SOLUTIONS OR ILLUSIONS?
AN ANALYSIS OF THE PALAEOGENETIC EVIDENCE
AT THE ORIGINS OF THE NEOLITHIC IN THE IBERIAN PENINSULA

Íñigo García-Martínez de Lagrán
Arcadia-General Foundation of the University of Valladolid (Spain). igmtzl@gmail.com

Eva Fernández-Domínguez
Department of Archaeology, Durham University (United Kingdom).

Manuel A. Rojo-Guerra
Department of Prehistory, University of Valladolid (Spain).

1) INTRODUCTION

If there is something that can be defined as the “Sisyphus’ rock of human sciences that would be the wide array of circumstances, interpretations and processes generated by human beings and their numberless organizations, with regard to groups, communities, collectives, languages… “Cultures”. Thus, when (pre)historians have the chance to resort to hard sciences, they cling tightly to it, sometimes even with a blind trust, in the hope of finding a reference point for their intrinsically human –and therefore multidimensional and multifocal– interpretations, open to different possible solutions.

In recent years the study of the past has placed its trust in dating techniques, –particularly in 14C–, in statistical analyses and in genetic studies, among other such possibilities, in a way that they could be considered as the new “guide fossils” in prehistorical research. For instance, the chronology of a domestic animal or plant would be taken as an undeniable marker of the “Neolithic” period, and something similar could be said of the hypothetical presence of a Near Eastern haplogroup in a VI millennium BC European individual. However, from a (pre)historic viewpoint, things are not so clear or so “hard”.

The present paper focuses on the analysis of the palaeogenetic data available for the Neolithisation process and the Early Neolithic period in the Iberian Peninsula. We have considered the information from a historical and anthropological perspective in an attempt to understand and complement the interpretations derived from the geneticists themselves. Moreover, a part of the paper is devoted to explain the archaeological consequences of the types of analysis carried out and the results of such studies. The final goal would be to broaden the possible interpretations of the Neolithisation process and the Early Neolithic period in this territory.

2) PALAEOGENETIC EVIDENCE OF THE MESOLITHIC-NEOLITHIC TRANSITION IN EUROPE

Understanding the jargon commonly used in genetics can be challenging for an unfamiliar audience, this being probably the main cause of misinterpretation of genetic data by non-experts. With this in mind, the main concepts used in population genetic studies have been compiled in supplementary information and explained in a simple way, starting from the very basics (Supplementary Information 1).

During the last few years we have witnessed an exponential increase in the quantity (and quality) of prehistoric human paleogenetic data. Two main factors have contributed to this fact, namely the development of a new genotyping technique -the so-called Next Generation Sequencing or NGS- and a new sampling strategy targeting the densest of the bones -the petrous...
(pars petrosa of the temporal bone)-, which proved to contain up to 183 times more DNA than any other skeletal elements (Gamba et al., 2014). These advances have not passed unnoted for the archaeological community and our knowledge of different cultures and periods has greatly increased, including the old question of the nature of the spread of the Neolithic Europe. In this regard, genetics has come to challenge us in different ways and while these scientific techniques are overall very welcome, their interpretations have left some scholars with discomfort (Hofmann, 2014; Linden, 2016).

Inspired by Thomas Lindahl’s article “Facts and Artifacts in ancient DNA” (Lindahl, 1997), we will focus in this section on the “facts” aka main conclusions of ancient DNA studies up to date on the European Mesolithic-Neolithic transition to further discuss some of their “artifacts”, understood in this case as partial interpretations needing of a wider archaeological discussion, with a particular focus on Iberia.

2.1. The pre-Neolithic genetic background

Early mitochondrial DNA studies already suggested that the genetic background of pre-Neolithic populations might have been dramatically different from that found in Early Neolithic populations (Bramanti et al., 2009), a suspicion that became a fact once that more and more data of both groups has been made available (Gamba et al., 2014; Hervella et al., 2015; Posth et al., 2016; Sánchez-Quinto et al., 2012). These were indeed good news, as this finding would potentially allow for phenomena of acculturation, admixture, assimilation or introgression to be identified, something that would have been unachievable if both groups had been genetically homogeneous or even similar.

These studies painted a picture of a uniform hunter-gatherer population across Europe, mainly dominated by U-type haplogroups: U, U2, U4, U5 and U8 (Bramanti et al., 2009; Gamba et al., 2014; Lazaridis et al., 2014; Sánchez-Quinto et al., 2012). More recent contributions have proposed a more complex picture, with a higher mtDNA diversity in pre-Late Glacial Maximum samples, including representatives of U haplogroups but also members of the haplogroup R in Northern Italy and, surprisingly, of the macro-haplogroup M in France and Belgium, a clade that is nowadays restricted to Asia, Australasia and America (Posth et al., 2016). This post-LGM reduction in diversity was interpreted by the authors as evidence of 1) a genetic bottleneck during the LGM resulting from extreme climatic conditions, followed by a re-expansion after the ice sheet retracted and 2) a genetic turn-over of the post-LGM population by a distinct population of an unknown origin, maybe from a different LGM refugia, dominated by haplogroup U5 (Posth et al., 2016). These last would be the immediate ancestors of the populations that would later interact with the Neolithic farmers. Even though this interpretation is based on a small subset of complete mitochondrial sequences (N=55) from different locations, cultures and ages, it raises concerns about the use of any hunter-gather sequence regardless of age as proxies of Mesolithic populations (Bramanti et al. 2009).

Genome-wide DNA studies contradict this apparent uniformity, showing a clear distinction between Western, Eastern and Scandinavian hunter-gatherers (Haak et al., 2015; Lazaridis et al., 2016; Olalde et al., 2015), even though it is important to note the differences in sample size, chronology and geographical representativity of the three markers.

Figure 1 illustrates the location of archaeological sites that have provided paleogenetic data from Mesolithic and Early Neolithic periods. As it can be seen in the figure, data are still very scarce and biased towards certain regions. Even though we acknowledge that the only way forward is to work with the available information, its limitations and uncertainties should be clearly stated, especially when attempting direct comparisons between Mesolithic and Early Neolithic specimens and the former comes from different periods and regions. In fact, Olalde et al., 2015 recognize that the Early Neolithic sample of Cova Bonica (Catalonia, Spain) is genetically closer to a Hungarian Mesolithic-type DNA than to the two samples of La Braña-Arintero.

Figure 1. Mesolithic and Neolithic European and Near Eastern archaeological sites from where palaeogenetic data has been obtained. Red: Mesolithic sites, Green: PPNB sites, Light green:
2.2. The Neolithic genetic background and possible interactions with hunter-gatherer populations

The Early Neolithic mitochondrial DNA landscape is overall dominated by haplogroups K, N*, N1a, T2, X2, H and J, most of which can be traced back to Anatolia and the Levant (Figure 2). In the absence of information on the local Mesolithic genetic background, assumptions need to be made based upon similarities with the genetic stock of the first farming Near Eastern populations, which has only become available very recently (Fernández et al., 2014; Kilinc et al., 2016; Lazaridis et al., 2016; Mathieson et al., 2015). Even though this dataset is still scarce and does not account for the geographical and chronological complexity of the first emergence of the Neolithic in the Near East, interesting patterns arise: 1) the farmers from the Levant, Anatolia and Zagros represent three different genetic stocks (albeit with certain common mtDNA haplogroups) clearly distinguished at genome-wide and Y chromosome levels and 2) Early Neolithic populations across Europe studied up to date show genome-wide striking similarities to the Anatolian Neolithic (Boncuklu Höyük, Menteşe and Barcı), but not to the populations of the Southern Levant (Ain Ghazal, Motza) and the Zagros region in Iran (Ganj Dareh). Even though a Southern Levantine ancestry of Early Neolithic European populations has been excluded considering the information above, mitochondrial DNA analyses of samples of the Northern Levant have shown intriguing similarities with both LBK and Cardial/Epicardial groups, proposing an early dispersal of farmers through Cyprus and the Aegean sea (Fernández et al., 2014). Southern and Northern Levantine populations share haplogroups K and R0, and both also lack N1a and X2, haplogroups characteristic of Anatolian populations.

Figure 2. Mitochondrial haplogroup frequencies of Early Neolithic populations of Europe and the Near East. 1: Pre-pottery Neolithic B/C from Jordan and Israel (Lazaridis et al., 2016); 2: Pre-pottery Neolithic B from Syria (Fernández et al., 2014); 3: Anatolian Neolithic (Mathieson et al. 2015); 4: Starčevo-Cris-Körös (Gamba et al. 2014; Hervella et al. 2015; Szécsényi-Nagy et al., 2015); 5: Transdanubian LBK (Gamba et al., 2014; Szécsényi-Nagy et al., 2015); 6: German LBK (Brandt et al., 2013; Brotherton et al., 2013; Haak et al., 2015, 2010); 7: Cardial/Epicardial Aragón (Gamba et al., 2012; Haak et al., 2015); 8: Epicardial Cantabria (Hervella et al., 2012); 9: Cardial/Epicardial Catalonia and Valencia (Gamba et al., 2012; Lacan et al., 2011; Olalde et al., 2015); 10: TRB Scandinavia (Mälmstöm et al., 2009; Skoglund et al., 2012, 2014).

Moving towards Europe, the most Southwestern point with Early Neolithic DNA data is the site of Revenia, in Northern Greece (cal. 6,438–6,264 BCE). Mitochondrial data is also available from two individuals of the Mesolithic site of Theopetra in Thessaly (7,605–7,529 cal BCE and 7,288–6,771 respectively) (Hofmanová et al., 2016). The two Neolithic individuals harbour mtDNA haplogroup X2 and Y Chromosome haplogroup G2a, both present in Anatolian Pottery Neolithic sites (Barcı and Menteşe), while the two hunter-gatherers from Theopetra have been identified as members of mitochondrial haplogroup K1. No genomic or Y chromosome data could be retrieved from these two hunter-gatherer individuals. While an Anatolian origin has been claimed for the Early Neolithic Revenia sample based upon their similarities at genome-wide level, it is surprising that the two studied Mesolithic individuals carry typically Neolithic mytotypes. As pointed out in (Fernández-Domínguez and Reynolds, 2017), in the absence of Mesolithic genomic data from the region both a common-similar Mesolithic genetic background for Anatolia and Greece and the Greek samples being direct descendants of Neolithic Anatolians as a result of migration are suitable explanations.

From there on, the bulk of data corresponds to the Danubian route of Neolithic spread, with representatives of the Körös, Starčevo, Čris, Transdanubian (Hungarian) LBK and German
LBK (Brandt et al., 2013; Brotherton et al., 2013; Gamba et al., 2014; Haak et al., 2015, 2010; Hervella et al., 2015; Szécsényi-Nagy et al., 2015). The mitochondrial DNA composition of the Anatolian datasets mirrors the Starčevo-Čris-Körös, Transdanubian and German LBK data, displaying a set of common haplogroups: K, N1a, T2, X2, H and J (Figure 2). These populations also harbour a high frequency -usually over 50%- of the Y chromosome haplogroup G2a. Similarly, the genomic DNA for the Körös and Starčevo samples tightly clusters with the Anatolian Neolithic (Haak et al., 2015; Mathieson et al., 2015). The German LBK samples, however, show a higher level of diversity at genome-wide level. As it would be discussed later, some of these display striking similarities with some Cardial/Epicardial Iberian samples (see Figure 2 of (Haak et al., 2015; Hofmanová et al., 2015; Olalde et al., 2015). This pattern overall, has been interpreted as an evidence of direct colonisation of Early Neolithic farmers from Anatolia, but a couple of detailed statements should be made on that regard 1) data from local Mesolithic populations is restricted to one individual in Hungary and 14 individuals in Germany, and none of them provided genome-wide data 2) in the case of Germany, late hunter-gatherers and Mesolithic populations from other regions have been used as proxies for the last hunter-gatherers of the region and 3) there is a direct genetic evidence of acculturation in the Danubian route, as the genome of individual KO2 from the Neolithic Körös culture in Hungary showed a characteristic Mesolithic genomic make up. Other lines of evidence, like the presence of basal levels of Y chromosome haplogroup I and its subclade I2a1, both present in Mesolithic hunter-gatherers from Scandinavia (Motala, Sweden) and Luxemburg (Loschbour), might also suggest that acculturation or admixture events may have been more prevalent in this area than the current data shows.

In contrast with the evidence presented above, the reconstruction of the overall dynamics of the Neolithic route of expansion through the Mediterranean, linked to the Impressa/Cardial cultures is much more complex. As it can be seen in Figure 1, there is virtually no genetic information available in Mediterranean Europe, from Greece to the southwestern coasts of France. Our knowledge is therefore restricted to the analyses performed in Iberia, all of them focusing exclusively on mitochondrial DNA with the exception of one individual from Cova Bonica, from which genome-wide data could be obtained (Olalde et al., 2015).

The available Early Neolithic genetic data within the Iberian Peninsula is also geographically disperse: two mitogenomes from Portugal (Galeria da Cisterna), two low coverage mitogenomes from Cova de l’Or and Cova de La Sarsa (Valencia) and the main bulk of data concentrated in the Northeast: Catalonia (Cova de l’Avellaner, Cova Bonica, Can Sadurní and Sant Pau del Camp), Huesca (Chaves and Cueva de Els Trocs) and Navarre (Los Cascajos and Paternanbidea) (Gamba et al., 2012; Haak et al., 2015; Hervella et al., 2012; Lacan et al., 2011; Olalde et al., 2015) (see Table 1 and Figure 3).

Table 1. Available paleogenetic data of Paleolithic, Mesolithic and Early Neolithic samples in the Iberian Peninsula. Only mitochondrial haplotypes and haplogroups are included. Chronological information in Table 4 and geographical location in Figure 7.

Figure 3. Map of Iberian sites with Palaeolithic, Mesolithic and Early Neolithic data mentioned in the text. Squares - Magdalenian: 1- La Pasiega (Puente Viesgo, Cantabria); 2- La Chora (San Pantaleón de Aras, Cantabria); 3- Erralla (Cestona, Guipúzcoa). Triangles - Mesolithic: 4- La Braña-Arintero (Valdelugueros, León); 5- Aizpea (Abaurrea Alta, Navarre). Circles - Early Neolithic: 6- Los Cascajos (Los Arcos, Navarre); 7- Paternanbidea (Ibero, Navarre); 8- Chaves (Bastarás, Huesca); 9- Els Trocs (Bisaurri, Huesca); 10- L’Avellaner (La Garrotxa, Gerona); 11- Can Sadurní (Begues, Barcelona); 12- Cova Bonica (Valliran, Barcelona); 13- San Pau del Camp (Barcelona); 14- Cisterna, Almonda (Santarem); 15- Sarsa (Bocairent, Valencia); 16- Or (Beniarrés, Alicante).
The available Mesolithic palaeogenetic data is also very poor, restricted to two individuals from La Braña-Arintero (León) and one individual from Aizpea (Navarre).

While the three Mesolithic individuals display typical hunter-gatherer U lineages, the transition to the Neolithic is characterised by the appearance of new mitochondrial haplogroups: HV, I, J, U*, K, X in Navarre and H, K, J, T2, X in Catalonia, Aragón and maybe also Valencia. While haplogroups X, J, T2, K, X, N1 and N* have parallels in the Near East, the situation is less clear for types HV, I and U*, and a possible pre-Neolithic ancestry has been proposed to explain their presence in the Cantabrian fringe. To sum up, the different datasets have provided different regional interpretations, from a pioneer farming colonisation in Catalonia to a random model of dispersal of Early farmers in Navarre.

On the other hand, the genome-wide data obtained from sample CB13 from Cova Bonica has provided a more detailed insight into the genetic pool of the first Cardial communities of Iberia. This sample shows some similarities to German LBK and Early Anatolian Neolithic, which has been interpreted as an evidence that both LBK and Cardial/Impressa can trace their ancestors to the same source population.

With all these data in hand, we consider that a reinterpretation of the palaeogenetic evidence available for the Iberian Peninsula, accounting for a more detailed archaeological discussion, is not just possible but needed, and this would be the objective of the following section.

3) DISCUSSION: SOLUTIONS OR ILLUSIONS?

In our opinion, the following aspects should be taken into account in the interpretation of the Palaeogenetic studies concerning the Neolithisation process in Europe in general and the Iberian Peninsula in particular:

3.1. The chronology of the DNA samples and the interaction phenomena during the Neolithisation Process

Research in recent years has qualified both the dating and the leading characters of the Neolithisation Process in the Iberian Peninsula. On one hand, a Mediterranean maritime colonisation by pioneer, fully Neolithic groups characterised by pottery with a clear parallel to Ligurian Impressa collections of Italian origin has been discovered (Bernabeu-Aubán, Molina-Balaguer, Esquembre-Bebiá, Ortega, & Boronat-Soler, 2009; García-Borja, Aura-Tortosa, Bernabeu-Aubán, & Jordá-Pardo, 2010; Guilaine & Manen, 2007; Guilaine, Manen, & Vigne, 2007). The archaeological record, from Italy to France and Spain, shows a clear maritime spread of these groups, which started arriving to Iberia from ca. 5778-5541 cal BC1 (García-Martínez de Lagrán, 2017). Some of these sites can be detected in inner areas, revealing other possible ways of Neolithic dispersal through the Pyrenees or/and terrestrial ways following the main rivers. One example can be the expansions through the Ebro Valley (Fernández-Eraso et al., 2015; García-Martínez de Lagrán, 2015; Utrilla-Miranda, 2002). At the same time, North African influences and contacts have been identified in the Neolithisation of meridional Iberian areas (Linstädter et al., 2012; Manen et al., 2007; Morales et al., 2013). All of these situations define the Iberian Peninsula as a Neolithisation melting pot of influences and groups, as the geographical and historical end of a continental process (García-Martínez de Lagrán, 2015).

In most of Iberia the Early Neolithic begins between 5405-5334 (García-Martínez de Lagrán, 2017). In this moment, sensu lato, the traditional Franco-Iberian Cardial and

---

1 The chronological references in the text are cal BC.
Impressed/Incised-Channelled (I/I-C) (Epicardial) groups appear. Recent studies show the same chronology and mutual influences in pottery style of these “groups” or “phenomena” (Bernabeu-Aubán & Martí-Oliver, 2014; Bernabeu-Aubán, Rojo-Guerra, & Molina-Balaguer, 2011; Rojo-Guerra et al., 2012). These data, and others, point out a more complex cultural, ethnographical mosaic in Iberian Early Neolithic.

However, other scholars propose a different scenario: the initial Neolithic expansion across the Western Mediterranean would have been associated with Cardial Ware but not with other collections of Italian or African origins. This expansion would have taken place between 5600 and 5400, from Languedoc to the western Portuguese coasts (Martins et al., 2015, pp. 126-127; Zilhão 2011, 2001).

These questions regarding the chronology, the type of colonization and the length of the Neolithisation process play a significant role in the interpretation of the current palaeogenetic evidence, particularly with regards to those samples dated back to the Early Neolithic period. As Table 2 and Figure 4 show, we have really few dates of the Neolithisation process. For Mesolithic contexts we have just the individual Aizpea 1(5647-5468 2σ) and with very limited percentage La Braña 1 (5983-5747 2σ). For Neolithic contexts there are just two dates of individuals with DNA analysis: Cova Bonica 13 (5469-5327 2σ) and Or H3C6 (5464-5298 2σ), and the context c18 of Can Sadurní (5486-5321 2σ) and Chaves Ib (5471-5304 2σ).

Other dates could be added to this list like individuals Cisterna G21 (5351-5223 2σ) and Sarsa 4675 (5360-5217 2σ), and context like San Pau del Camp F1 (5372-5076 2σ), but their chronological lapses beyond 5334 are limited.

To put it clearly, as of today we have very scarce, effectively almost nule, palaeogenetic analyses of individuals directly involved in the Neolithisation process. In consequence, the majority of the ancient DNA evidence available for the Iberian Peninsula either belongs to hunter-gatherer communities or to Early, Middle and Late Neolithic groups (Alt et al., 2016; Carvalho et al., 2016; Gamba et al., 2012; Gómez-Sánchez et al., 2014a, 2014b; Günther et al., 2015; Haak et al., 2015; Hervella et al., 2012; Lacan et al., 2011; Olalde et al., 2015; Sánchez-Quinto et al., 2012). Therefore we have no direct information regarding the genetic background of the earliest, maybe exogenous Neolithic peoples arriving at the Iberian Peninsula. The available evidence belongs to individuals resulting from the Neolithisation process, a period of social, cultural and probably also of genetic exchange and interaction among different human groups.

**Figure 4.** Compared dates of analysed individuals and contexts and dates from the Early Neolithic period in the Iberian Peninsula. In blue/purple the Neolithisation period.

**Table 2.** 14C dates of sites of the Table 3.

With all this in mind, a key concern would be to try to understand the full magnitude and the consequences of this period of interaction and exchange. The difficulty of the analysis lies in the fact that there are two different, merging processes taking place at the same time: a historical and socio-economic process among communities, and a mixing process among individuals. This two-fold variability factor multiplies almost ad infinitum the different possible results. At the same time, it calls for a more holistic approach to the Neolithisation process. For example, the presence of Near-Eastern haplotypes in an individual (a part of the whole) does not
necessarily mean that the community (another part of the whole) to which he or she belongs can be considered a Neolithic group, or that the Neolithisation process of this community or region (in this case, the whole) followed a colonization by exogenous fully Neolithic settler groups. It could also be the case of an individual with Near-Eastern haplotypes living in the midst of a hunter-gatherer community.

The main conclusion is that palaeogenetic data from the Early Neolithic period in the Iberian Peninsula could have been determined and modified by a period of massive interaction and exchanges: the Neolithisation process. Being aware of these circumstances should serve us as a warning against the simplistic interpretations we have accepted until now.

3.2. Types and features of palaeogenetic analyses

As we have seen in section 2, genetic analyses can focus on different genetic markers. Their features and object of analysis have a very significant impact on the interpretation of the results.

For haploid markers (mtDNA and Y chromosome), the relationship between the temporal origin and the time of coalescence of the haplogroups should be taken into account for their (pre)historical interpretation. Haplogroup coalescence times, that is, the estimated haplogroup datings, have nothing to do with their assignment to a particular period or culture. For instance, haplogroups J, T and K’s coalescence times go farther back in time than the Neolithic period. However, they are considered characteristic of Neolithic groups because they are present in Early Neolithic populations from the Near East while absent in local European Mesolithic populations, so it is assumed that they must have been introduced into Europe from the Near East together with the spread of the Neolithic. And at the same time, for example, a haplogroup may have a generations-old Mesolithic origin, while the sampled individual is fully Neolithic from a cultural, socio-economic and collective perspective.

Secondly, attention must be paid to the directionality of the genetic contribution, since it differs in each type of analysis.

On the one hand, mitochondrial haplogroups are inherited exclusively from the mother, with no male contribution whatsoever. This fact conditions the interpretation, since in the case of a genetic exchange between a man with a Mesolithic mitochondrial haplogroup and a woman with a Neolithic mitochondrial haplogroup, all the descendants from the maternal line would inherit the Neolithic mitochondrial haplogroup (Figure 5a). This could be taken as a population replacement, as the Mesolithic contribution cannot be detected; while in fact it is a blend. At the same time, taking the genetic information alone, it could be inferred that this is a fully Neolithic community from an economic viewpoint, when that is not necessarily the case. It could be also a situation of individuals with Neolithic haplogroups that lived in the midst of a hunter-gatherer group.

On the contrary, a genetic exchange between a man with a Neolithic mitochondrial haplogroup and a woman with a Mesolithic mitochondrial haplogroup would result in an offspring with a Mesolithic mitochondrial haplogroup. The descendants from such a couple will inherit Mesolithic mitochondrial haplogroups, although they could live in a farming community.

Something similar can be said for Y-chromosome analyses, with the difference that in this case we will be seeing only the result of the introduction of Neolithic or Mesolithic DNA through the paternal side (Figure 5b). Any external input through the maternal line would be undetectable.

As detailed in section 1, another risk of using haploid markers to study the biological interaction between Mesolithic and Neolithic individuals is that several lineages may become
extinct in subsequent generations at random (genetic drift). This is particularly significant in the case of Iberia, where most of the available palaeogenetic evidence does not belong to the initial stage of Neolithisation, but several generations after.

A combined analysis of both maternal and paternal lineages, as the one attempted in (Lacan et al., 2011), for example, would sort out the directionality problem, but not the problem of genetic drift. Consequently, it would be advisable for future analyses to focus on autosomal nuclear DNA, as it reflects the mixed, shared inheritance of both progenitors even after several generations (Figure 5). This approach has generally a lower success rate than the mtDNA and requires the specimen to be exceptionally preserved, thus not being affordable for all type of samples.

Figure 5. Observable genotypes after two generations following the introduction of Mesolithic DNA (in black) into the Neolithic gene pool (in white) through the paternal line, depending on the analysed genetic marker: a) Mitochondrial DNA, b) Y chromosome c) Autosomal DNA. Black areas indicate Mesolithic DNA and white areas Neolithic DNA.

3.3. The interpretation of Palaeogenetic data and archaeological contexts

To the problematics cited above we must add the archaeological interpretation. Depending on the archaeological record available for each context and site, its definition as Neolithic or Mesolithic may vary. Given that diverse and even a priori contradictory situations may emerge from the palaeogenetic data alone, these should be always combined with the archaeological information. Examples of these situations will be provided next using the genetic information obtained from the site of Los Cascajos and the cave of Els Trocs.

As stated by (Hervella et al., 2012), p. 6, referring to the mitochondrial palaeogenetic data of Los Cascajos and Paternanbidea “the sites of Los Cascajos and Paternanbidea, from Early Neolithic, although they showed a substantial Neolithic cultural influence, they seem to show a lower [than southern France or Central Europe] genetic contribution of female migrants from the original areas where Neolithic first developed”. This statement is based upon the high frequency of haplogroups H (45%), U (29.6%) and U5 (12%) and the low frequency of K (8.2%), J (4.08%), HV, I, T and X (±2% combined) (Hervella et al., 2012 pp. 3-5). Therefore, in light of this mtDNA evidence, it could be assumed that most (female) ancestors of the Neolithic settlers in this site were local hunter-gatherers. Such evidence would confirm a Neolithisation process with a minor contribution from foreign population and a substantial local participation, as some authors argue based on the archaeological record (Alday-Ruiz, 2012, 2009, 2006, Alday-Ruiz et al., 2012a, 2012b; Barandiarán-Maestu and Cava-Almuzara, 2000; Montes-Ramírez and Alday-Ruiz, 2012). The main mechanism of this process would have been then the adoption, exchange and circulation of information among local groups.

However, a complete different interpretation is also possible. Resorting to the very same archaeological, chronological and genetic data, it can be proposed that the community of Los Cascajos is an exogenous group who settled in this territory and whose genetic features have been fully Neolithic during several generations. Consequently they did not need to establish any relationship with local hunter-gatherer groups.

Our hypothesis begins in the Southeastern French coast where a fully Neolithic group arrives: Point 1 in Figure 6 and Table 3. They could have Mesolithic mtDNA haplogroups or not. For a certain time, this group interacted with other Mesolithic or/and Neolithic communities (Doted white circles). These interactions configured the mtDNA of the group through exchanges with these communities, including, for example introgression of Mesolithic DNA. As
a result of these interactions, some Mesolithic mtDNA haplogroups may appear in the DNA background of this community which already had Near Eastern haplogroups.

Some generations after, the group settles in a new territory through a leapfrog colonization process (Point 2) and new interactions begin. The genetic background and socioeconomic features of the group are the same than in point 1. The same process will occur in point 3.

In point 4, there is no longer interaction with other groups (closed white circle), but the genetic (Mesolithic and Neolithic mtDNA admixture) and socioeconomic (fully Neolithic) features are, again, the same than in the preceding points.

Finally, the group settles in the region of the Upper Ebro valley (Point 5), thus becoming a foreign, colonist group with a set of mitochondrial haplogroups both of Near-Eastern and hunter-gatherer origin. Yet, even though this population would be undoubtedly interpreted as Neolithic from a chronological, archaeological and historical perspective although they still retain some Mesolithic haplogroups. Moreover, they could have become part of the genetic code of the community several generations back in other geographical areas (Points 1, 2 and/or 3) and are not necessarily derived from the direct contact with local Mesolithic groups from the present-day region of Navarre.

In this hypothesis, we must take into account that the Iberian Peninsula lies in Europe’s western end, and therefore the pioneer groups reaching this territory have first travelled across the continent for several millennia. In that lengthy journey across Europe these Neolithic groups could have mixed with local Mesolithic populations, who would have therefore contributed to their genetic background. For instance, the ancestors of the first Neolithic groups that reached Spain’s Mediterranean coasts probably had, in previous generations, mixed both with other Neolithic communities and hunter-gatherer groups, from Anatolia to the Italian coasts through Greece, or in southern France, the Mediterranean islands, the Balkan region, the Alps, or even in the North African coasts. The same could be said for the Central European route. Regardless of their origin, are the lifestyle and genes of these communities Neolithic or Mesolithic? In other words, we might wonder whether the haplogroups H, U and U5 that such groups (for example in Los Cascajos) introduced into the Iberian Peninsula can or must be interpreted as haplogroups of hunter-gatherers and “local contributions” or just the opposite.

Figure 6. Possible interpretation of the origin of the Neolithic community in Los Cascajos.

Table 3. Hypothesis of Los Cascajos community’s story.

Similarly, individual CSA24 from Can Sadurní, displaying mitochondrial haplogroup U5, was interpreted as evidence of admixture with local hunter-gatherer groups, therefore (Gamba et al., 2012). Applying the same logical above, other possibilities should be considered, like this haplogroup resulting from an admixture event in other region long before the arrival of this group to the coasts of Iberia. In the light of these results, we should revise perhaps the absolute terminology used to define putative Mesolithic and Neolithic variants provided that, as demonstrated, this nomenclature can have no relation at all with the economic and social reality of the group under study.

The obtention of genomic palaeogenetic data from these samples would definitely help in a fine identification of the basal Mesolithic ancestry of these Early Neolithic communities ultimately shedding light on the Neolithisation process in these areas. However, even if this is
the case and as pointed out above, in order to obtain an accurate picture of these interaction phenomena it would be paramount to know the Mesolithic genetic background of the regions under study.

The cave of Els Trocs (Bisaurri, Huesca) offers several exceptional aspects to understand the Neolithisation of Iberia, including both mitochondrial and nuclear palaeogenetic results of some of the individuals unearthed in the cave. With regards to the mtDNA, we will focus on two haplogroups: the haplogroup N1a of individual 5 and the haplogroup K1a2a of individual 4 (Haak et al., 2015).

In the first place, the presence of the haplogroup N1a could be accounted for as a direct relation of this community with Central European LBK groups, where this haplogroup is quite frequent (Brandt et al., 2013; Haak et al., 2015, 2010, 2005; Lazaridis et al., 2014). This mtDNA “identity” between Els Trocs and LBK samples can be also appreciated at the nuclear DNA level in two individuals, that show striking genetic similarities with their LBK counterparts (Haak et al., 2015; Hofmanová et al., 2015). The chronological occurrence of both groups may point out that the arrival of people from outside Iberia at the beginning of the Early Neolithic, beyond 5400/5300 (Bernabeu-Aubán et al., 2016; García-Martínez de Lagrán, 2015; Gronenborn et al., 2014). It could even be suggested that the directionality of migrations among Neolithic communities was multiple and there were movements from the Iberian Peninsula and southern France towards Central Europe at the beginning of the Neolithic, as it was proposed some decades ago regarding phenomena such as La Hoguette and Limbourg (Constantine et al., 2010; Gronenborn, 2007; Manen, 1997; Manen and Mazurie de Keroualin, 2003; Van Berg, 1990) and suggested by (Fernández-Domínguez and Reynolds, 2017). Archaeological evidence such as the presence of Mediterranean shells and poppy seeds and particularly pottery (Manen and Mazurie de Keroualin, 2003, p. 166) supports such contacts. Resemblances can be found in some decorative techniques –double and triple combs, plain or decorated strips– and several iconographic patterns, above all the laying out of ornaments in a “T” shape and the horizontal and/or vertical parallel lines flanked by dots or small strokes (Figure 7). These last patterns are found in an extremely wide geographic area from inner Iberia to Central Europe, both in the case of impressed/incised-channeled (Epicardial) pottery and Cardial, La Hoguette, Limbourg and LBK pottery (Manen, 1997; Manen and Mazurie de Keroualin, 2003; Van Berg, 1990; Van Willigen, 2004).

![Figure 7. Selection of potteries of Els Trocs cave.](image-url)

The haplogroup K1a2a found in another individual of Cueva de Els Trocs, is also present in Cova Bonica (CB13), albeit both samples have different haplotypes (Olalde et al., 2015). Therefore, it can be suggested that both Central European and Mediterranean influences can be found in Trocs. Interestingly, pottery is witness to the same phenomenon: Early Neolithic ceramic collections from the cave and other nearby sites show techniques and decorative styles typical of the cardial communities in the Mediterranean coast and of impressed/incised-channeled collections from the Ebro valley and the inner Peninsula (Alday-Ruiz et al., 2012b; García-Martínez de Lagrán et al., 2011; Rojo-Guerra et al., 2013).

As an alternatively to this explanation, the presence of haplogroups N1a and K1a2a could also point to different Neolithic populations (LBK, Cardial, I/I-C) with the same genetic stock that arrived in Europe from the same region in the Near East or Eastern Europe, as, for instance, Olalde et al. have suggested (Olalde et al., 2015, pp. 4-5). In fact, haplogroup N1a is
also present in Early Neolithic LBK Hungarian populations (Szécsényi-Nagy et al., 2015) and can be traced back to Anatolia, being its earliest occurrence in the PPNB population of Boncuklu Höyük (Kılınç et al., 2016). Therefore, if this hypothesis is accepted, then the observed parallels in the archaeological record will be a result of migrations and contacts in a second stage in the Neolithisation process or at the beginning of the Early Neolithic. However, if we assume that a common genetic background existed for both Cardial and LBK cultures, this characteristic should be observed in all samples of both populations. However, both mitochondrial and genomic DNA data show variability within both groups, even in the same site. Again, this data could be a result of different level of admixture with local (or distant) Mesolithic groups in the different individuals, a different level of biological interaction with other Neolithic groups or a combination of both.

3.4. The interpretation of palaeogenetic data and the Neolithisation process

Current palaeogenetic data from the Iberian Peninsula can support different hypotheses regarding the Neolithisation process. As we have just seen in the case of Los Cascajos, on the basis of diverse interpretations of the available archaeological record, we can propose a local or even regional (in the Upper Ebro area) process of Neolithic colonisation, a model of interaction among different groups or even a Neolithisation process carried out by local Mesolithic groups. Similarly, the genetic parallels found between certain Iberian and LBK samples could be either explained through a model of continental migration from a single source or through local contacts between both cultures. In sum, these and other cases presented above are fine examples of the limits imposed by palaeogenetic analyses and of the cautions to be taken when offering our interpretations. In this sense, it is noteworthy to verify how the recent debate regarding the Neolithisation process in Iberia has focused precisely on the archaeological contexts and on the theoretical premises supporting the participation of hunter-gatherers in the process. In fact, the archaeological record does not provide direct proof of such participation (proposals on the contrary in Alday-Ruiz, 2012; Montes-Ramírez and Alday-Ruiz, 2012 for the Ebro valley). As a result, palaeogenetic data have virtually become the only evidence in this regard.

One interesting interpretation of the Neolithisation process in light of the available mitochondrial genetic data is the model of a random Neolithisation of the territory. This view is based on the “different genetic impact of Neolithic groups on various regions”, as stated Hervella et al., 2012 p. 1). In our judgment, it is possible that the same data can derive from precisely the opposite, that is, from a planned search and occupation of certain geographic and environmental areas by Neolithic communities that has been labelled as leapfrog colonization. The main characteristic of this type of colonization is the existence of an isolated movement in one direction, whose goal would have been to occupy concrete biotopes: endoreic areas, wetlands, with lagoons or interfluves that favour the development of agriculture and provide grazing ground (Bernabeu-Aubán et al., 2015; García-Martínez de Lagrán, 2008; Sherrat, 1980; Van Andel and Runnels, 1995). This colonisation model resulted in the sparse occupation of space though it was geographically extensive from the absolute point of view, enabling new contacts with different communities (Mesolithic or Neolithic) over a wide area which, in turn, would favor the quick Neolithisation of the territory. This would be the case in the example of Figure 6 and Table 3.

To sum up, we must be aware that it is possible that historical processes determine the general view of palaeogenetic data, thus providing us with an additional variable when analysing all the information and formulating our hypotheses.
4) CONCLUSIONS

1-) Current palaeogenetic data regarding the Neolithisation process and the Early Neolithic period in the Iberian Peninsula are heavily conditioned by the following factors:
   a) the scarcity of the sample, that leads to a geographic and analytical bias.
   b) the study of Early Neolithic individuals, subsequent to the Neolithisation process itself, and the extreme paucity of analyses of Late Mesolithic indigenous hunter-gatherer communities of the same areas that would have played a leading role in the Neolithisation process.
   c) the tendency to generalise the data, using evidence from Middle or Late Neolithic communities with different archaeological and genetic realities as “proxies” for the earliest Neolithic groups in the region. (Günther et al., 2015; Sampietro et al., 2007).
   d) the prevalence of mtDNA analyses, with the problems and interpretative limits mentioned previously. In this regard, nuclear DNA studies are extremely rare, just circumscribed so far to Cueva de Els Trocs and Cova Bonica.

2-) Current palaeogenetic data has confirmed the existence of colonisation phenomena in the Neolithisation process in the Iberian Peninsula, a key aspect to propose historical explanations for the process. The big question is how much migrant population is needed for Near Eastern variants to appear in the Iberian Neolithic several millennia after their origin in the Fertile Crescent and after crossing a whole continent. On the other hand, these migrations can confirm some proposals about the same process, for instance, the need for a direct contact between teacher and learner in the spread of farming and stockbreeding technology, as well as the level of development or complexity in the Early Neolithic (Rojo-Guerra et al., 2008; Stika, 2005; Zapata et al., 2004).

3-) The same data has also demonstrated the presence of local hunter-gatherer groups in this process. From an archaeological perspective, this is a rather significant finding, since the evidence does not clearly show their participation. The presence of Paleo-Mesolithic DNA in some of the studied individuals would point to a considerable participation of local groups in the Neolithisation process. Nonetheless, in this case, as it happens in the previous one, it is hard to define and quantify the term “considerable”. Once again, since the definition of Neolithic and Mesolithic constitutes a key aspect, we must gather information about the genetic substratum of each region and of the earliest colonising Neolithic communities. As we have seen, the use of mtDNA is quite controversial, for just the maternal genetic material is detected and some lines can disappear with time. On the contrary, nuclear DNA is far more detailed, but the available data is still too incomplete.

4-) Available palaeogenetic information, while scarce, suggests that the Iberian Peninsula was a point of confluence and arrival of different Neolithisation currents. Archaeological evidence dating from the Neolithisation process and Early Neolithic already showed that aspect. Pottery in particular has corroborated different influences and ways of access of Neolithic into the Iberian Peninsula. In this case, as we have seen in a previous section, genetic data could support these hypotheses, although alternative interpretations are possible. In sum, current data seem to add even more questions and open new possibilities to an already variable and diverse process.

Bibliography


