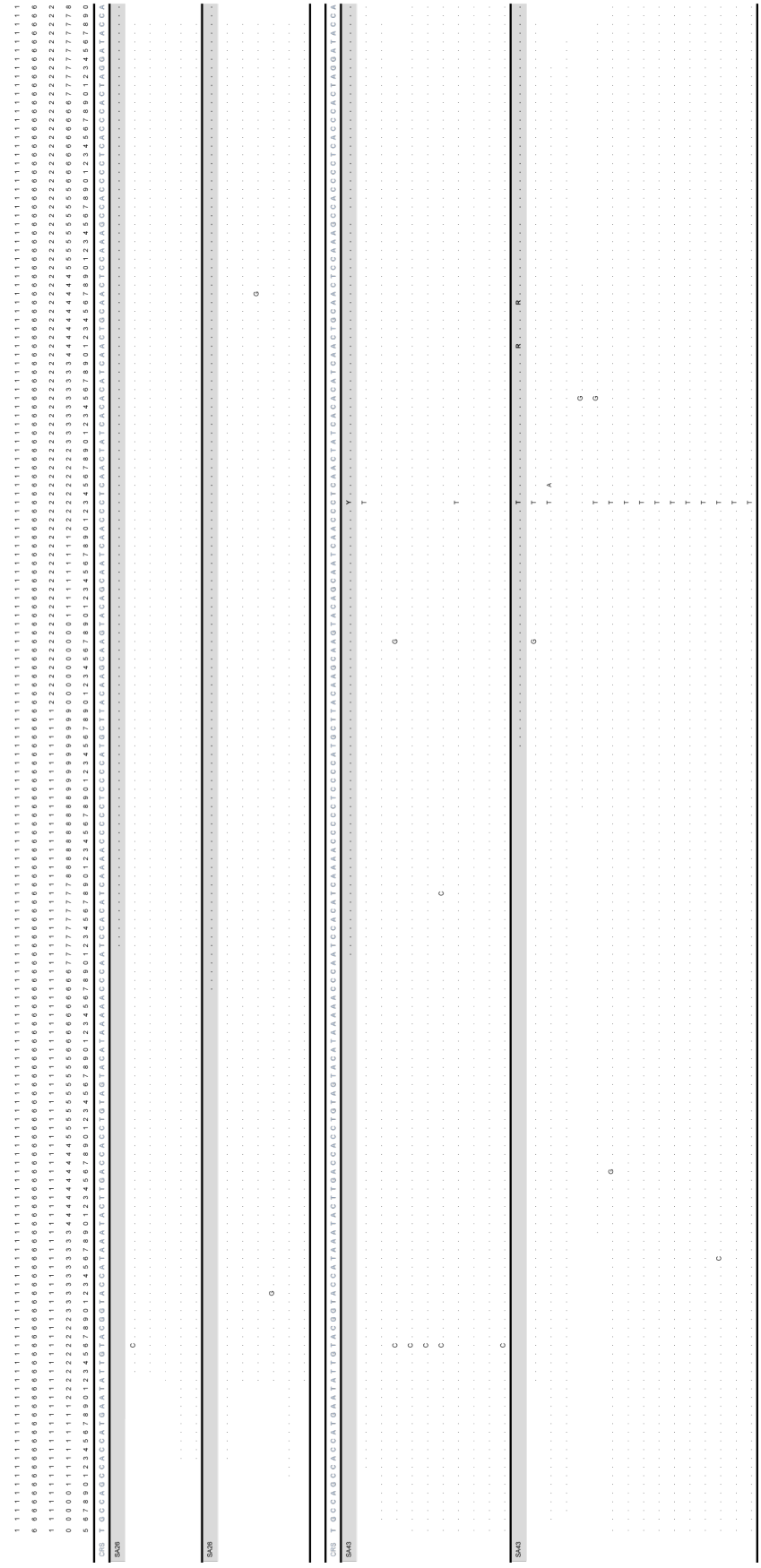
Appendix 1. Continued



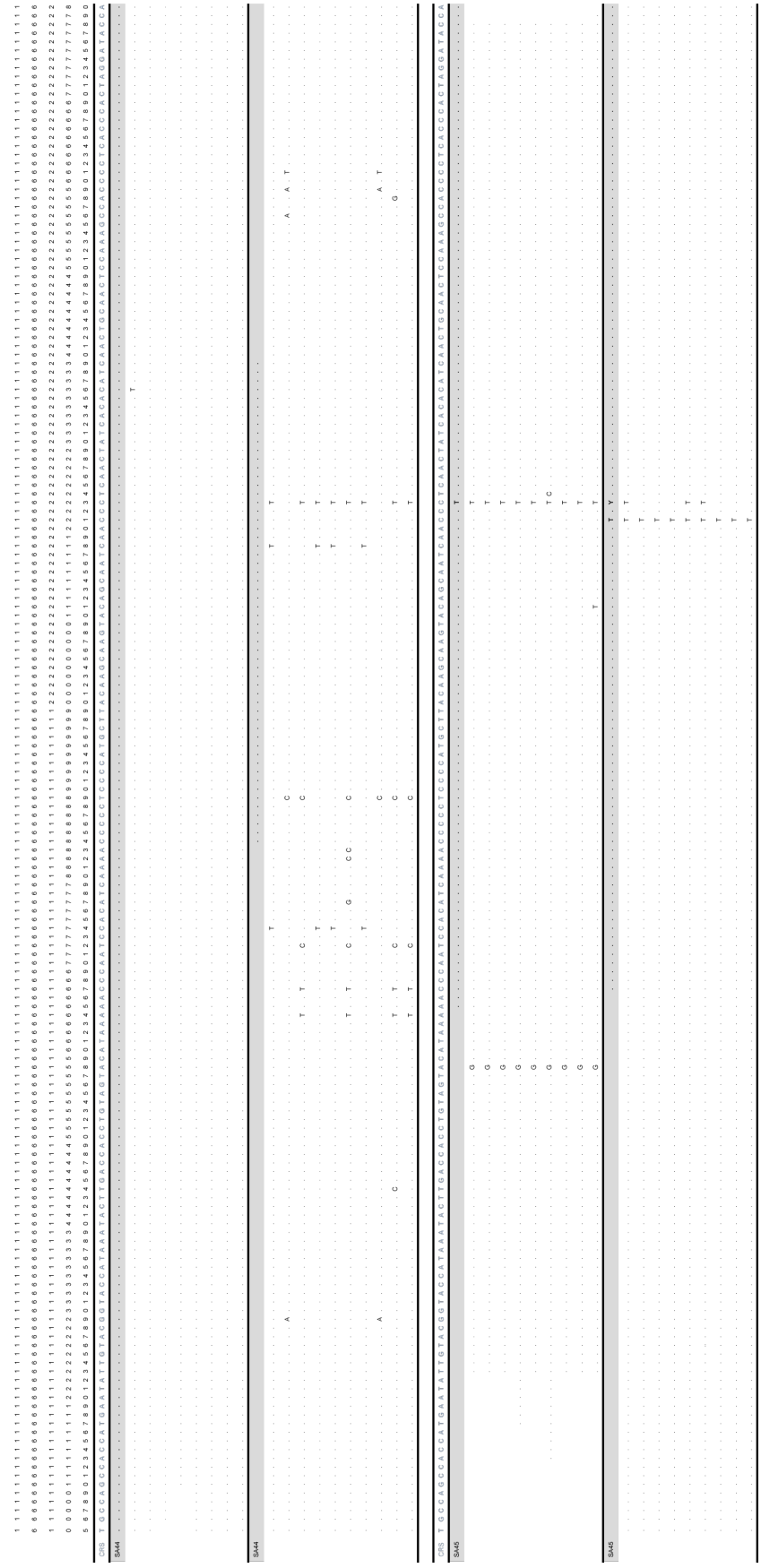
Appendix 2. Clone alignments for mitochondrial hypervariable region I base positions 16105-16280. Grey rows represent direct PCR sequences



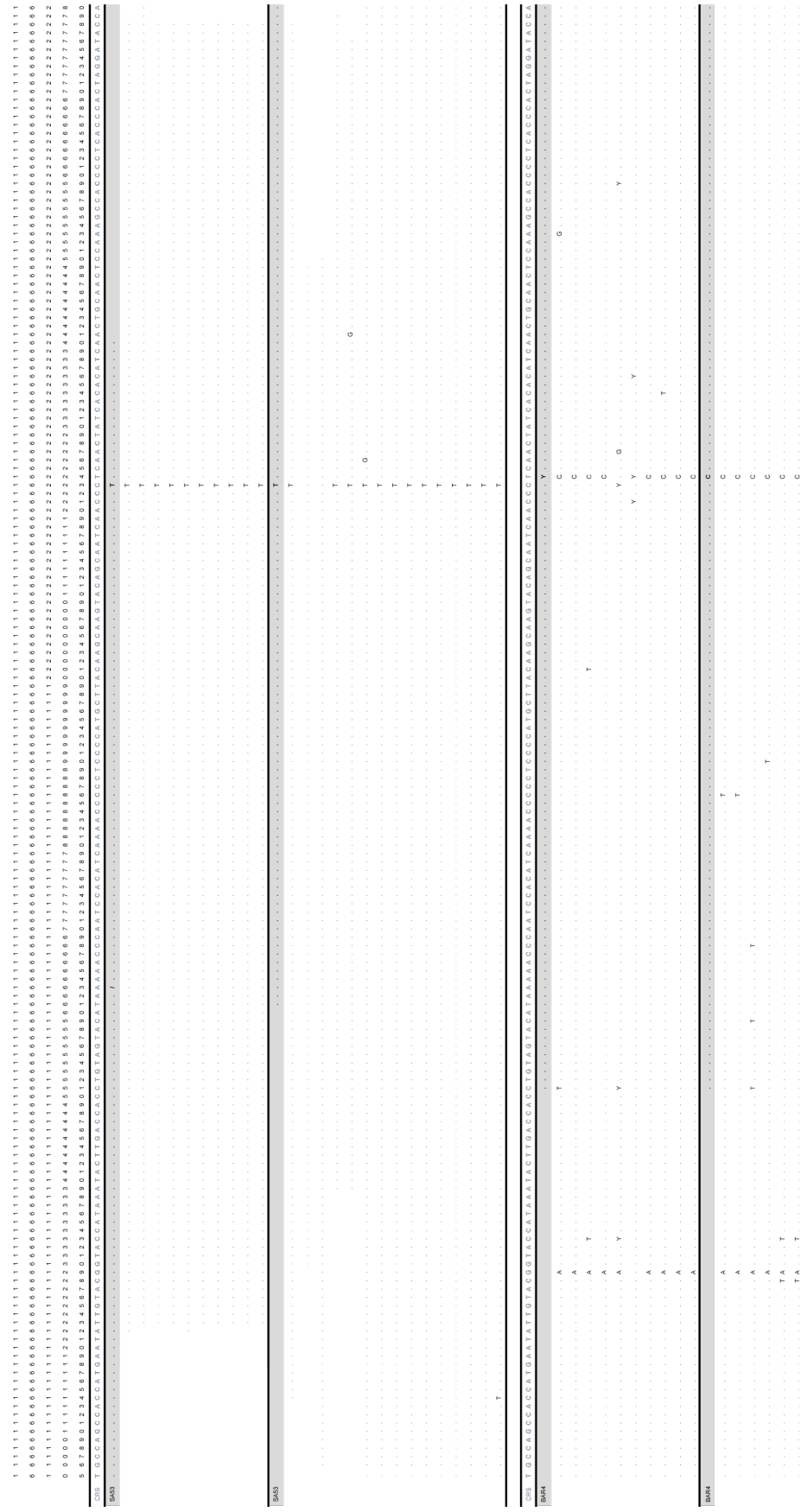
Appendix 2. Continued



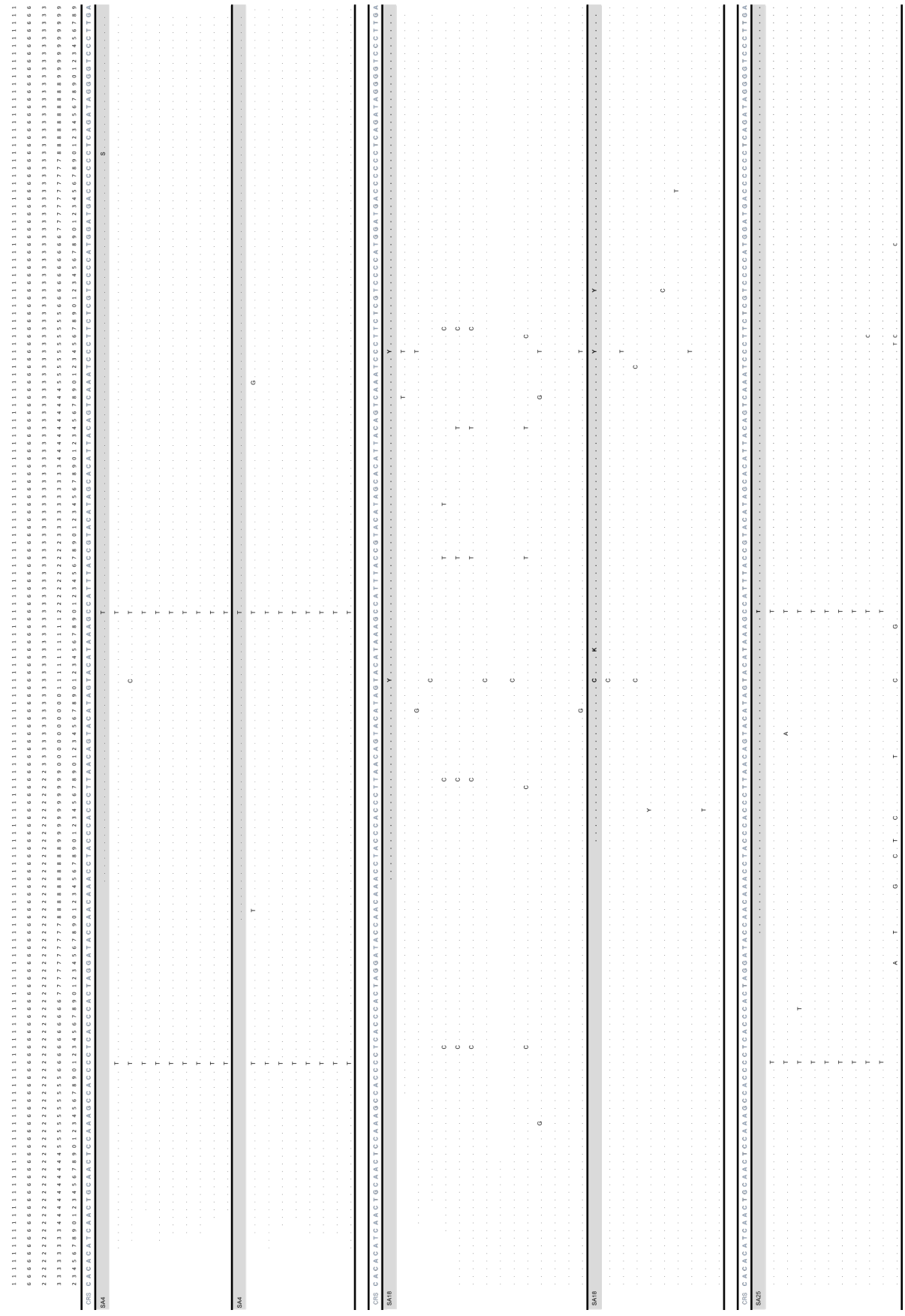
Appendix 2. Continued



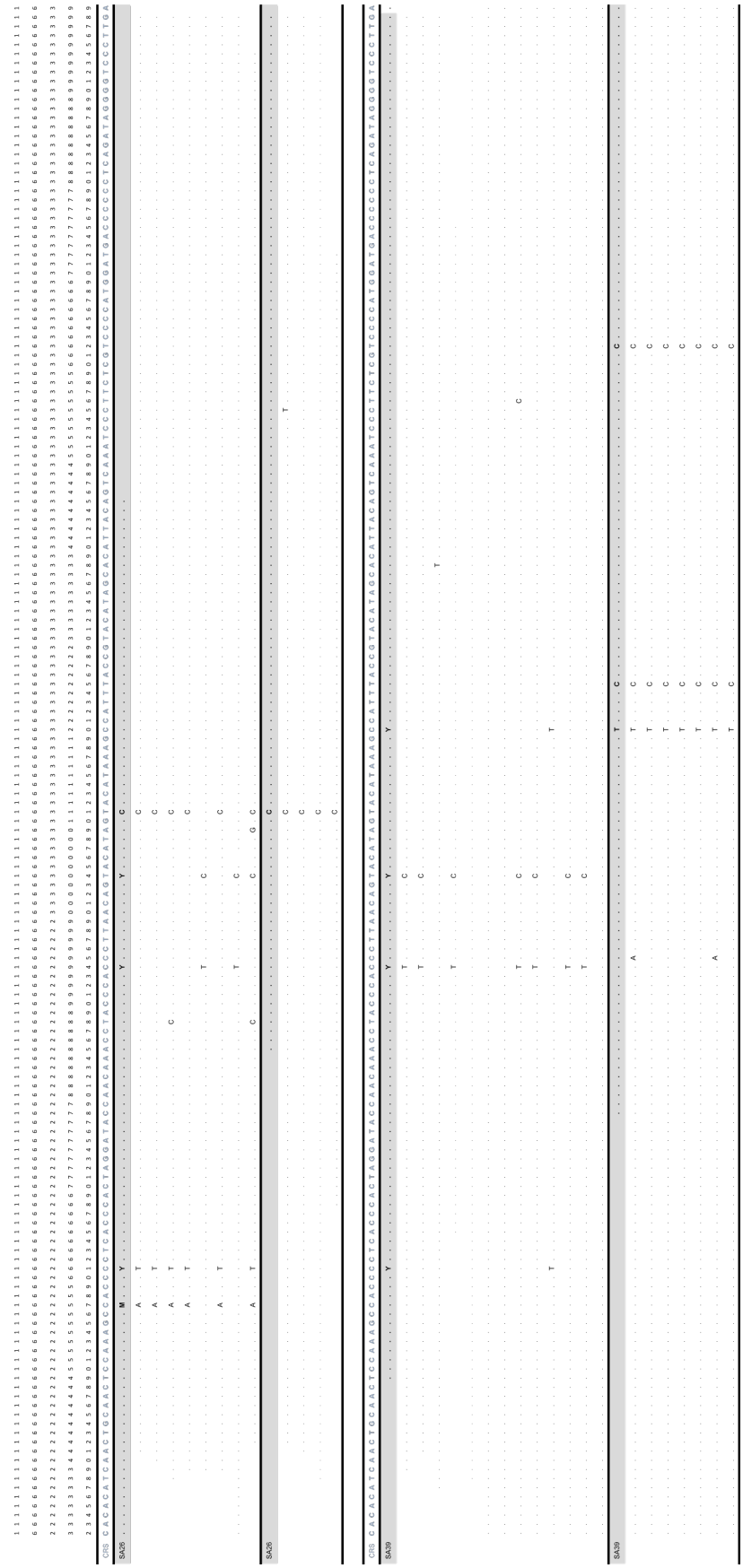
Appendix 2. Continued



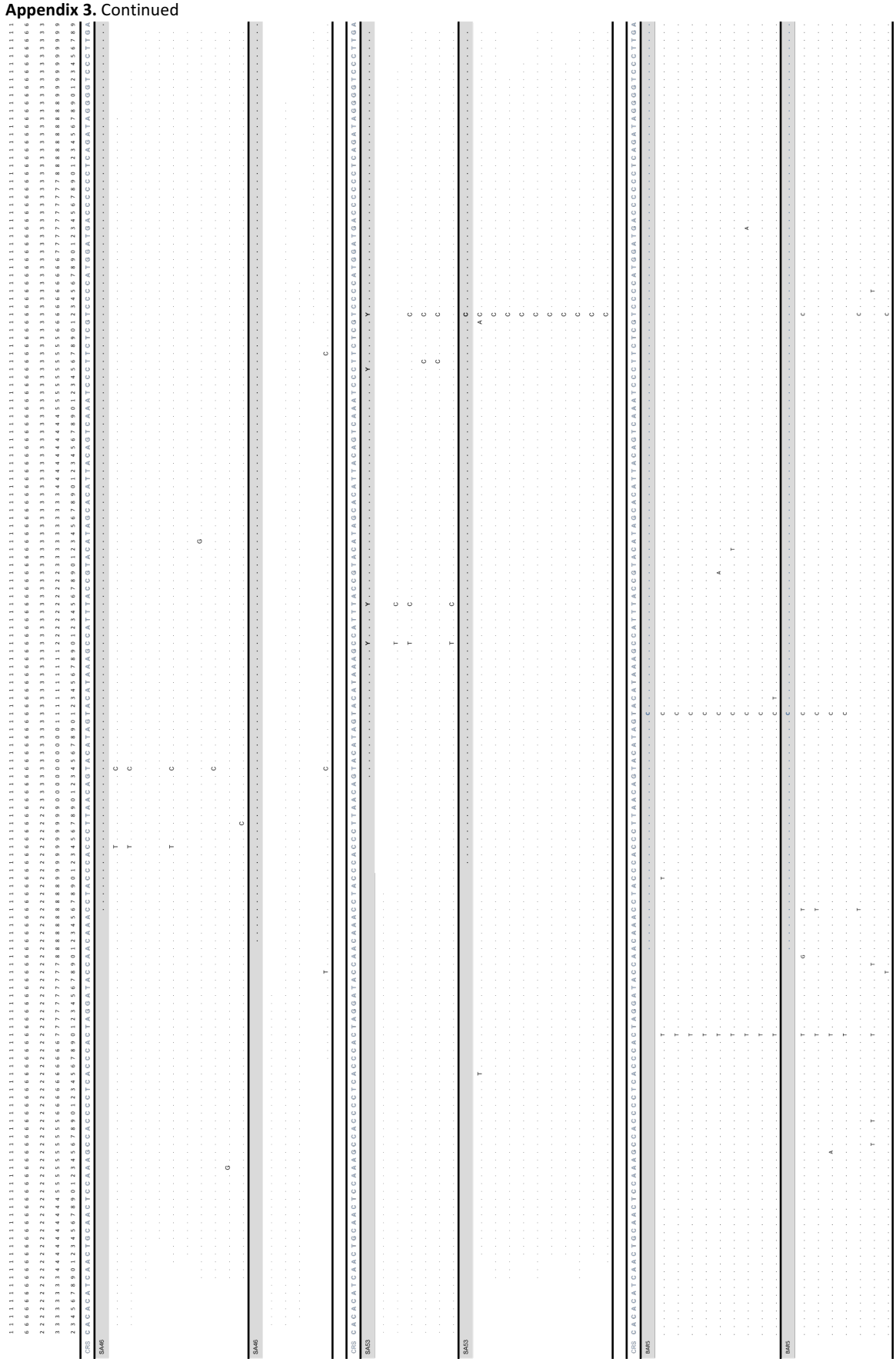
Appendix 3. Clone alignments for mitochondrial hypervariable region I base positions 16232 – 16399. Grey rows represent direct PCR sequences



Appendix 3. Continued



Appendix 3. Continued



Appendix 3 Continued