EVOLUTIONARY GENETICS OF *HOMO NEANDERTHALENSIS*: ADAPTIVE TRAITS AND METHODOLOGICAL PROBLEMS

Elena Gigli

TESI DOCTORAL UPF / 2011

Director de la tesis:

Dr. Carles Lalueza-Fox

Institut de Biologia Evolutiva (CSIC-UPF)

UNIVERSITAT POMPEU FABRA

A
Leda e Silio
Fedora e Federico

Margherita e Sauro

Fabio

Ringraziamenti

El segundo día en Barcelona me perdí de camino a casa desde el supermercado. Después de dar varias vueltas, las bolsas de la compra se rompieron. Con toda la compra por el suelo me puse a llorar pensando que NO PODIA MÁS!!!!!!.. Pero con un poco de paciencia todo se arregla.... Una pareja de abuelos me acompañó a casa y desde ese día ya han pasado casi 4 años ... Llegado el final de este camino de estudio y de vida, quiero agradecer en primer lugar a tres personas: a mi tutor, Carles Lalueza-Fox, no solo por sus conocimientos científicos que me ha trasmitido, sino por ser *una persona gentile* que ha entendido todas mi necesidades; a Jaume Bertranpetit, que me dio tiempo para reflexionar en un momento de locura cuando yo quería salir y dejar el doctorado; a Oscar Ramirez *il mio secondo tutor* por su amistad y por el tiempo que me ha dedicado.

Gracias a todos los compañeros de trabajo que siempre me han demostrado afecto y amistad. En particular a Natalia, Begoña, Carla y Gemma por las charlas realizada en los buenos y no tan buenos momentos. Gracias a Txema y Ángel por las muchas veces que han arreglado mi ordenador estropeado (o no) por mi culpa. A la *pequeña familia* Italiana, Ludovica, Giovanni, Francesca y Valeria que siempre han estado cerca. *Vale grazie di tutto*.

Un agradecimiento especial a Silvia y sus padres que me han tratado como un miembro más de la familia desde el primer día que llegué. Su cariño ha sido de gran ayuda y apoyo. Gracias también a Mónica y Roser por todo el tiempo que hemos pasado juntas.

Grazie ai miei amici di sempre che fanno il tifo per me incondizionatamente. A Giada, Katy Sceila Elisa Elena e Michela, la vostra amicizia e' molto importantre. A Sara che dal primo giorno di universita' siamo state partecipi l'un l'altra degli eventi importanti che ci sono capitati. A Renzo per le dissertazioni non prorpio ortodosse sull'evoluzione umana. Al gruoppo degli antropologi di Firenze, che non sono piu' un gruppo ma tanti amici con occupazioni diverse che continuano a fare parte della mia vita.
La mia famiglia e' una piccola tribu' fatta di legami forti. E' grazie anche al loro affetto ed al loro amore che sono arrivata alla fine di questo dottorato. Grazie ai miei genitori per il loro supporto morale

in ogni momento. Senza di voi mi sentirei persa. Grazie a mio fratello primo testimone della mia vita.

Gabriele grazie perche' questi anni non sono stati facili, perche' abbiamo superato tante cose, perche' mi sei stato vicino anche se lontano.

Questa tesi e' dedicata ai miei nonni, l' affetto, e la tristezza piu' grande nella loro assenza, che abbia provato.

Abstract

The evolutionary history of *H. neanderthalensis*, interwoven with that of *H. sapiens*, has always fascinated the scientific world. Recent adavncess in paleogenetics shedds new light on the phylogenetic relationship between Neandertals and modern humans. The studies developed in this thesis intend principally to control the contaminants through the development of an anti-contamination protocol for decreasing the human contamination in pre-laboratory phases. We designed a PCR-based method specific for reducing human contamination during the laboratory analysis, and we analyzed the fragmentation pattern of the ancient sequences by massively parallel sequencing technologies. Furthermore, we studied two nuclear genes, *TAS2R38* -associated to bitter taste perception- and *ABO* blood group system –involved in natural immunity- that provide specific information on aspects of the Neanderthal phenotype and adaptation.

Resum

La historia evolutiva *d'H. neanderthalensis*, imbricada amb la d'H. sapiens, ha fascinat sempre el món científic. Avenços recents en paleogenètica aporten una nova llum sobre la rel·lació filogenètica entre els neandertals i els humans moderns. Els treballs d'aquesta tesi intenten principalment controlar els contaminants mitjançant el desenvolupament d'un protocol d'anti-contaminació que disminueixi la contaminació humana de les mostres en la fase de pre-laboratori. Hem desenvolupat un mètode basat en la PCR específic per a reduïr els contaminants humans durant l'anàlisi en el laboratori, i hem analitzat el patró de fragmentació de les seqüències antigues amb tècniques de seqüenciació massiva en paral·lel. A més a més, hem estudiat dos gens nuclears, el TAS2R38 –associat a la percepció del gust amarg- i el grup sanguini ABO –implicat en la immunitat natural- que proporcionen informació específca sobre aspectes del fenotip i de les adaptacions dels neandertals.

Preface

Incessant research since the discovered of the Neanderthals has produced a great amount of data, both biological and cultural, placed in their environmental context. This has allowed to draw a picture fairly defined of their physical characteristics and some cultural aspects as lithic industries and perhaps body ornaments. Moreover the strict phylogenetic relationship between *H. neanderthalensis* and *H. sapiens* has made the former human group central to the debate of the meaning of humankind.

In recent years, the study of ancient DNA has assisted the analysis of the fossil record by shifting the focus on the information contained in mitochondrial DNA and in nuclear DNA. The latest technological developments, such as high-throughtput DNA sequencing technologies, have allowed to sequence genomes from late Pleistocene species. Despite many guidelines, designed to solve the problem of contamination which affects the aDNA field, discriminating between endogenous and exogenous sequences is still difficult. In fact, contamination, specially with modern human DNA, can take place during excavation, restoration and cleaning of samples, but also in the subsequent laboratory analyses.

In this context of technological innovation and great attention to authenticity of the results, the works presented in this thesis are intended to improve the control of contaminant DNA in both phases of pre-laboratory and laboratory. Moreover the retrieval of two nuclear genes, associated to bitter taste perception and natural immunity, can contribute to the understanding of the adaptation of *H. neanderthalensis* to their specific Eurasian environment.

Index

1. INTRODUCTION

1.1. Neanderthal world

1.1.1. Neanderthal features

Homo neanderthalensis is an extinct species of homininae, who occupied parts of Europe and western Asia in the period between about 130 ka and 30 ka. (Weaver et al. 2009; Hublin 1998) Although the precise level of phylogenetic and genetic distinction between these hominids and *Homo sapiens* is still debate, there is no doubt that Neanderthals are an entity characterized on the morphological and molecular level. Despite their geographical and individual variations they represent a coherent group (Facchini and Belcastro 2009).

The heritage of fossil hominids is very wide and includes findings that lead to some, but not all, features typical of Neanderthals. Looking at the wider context it is clear that the evolution of hominids in Europe during the last half million years is more complex than a simple linear trajectory. The most accepted model to try to explain the Neanderthal evolution is the ‘‘Accretion Model’’ (Rosas et al. 2006; Hublin 2009) which accounts for the progressive appearance of Neanderthal morphology through time, beginning around 450 ka. According to this hypothesis, the Neanderthal lineage became isolated in Europe due to the severe climatic conditions of the Pleistocene. The geographic range of Middle Pleistocene European hominins would have been restricted and isolated during glacial maxima by ice sheets covering northwestern Europe and by their associated permafrost zones. In these conditions of isolation, the Neanderthal morphology is thought to have become gradually fixed, partly through natural

selection as an adaptation to cold climate conditions (Weaver et al. 2009) but perhaps primarily through the process of genetic drift (Hublin 1998).

In Rosas et al (2006) four broad stages of Neanderthal evolution have been described:
Stage 1 includes ‘‘early pre Neanderthals”, the Middle Pleistocene archaic specimens. These hominins are considered to show incipient Neanderthal features mainly in the facial region.
Stage 2 specimens are termed ‘‘pre Neanderthals’’. They are thought to exhibit Neanderthal morphology more clearly, showing Neanderthal features also in the occipital area.
Stage 3 ‘‘early Neanderthal’’ specimens show most Neanderthal traits in the posterior cranium and some also in the temporal region.
Stage 4 comprises the ‘‘classic Neanderthals’’, showing fully expressed Neanderthal morphology.

According to the accretion hypothesis, Neanderthal evolution was a genetic process with no speciation event resulting in the appearance of this taxon. The Middle Pleistocene specimens from Europe and African Middle Pleistocene traditionally have been included in the species *Homo heidelbergensis*. This taxon is viewed by some as ancestral to both Neanderthals and modern humans (Rightmire 1998). An alternative interpretation (Lalueza-Fox et al. 2005) is that, after Middle Pleistocene populations became isolated in Europe, there was a speciation event leading to the evolution of Neanderthals. If this were the case, then the taxon *H. neanderthalensis* should include only the specimens postdating this event, while *H. heidelbergensis* can include both African and European Middle Pleistocene human fossil. Some researchers

consider the early skeletal remains from Gran Dolina Atapuerca (dated to780 ka) as a distinct species, *H. antecessor*, itself the last common ancestor of both Neanderthals and modern humans (Bermudez de Castro et al. 1997, Hublin 2009) (Figure 1).

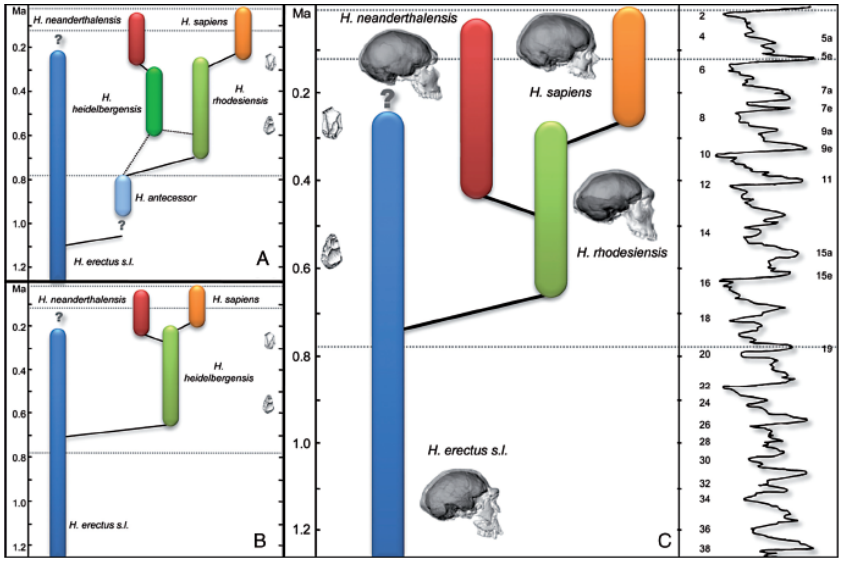


Figure 1: Hypothesis on Neanderthal origins: three evolutionary scenarios are rapresents: (*A*) depicts an early, (*C*) depicts a late, and (*B*) depicts an intermediate divergence time. The marine 18O isotope record is indicated on the right of tree (C). A biface on the chronological scales indicates the time of emergence of the Acheulean in Europe and a flake indicates the occurrence of fully developed *Levallois debitage* in Europe. Horizontal dashes indicate the limits of the Middle and Late Pleistocene. (Hublin et al. 2009)

For about a century and a half, researchers have documented morphological differences between the crania of Neanderthals and those of modern humans. It is true that for most features Neanderthal and modern human ranges of variation overlap, but on average, and when multiple features are considered in combination, most specialists agree that Neanderthal and modern human crania can be distinguished morphologically from one another (Weaver 2007; Weaver et al. 2009).

In general they show a long, relatively low and globular (en bombe = subspheroid or oval) braincase; large endocranial capacity ranging from 1,245 to 1,900 cc and averaging about 1,520 cc. (in *H. Sapiens* the endocranial capacity range between 1200-1850 cc) (MS Ponce de Leon et al. 2008)

1.1.1.1. Cranium

The **frontal bone** is low and receding with a continuous supra-orbital torus generally forming a double arch above the orbits and usually separated from the frontal scale (or squama) by a sulcus or gutter. The orbits are large and rounded, the chin (forward protrusion of the mental symphysis) is variably developed but usually absent (Klein 1999; Weaver et al. 2009).

Neanderthal **temporal bone** presents a supero-inferiorly low and antero-posteriorly short temporal squama, a robust zygomatic process with a relatively pronounced lateral projection, a strong

supramastoid crest, an elevated position of the external auditory meatus relative to the zygomatic process, a relatively small mastoid process, and a large juxtamastoid eminence (Balzeau and Radovcic 2007).

The **bony labyrinth** (Figure 2) or osseous labyrinth, consists of three parts: the vestibule, semicircular canals, and cochlea. In mammalian bony labyrinth tends to show a consistent, species-specific morphology and it can help identify a fossil afiliation (Hublin et al. 1996; Spoor et al. 2003) In Neanderthals the arc sizes of the vertical (anterior and posterior) semicircular canals are smaller than in modern humans and *H. erectus*, whereas its lateral canal is larger. Moreover, the position of the posterior canal was described as markedly inferior relative to the plane of the lateral canal (Hublin et al 1996; Spoor et al. 2003). Neanderthal have smaller anterior and posterior canals than modern humans and they are particularly striking if it is considered that body mass, which is positively correlated with canal size (Spoor 1993), was probably higher than in modern humans (Ruff et al. 1994). The labyrinth in Neanderthal is also characterized by a markedly inferiorly-positioned posterior semicircular canal (Hublin et al. 1996) and it is smaller than in modern humans. A trend of anterior and posterior semicircular canal enlargement has been functionally linked with the emergence of obligatory bipedalism (Spoor et al. 1994). That Neanderthals as habitual bipeds nevertheless have relatively small canals demonstrated that the link between canal size, head movements and locomotor behaviour is complex, possibly involving factors such as neck and ocular motility (Hublin et al. 1996).

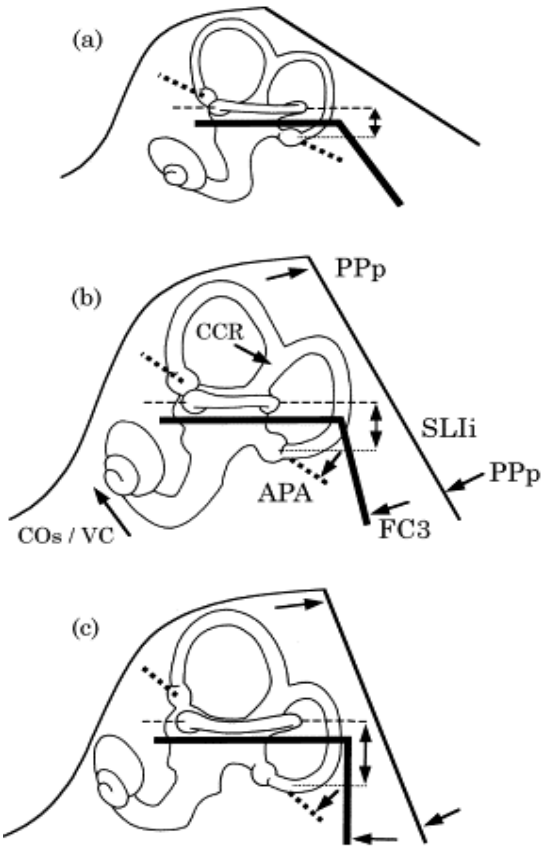


Figure 2: Lateral view of the bony labyrinths of (a) *Pan paniscus*, (b) a Holocene modern human and (c) La Ferrasie 1 Neanderthal. The labyrinths are aligned according to planes of their lateral semicircular canal. (LSCm, dashed line), and the course of the second and third parts of the facial nerve canal (thick line) and the endocranial petrosal contour at the level of the common crus are indicated. Single headed arrows indicated morphological differences comparing (b) to (a), and (c) to (b). The ampullar line (APA, thick dotted line), the third part of the facial nerve canal (FC3) and the posterior petrosal surface (PPp) are increasingly vertically inclined from (a) to (c), and the inferior component of the sagittal labyrinthine index increases (SLIi, double headed arrows). Neanderthal and Holocene humans are similar in having a common crus that is tilted posteriorly (CCR), and a cochlea that is positioned more superiorly (Cos and VC) than in non hominin primates (Spoor et al. 2003).

The characteristics of the Neanderthal **nasal region** distinguish these extinct human relative from *H. sapiens* (Schwarz and Tattersal 1996). In fact in the Neanderthal nose there is the presence of a rim of raised bone that projects from each side of the rim of the anterior nasal aperture just within its anterior edge, forming a secondary "internal margin" (Figure 3) (Schwarz and Tattersal 1996). This rim runs one-third to halfway up the inner nasal wall on both sides and then expands to become a wide, broad-based and bluntly pointed mass that protrudes medially into the nasal cavity. This medial projection fades superiorly into a low ridge that continues to frame the nasal cavity within its external margin. On its posterior side, the horizontal inferior margin of the rim may be rounded, and the vertical portion is bounded by an open lacrimal groove. Varying only in relative size, the vertical medial projection is present in all the adult Neanderthal specimens.

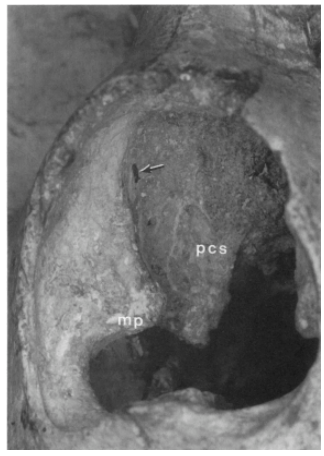


Figure 3: Frontal view of the nasal cavity of Gibraltar 1 cranium, illustrating the internal margin bearing a large internal projection (mp), behind which is the large swelling (pcs) within the posterior component of the nasal cavity, partly obscured in this view by matrix. The arrow indicates the hole in the wall of the swelling that reveals the enlarged maxillary sinus (Schwarz and Tattersal 1996).

Another unusual structure, discovered in Gibraltar 1 specimen, is a distinct swelling of the lateral nasal wall. In the more general and primitive configuration, there is no medial swelling of the posterolateral wall of the nasal cavity. The nasal cavity of extant terrestrial mammals is normally filled to varying degrees with two, three, or even four pairs of turbinates that derive and swell laterally from the ethmoid bone that lies in the midsection of the cavity wall, and given the peculiar nasal morphology in Neanderthals, it appears likely that the turbinates of these extinct hominids, and possibly also the ethmoid, were configured in an unusual manner. Whether the medial swelling is functionally a specialized nasal structure, as the medial projection appears to be, or whether it is merely a passive result of maxillary sinus expansion remains to be determined (Schwarz and Tattersal 1996). A least three notable apomorphies distinguish the structure of the Neanderthal internal nasal cavity not only from that of other hominids but also from primates in general. The apomorphies are the development of an internal nasal margin bearing a well-developed and vertically oriented medial projection, the swelling of the lateral nasal cavity wall into the capacious posterior nasal cavity, and the lack of an ossified roof over the lacrimal groove (Schwarz 1996). The large internal nasal margin may serve to expand the internal surface area, thus allowing for an increase in ciliated mucosal covering. The placement of this margin also suggests a location ideally suited to be the initial vehicle to confront inspired air or the last opportunity to interact with expired air. Both the internal nasal margin and second feature, as welling of the lateral nasal cavity, may be related to be expansions of the paranasal sinus system. The relationships/functions of an unossified roof over the lacrimal groove, is less clear. The absence of a rigid roof, however, would

clearly permit more expandability for components of the naso-lacrimal duct system (which, inhumans, contains a venous plexus forming erectile tissues and can, when engorged, obstruct the duct) (Laitman et al. 1996). Finally, some aspects of the Neanderthal nose, such as narrow superior internal nasal dimensions in specimens deriving from glacial periods, are consistent with the predictions of cold climate adaptation generated from empirical observation and clinical and mathematical modeling of nasal function in recent humans (Klein 1999).

One feature that deviates strongly from a well established pattern of cold climate morphology is the characteristically wide Neanderthal nasal aperture. This feature is usually associated with populations adapted to more tropical climates. Nasal breadth, as one component of the nasal index, exhibits a strong ecogeographic clinal distribution with recent modern humans indigenous to higher latitudes exhibiting, on average, narrower nasal breadths than recent modern humans found in equatorial regions (Weaver et al. 2009). This adaptive trend is a reflection of respiratory physiology dynamics associated with varying levels of humidity and temperature. A narrower nasal aperture functions better in cold, dry environments by moistening and warming air as it is inhaled as well as recapturing moisture during exhalation. In contrast, a wide nasal breadth functions to dissipate heat in more tropical environments (Holton et al. 2008). Given the relationship between climate and nasal breadth, one would predict that Neanderthals, associated with increasingly colder, drier conditions, should exhibit relatively narrower nasal apertures. Probably, nasal breadths in Neanderthals, which would have otherwise narrowed through climatic selection, were ontogenetically constrained by the

plesiomorphic retention of large anterior teeth and broad palates that characterized their Pleistocene ancestors (Holton et al.2008). Probably the wide Neanderthal nose is the result of a large anterior dentition with increased growth of the anterior palate in effect pulling apart the edges of the nasal aperture (Holton et al. 2008).

The evolution of life history (pace of growth and reproduction) was crucial to ancient hominin adaptations. The study of **dental development** facilitates assessment of growth and development in fossil hominins with greater precision than other skeletal analyses (Smith et al. 2007). Neanderthals have a distinctive suite of dental features and one type of the taxonomic diagnosis as *H. neanderthalensis* is based on the presence of a mid-trigonid crest, anterior fovea, and slight taurodontism. These traits show a significantly higher frequency in Neanderthals than in Upper Paleolithic and recent humans. Dental tissue conformation distinguishes Neanderthal molar (Figure 4) from those of modern humans and is better suited to intra-generic comparisons than enamel thickness (Smith et al. 2009). Crown formation time of the mesiopalatal cusp was 872 days, yielding ages at cuspal crown completion of 2.35 years and 2.22 years for the mesiopalatal and mesiobuccal cusps, respectively. Coronal extension rates for the first molar cusps were outside the range of modern human values but also show some similarity to chimpanzee values (Bailey et al. 2002).

Longer formation times in anterior teeth relative to modern humans may stem, in part, from the larger size of these teeth in Neanderthals.

In the case of Neanderthal molars, relatively short formation times may be explained, in part, by differences in cuspal enamel thickness, which are 60–90% of modern human mean values. Given that overall daily secretion rates do not vary between Neanderthals and modern humans. Neanderthal molar cuspal enamel formation times are likely to be shorter than in modern humans. First molars are often considered the most reliable predictors of life history and there is a strong correlation of emergence ages between permanent teeth, particularly for adjacent teeth such as first and second molars. At 8 years of age, the second molar is beyond clinical (gingival) emergence. Second molar eruption occurs on average at 10–13 years of age in global human populations, suggesting that Neanderthals show advanced second molar development relative to modern human populations. Analysis of third molar calcification (approximately three-fourths crown complete) shows that it was 2-3 years ahead of the expected developmental stage of other molar teeth. This pattern, however, does not appear to be unique to Neanderthals; early third-molar initiation ages have also been reported for modern humans of southern African origin (Smith et al. 2007).

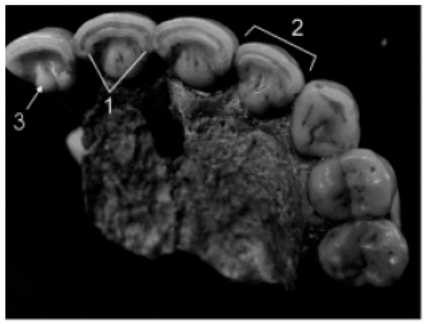


Figure 4: Marked shoveling (1), labial convexity (2), and lingual tubercles (3) on maxillary incisors and taurodontism (arrow) of mandibular molars (Bailey et al. 2002).

Neanderthals mandibular second premolars (Figure 5) presents a distinct configuration: a complex occlusal surface topography and an asymmetrical lingual contour. The tooth's complex topography results from a combination of a strong and continuous transverse crest, a high and well-developed metaconid, and extra fissures, ridges, and lingual cusps an asymmetrical lingual contour, marked transverse crest, a mesially placed metaconid and extra lingual cusps, the *H. erectus* and modern human all have symmetrical lingual contours and lack a prominent transverse crest (Bailey et al. 2002).

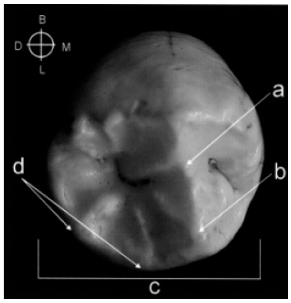


Figure 5: Premolar crown characters referred to in the text: transverse crest (a), metaconid (b), lingual contour (c), and lingual cusps (d). The symbol indicates orientation of the tooth: B, buccal or cheek side of the tooth; L, lingual or tongue side; M, mesial or toward the front of the mouth; and D, distal or toward the back of the mouth (Bailey et al. 2002).

The incisors are usually larger than those of earlier people and significally larger than those of modern people. The maxillary incisors are usually shovel shaped and mandibular and maxillary incisors tending to exhibit distinctive rounded wear on the labial lip surface. The root fusion with enlargement of the pulp cavity (taurodontism) is common in the cheek teeth. It is easily discernible the gap (retromolar space) between the posterior wall of the lower third molar and the anterior margin of the ascending ramus resulting from the combination of a long mandible, a short post canine dentition and a relative narrow ascending ramous. The mandibular foramen is usually horizontal-oval (Klein 1999).

The **occipital bone** (Figure 6) shows a bilaterally protruding transverse occipital torus with a depression in the mid-sagittal plane. The upper lip (supreme nuchal line) of the torus delimits the occipital plane superiorly and is very smooth, while the lower lip (superior nuchal line) is more clearly marked and indicates the inferior limit of the occipital plane (Bastir et al. 2010).

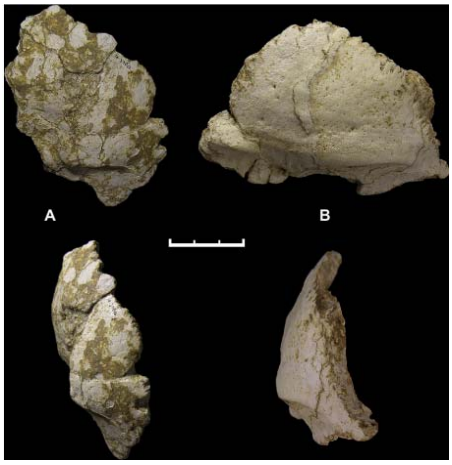


Figure 6: Occipital bone in original fossils from El Sidron cave (Asturia, Spain) in posterior (upper part) and lateral views (lower part). A) SD-370a, B) SD-1149. Scale bar is 3 cm. (Bastir et al. 2010).

Moreover, the occipital bone of Neanderthals contains the suprainiac fossa that appears as a horizontal oval-shaped depression located above inion on the occipital plane. It is generally formed by one depression, but some bony relief may separate it horizontally; it may also be composed of two vertically dissociated depressions. The suprainiac fossa (Figure 7) is variable in number, shape, depth, and size, but its presence is constant in Neanderthals. Its surface is variably pocked, a characteristic that could reflect resorptive activity (Balzeau and Rougier 2010).

Probably the fossa appears early in development and it is present in all immature Neanderthals (Hublin 1980). A depression above the occipital superstructures and/or various degrees of resorption of the external cranial surface in the medial part of the occipital bone have been observed in anatomically modern *H sapiens* from Upper Paleolithic European sites (Kramer et al. 2001). Suprainiac fossa was described in non-European anatomically modern *H. sapiens* from the Late Pleistocene (Trinkaus 2004). The possible presence of the suprainiac fossa in non-Neanderthal hominins raises the question of the autapomorphic status of this feature in Neanderthals (Caspari 2005). Variations in absolute bone thickness in the area of the Neanderthal suprainiac fossa are correlated with variations in the representation of the three structural layers of the bone. The suprainiac fossa of a large sample of Neanderthal specimens corresponds to a thinning of the diploic layer, whereas no profound remodeling non variation in thickness of the external table occurs in this area. To the contrary, the depressions on the occipital bones of anatomically modern *Homo sapiens* from Europe and Africa systematically correspond to a resorptive surface of the external cranial surface, which affects only the external table and not the underlying layer. Moreover,

these depressions present some differences in their shape, disposition, and surface with the suprainiac fossa observed in Neanderthals. This demonstrates that the depressions observed on anatomically modern *Homo sapiens*, as well as on other hominin species, are different from suprainiac fossa and these features can no longer be considered as homologous. As a consequence, the Neanderthal suprainiac fossa appears to be an autapomorphic trait for this fossil population (Balzeau and Rougier. 2010).

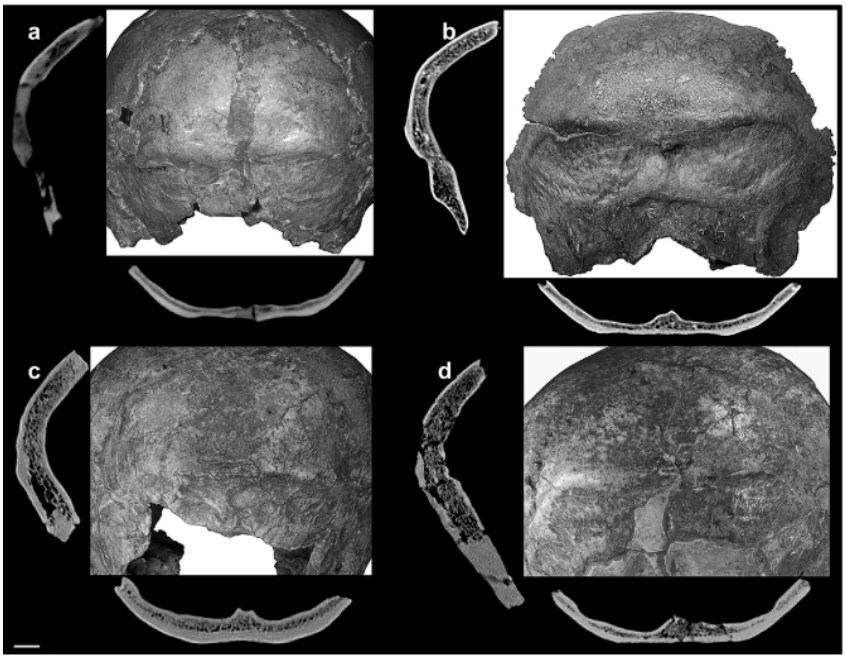


Figure 7: Posterior view of original Neandertal fossils, together with orthogonal CT slices cutting along the largest extension (bottom) and the median vertical extension (left) of their occipital depression. a) La Quina H5, b) Salzgitter-Lebenstedt 1, c) Spy 1, d) Spy 10 (scale bar ¼ 1 cm) (Balzeau et al. 2010).

The occipital '*bun*,' or chignon, (Figure 8) is one of the most frequently discussed Neanderthal traits. It has been cited widely as a Neanderthal characteristic. It is situated in the occipital bone and it is described variably as a posterior projection or a great convexity of the upper scale of the occipital bone. It is often associated with the presence of a depression of the area around lambda (''lambdoid flattening'') on the occipital and parietal bones. Not only does the 'bun' involve a lambdoid flattening, but it is also associated with a flattened parietal and flat nuchal part of the occipital. Its posterior projection is expanded laterally in Neanderthals and it is manifested in both the exterior and interior aspects of the occipital bone below lambda and above the internal occipital protuberance. Probably, the convexity of the occipital squama and the flattening of the lambdoid region are not the salient features of Neanderthal 'buns,' and do not greatly contribute to the metric separation of modern humans and Neanderthals, instead, it is the combination of occipital squama convexity/lambdoid flattening with occipital bone orientation relative to the rest of the cranium that differentiates Neanderthals from modern humans (Gunz et al.2006).

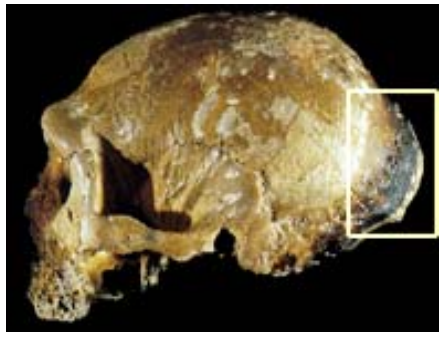


Figure 8: In the box, typical occipital bun in Grotta Guattari (Italy) Neanderthal specimen (Piperno and Giacobini 1992; Stiner 1992).

Other important Neanderthal features are the mastoid process usually small with a pronounced mound or pump of bone behind the auditory aperture. The skull base is flatter between the hard palate and the forame magnum than that of modern people (Klein 1999).

1.1.1.2. Post Cranium

The Neanderthals were extremely robust and heavily muscled .The most recent studies on the shape and size of the Neanderthal **thorax** suggest that these humans had a large thorax for their stature, but probably, this characteristic in Neanderthals reflects greater body mass (compared to modern human populations). The link between the wide bi-iliac breadths of Neanderthals and a larger inferior thoracic circumference, invoke the possibility of multiple factors, including cold-climate adaptation, heightened activity, and genetic drift, for this complex (Weinstein et al. 2008; Gomez-Olivencia et al. 2009). Among Neanderthals, Kebara 2 (K2) (Valladas et al.1987; Schwarcz et al.1989) preserves the most complete rib cage found to date and has been used as the model for the thorax of the articulated complete Neanderthal skeleton. It shows a capacious thorax with a dome-shaped superior rib cage and a flaring lower rib region that resulted in a bell-shaped thoracic cage that did not conform to the expected barrel-shaped form.

The ribs of K2, together with other Neanderthal rib remains, have also been used to hypothesize the presence of eco-geographical patterning in the upper thorax of Neanderthals. Based on both the first and second rib remains in Regourdou 1 (Vandermeersch and Trikaus 1995) and La-Chapelle-auxSaints (Trinkaus 1985; Tappen

1985) on the one hand, and Tabun C1 (McDermott et al. 1993; Coppa et al. 2005) and Kebara 2 on the other, it was suggested that western European Neanderthals exhibited greater anteroposterior upper chest expansion compared to their Near Eastern counterparts (Weinstein et al. 2008).

Analyses of Neanderthal **clavicular** proportions had also previously suggested the possibility that Neanderthal upper thoraces were more anteroposteriorly expanded compared to modern humans, and that this expansion might have been somewhat greater in European than Near Eastern Neanderthals (Klein 1999). The relationship between a large skeletal rib cage and a large pulmonary capacity is not straightforward and vital capacity depends a great deal on other non skeletal tissues. On the other hand, studies of modern human high altitude adapted populations show both higher vital capacities and greater chest circumferences (and lower stature and body weights) when compared to lowland populations (Weinstein et al. 2001; 2007; Gomez-Olivencia et al. 2009). In conclusion the large chest of Neanderthals likely reflects relatively large ventilatory capacities and the large thoraces would reflects a need for increased oxygen consumption derived from high daily energy consumption (Gomez-Olivencia et al. 2009).

The morphology of the **lumbar spine** (Figure 9) is crucial for upright posture and bipedal walking in hominids. As the size of the vertebral end plates and pedicles correlates with the loads acting on the lumbar spine (Been 2010) and no significant differences were found between Neanderthal and modern humans, it is likely that the axial weight transmission in the lumbar spine was similar

as well. To analyze the lumbar spine characteristic, Kebara 2 (K 2) was taken as a model. The similarity in body length and inter-articular distance suggests no difference in sagittal plane motion. The sagittally oriented laminae at L5 (L=lumbar) and S1 (S=sacral) suggests increased resistance of the ligamentum flavum to ventral flexion at the lower lumbar spine; and the less dorsal orientation of the transverse processes of K 2 at L2-L4 suggests a decrease in lever arm length, and thus in the ability of the erector spinae to perform spinal extension. There is strong evidence for increased lateral flexion in the lumbar spine of K 2 compared with that of modern humans. The craniolateral orientation of the transverse processes of the upper lumbar vertebrae, coupled with the lateral projection of the transverse processes at L2 and L4, are advantageous for powerful lateral flexion of the lumbar spine (Shapiro et al. 1993; 2007; Been et al. 2010). Both of these bring the attachments of the erector spinae and quadratus lumborum muscles further away from the axis of lateral flexion motion (located at the lateral side of the vertebral body) and increase their ability to laterally flex the lumbars. The more cranial orientation at L1and L3 also allows the lower spine to be laterally flexed without bony impingement of adjacent vertebrae (Shapiro et al 1993; 2007, Been et al. 2010). It is not clear whether the relatively wide span of the lumbar transverse processes of K 2 corresponds with its broad pelvis. Pelvic obliquity (which results from lateral flexion of the lumbars when the thorax is stable) is essential for foot clearance during the swing phase of bipedal walking; it also helps to minimize centre of mass displacement during walking.

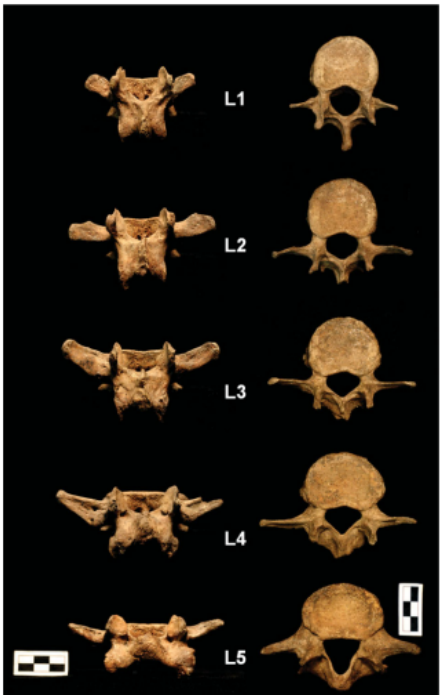


Figure 9: Superior and posterior view of the lumbar vertebrae of Kebara 2. The long side of the black rectangle equals one centimeter (Been et al. 2010).

Pelvic obliquity is produced mainly by the combined unilateral action of the quadratus lumborum and the spinae. The increased overall effectiveness of these muscles in producing pelvic obliquity corresponds with K 2 wide pelvis and it enables sufficient obliquity of the wide pelvis without increasing the load on the intervertebral joints. The similarity in body heigh coupled with the similarity in body length and width between K 2 and modern humans suggests that the spine of K 2 was as stable as the lumbar spine of modern humans (Been et al. 2010). The excessive kyphotic wedging of the vertebral bodies found in the lumbar vertebrae (L3, L4) of K 2 indicates a smaller lordosis angle in the lumbar spine of K 2 compared with that of modern humans (Figure 10).The decreased

lordosis angle in the lumbar spine of K 2 is based on the wedging of the vertebral bodies and does not include the wedging of the intervertebral disk spaces (Whitecome et al. 2007; Been et al. 2010).

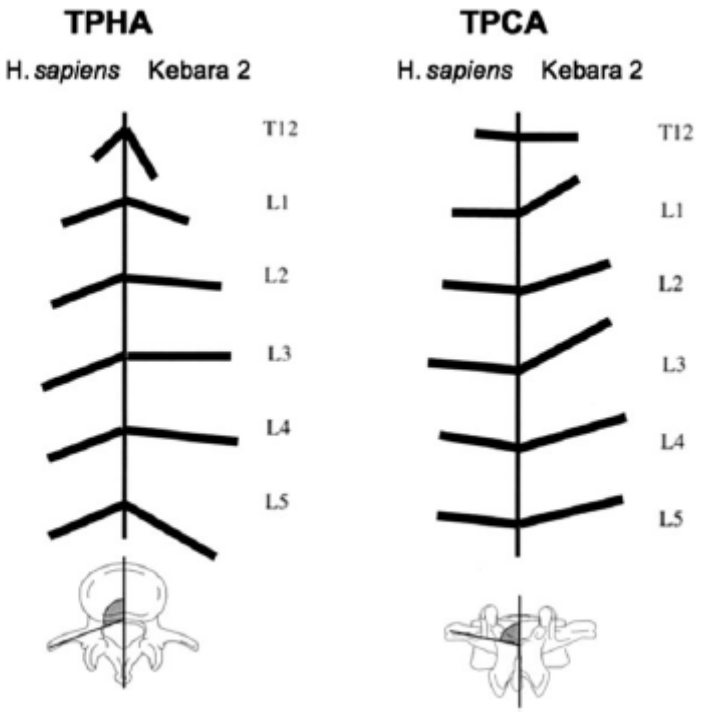


Figure 10: Schematic illustration of the length and orientation of the transverse processes of modern humans and Kebara 2. T=torax; L=lumbar (Been et al. 2010).

Neanderthals and *H sapiens* have similar pelvic areas but their **birth canal** (Figure 11) shapes differ considerably (Weaver et al. 2008). Both female and male humans typically have transversely oval inlets and anteroposteriorly oval outlets, but human females tend to have lower inlet and outlet indices than males. To analyze the Neanderthal birth canal and the birth mechanism, Tabun was taken as a model. Tabun's outlet index is completely outside the

range of variation of humans. Unlike humans, who have an anteroposteriorly oval outlet, Tabun has a transversely oval outlet and the midplane also appears to be transversely oval, but poor preservation of the ischial spine leaves open the possibility that it is round or perhaps anteroposteriorly oval (Weaver et al. 2008).

Nevertheless, because neonatal rotations are mediated by physical resistance within the birth canal, this transversely oval outlet indicates that Tabun has a different birth mechanism than humans. On reaching the outlet, a neonate passing through Tabun's birth canal would align its anteroposterior head dimensions transversely, leading to an occiput transverse exit position (Weaver et al. 2008).

For successful childbirth, both human and Neanderthal females need transversely wider inlets than are found in males, which can be achieved either by having pubic bones that are about the same length as those in males but more coronally oriented or by having pubic bones oriented similarly to those in males but longer. Neanderthals would not be expected to be sexually dimorphic in pubic length. Even though Neanderthals appear to have a different birth mechanism than humans, Tabun's pelvic areas are similar to those of human females suggesting that a human sized neonate would have been able to pass through Tabun's birth canal. This perhaps is not surprising, given that Neanderthals had similar neonatal and adult brain sizes as humans. In addition, the neonate's anteroposterior head dimensions have 132 mm of space in Tabun's outlet (Tabun's transverse outlet dimensions), compared with 122 mm in a human outlet (human female mean anteroposterior outlet). Probably, the childbirth was about as difficult in Neanderthals as in humans (Weaver et al. 2008).

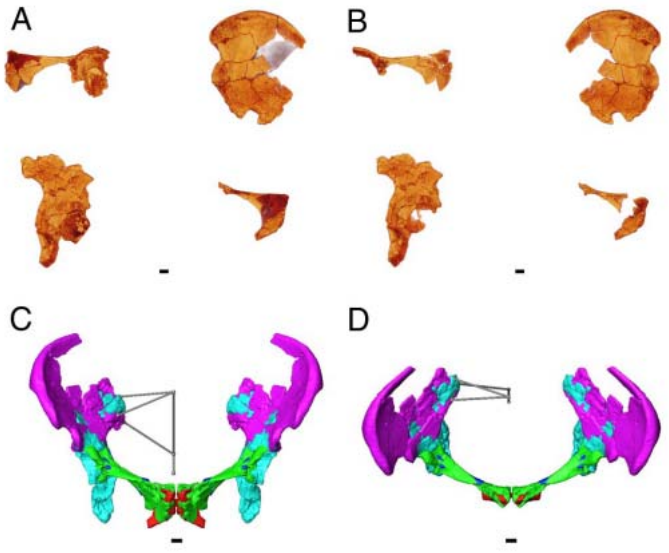


Fig 11: Virtual reconstruction of the Tabun pelvis. Original fragments (A) and after mirroring of the left-sided ilium and acetabular-pubic fragments and segmentation (B). Colors reflect higher (yellow) to lower (red) density. Anterior (C) and anterosuperior (D) views of the completed reconstruction (Scale bar 1 cm.) (Weaver et al. 2008).

Generally Neanderthal upper and lower limbs are short and robust with muscle and ligament attachment areas extremely large and well developed. The scapula show a dorsal groove (ventral in *Homo sapiens*) (Facchini and Belcastro 2009) and the glenoid fossa is long, narrow and shallow. The femur is large and round without the longitudinal bony ridge (pilaster) found on the dorsal surface of modern femoral shaft. Adaptation to cold probably explains the great breadth of the Neanderthal trunk, the shortness of the limbs and especially the shortness of the forearm and lower leg.

This inference follows from the ecological generalizations known as Bergmann's and Allen's rules, which state that all other things equal individuals of a warm-blooded species will tend to have larger body cores and shorter limbs in colder climates. A large body core is advantageous because since core size increases volume which conserves heat increase more rapidly than the skin area which dissipates it. Shorter limbs further reduce heat loss. Both Bergmann's and Allen's rules apply broadly to living humans, in whom trunk breadth tends to increase in distal limb length tends to decrease from equator to the poles. European and west Asian Neanderthal limb bones suggest that European populations had especially short limbs in keeping with their exposure to severe cold (Klein 1999).

1.1.2. Diet

Since the remains of animals living in Mousterian sites are, widely distributed geographically, we can assume that Neanderthals were omnivores. The collection was an important activity which completed the acquisition of meat by one or more specialized hunting prey. Once killed, the carcasses of the animals were usually brought to a base camp for consumption. The strategies included in the Neanderthal diet were flexible in space and time. The prey species varied according to what was available. In times of plenty it was possible to pick plant foods, and hunting activities were most important when the plants were scarce (Facchini and Belcastro 2009).

During the last years, the study of stable isotopic composition of fossil bone collagen has become increasingly important in paleonutrition research (Richards et al. 2008) (Figure 12).

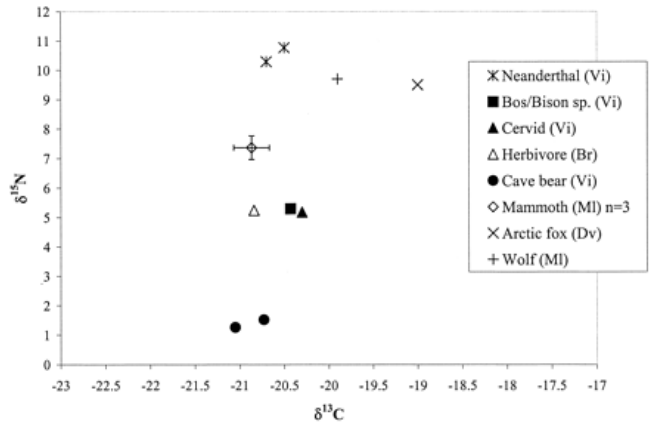


Figure 12: Bone collagen $\delta^{13}C$ and $\delta^{15}N$ values of Neanderthals and associated fauna from Vindija Cave, Croatia (Vi), dated to ≈28,500 years B.P. Included is a single faunal value from the site of Brno-Francouzská (Br), Czech Republic (≈24,000 years B.P.) (Richards et al. 2000)

This new approach is a useful tool for reconstructing past human and nonhuman animal diets. The method is based on the principle that the carbon (d13C) and nitrogen (d15N) ratios in bone collagen are related to the carbon and nitrogen isotopic ratios in foods that have been consumed over the lifetime of the human or nonhuman animal. Tooth dentine has relatively little turnover, and therefore the isotope values will reflect the diet at the time of dentine formation, usually during later childhood (Richards et al. 2008). Numerous controlled feeding experiments and field studies have established that, for mammals with diets with adequate amounts of protein, the d13C value is most related to dietary protein (Ambrose and Norr 1993).

Nitrogen is only present in dietary protein, so the d15N value must reflect dietary protein. There is an approximate 5% offset between the d13C values for mammalian collagen and dietary protein, and similarly, there is an increase of approximately 2% and 4% in d15N values of consumer collagen compared to dietary protein. As mammalian bone collagen constantly turns over, the measured d13C and d15N values reflect a long-term average of dietary protein, likely over many years (Richard et al. 2008). The application of mammal collagen carbon and nitrogen stable isotope measurements in archaeology has focused mainly on three areas. The first is the use of carbon isotopes to discriminate between the consumption of marine and terrestrial resources, as there is a difference in d13C values between these two ecosystems. The second application is the use of d13C values to discriminate between the consumption of C3 and C4 plants, as there is a difference in the d13C values between plants using these photosynthetic pathways. Finally, the third application of this

method is the measurement of d15N values to determine the trophic level of organisms within a food web, especially to discriminate between herbivores, carnivores, and omnivores. This latter application is of relevance to determining the diets of omnivores like Neanderthals and modern humans in Europe (Richards et al. 2008).

Several studies (Bocherens and Druker 2005; Simon et al. 2006, Richards et al. 2008) carried out with stable isotope, show a new light on the dietary habits of Neanderthals. It appears that the old and new Neanderthals, from the colder steppes or temperate forest areas, show isotopic trace of predators of herbivores, animals that they actively hunted. Traces of plant food or derived from fish consumption, are very low, almost absent. But at this regards a recent study by Henry et al. (2011) (Figure 13) reports direct evidence for Neanderthal consumption of a variety of plant foods, in the form of phytoliths and starch grains recovered from dental calculus of Neanderthal skeletons from Shanidar Cave, (Iraq), (Stewart 1977; Agelarakis 1993) and Spy Cave, (Belgium) (Hotton et al. 1976). Some of the plants are typical of recent modern human diets, including date palms (*Phoenix* spp.), legumes, and grass seeds (Triticeae) and other not heavily used today. Many of the grass seed starches showed damage that is a distinctive marker of cooking. The results indicate that in both warm eastern Mediterranean and cold north-western European climates, and across their latitudinal range, Neanderthals made use of the diverse plant foods available in their local environment and transformed them into more easily digestible foodstuffs in part through cooking them, suggesting an overall sophistication in Neanderthal dietary regimes.

Further evidence of this data comes from the micro-tooth wear. Lalueza-Fox et al. (1996) have conducted studies on dental micro wear. They analyzed micro molar tooth wear of modern hunter-gatherer groups with different diets to recognize different diets in relation to the degree of micro wear. The analysis shows, including several samples of Pleistocene hominids, that the Neanderthals are placed with modern hunters, while modern and archaic Homo Sapiens show signs of an abrasive diet rich in vegetables (Facchini and Belcastro 2009).These data support the emerging picture from isotopic studies that Neanderthals have a similar dietary adaptation over a wide range of environments and over a relatively long period of time.

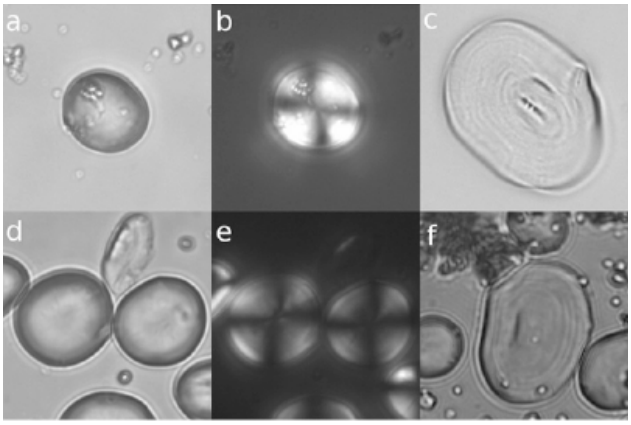


Figure 13: Triticeae cf. Hordeum (barley and close relatives) starch grains recovered from the dental calculus from Shanidar III compared with modern Hordeum starch grains. Each box is 50 μm on a side. (A and B) Hordeum spp. starch grain from Shanidar tooth 4 under brightfield and cross-polarized light. (C) Cooked Hordeum spp. starch from Shanidar tooth 5. (D and E) Starch from modern Hordeum hexastichon under brightfield and crosspolarized light. (F) Starch from modern Hordeum vulgare (domesticated barley) boiled for 5 min (Henry et al. 2011).

1.1.3. Technology and Behaviour

Neanderthals are most commonly, though not exclusively, associated with the Mousterian lithic technology, named after the site of Le Moustier in the Dordogne, (France). Typical of Mousterian industries was the use of both Levallois and discoidal flaking techniques for the production of flakes that could be converted to a wide range of shapes, including various kinds of side scrapers, retouched points, denticulates, notches, and sometimes small handaxes (Mellarset al. 1996; Shea and Brooks 2000). In addition to Europe, the Mousterian is found in the Caucasus, the Near East (where it is associated with both Neanderthals and early modern humans) and North Africa (where it is not associated with Neanderthals) (Klein 1999).

Mousterian industries appear in Europe as early as 200–150 ka and possibly earlier in the Near East, but most sites are dated to the interval from 130 to 30 ka. The lithic raw material used for the production of tools in most Mousterian sites tends to be available locally (Mellars et al. 1996). There is a lack of specialized use of different types of raw materials in the Mousterian, as well as a lack of specialized quarries. Very few bone tools are known. Some points appear to have been hafted and were probably used as spear points (Mellars et al. 1996). Wooden tools were probably also made, as is evidenced by several well preserved wooden spears discovered in Schoningen and dated to approximately 400 ka and by parts of similar implements from Clacto on Sea (possibly ca. 350 ka) and Lehringen (ca. 130–110 ka) (Mellars et al.1996; Facchini and Belcastro 2009).

Neanderthal sites show relatively little structure compared to later Upper Paleolithic sites. Hearths are well defined and were probably central in tool production and bone processing but are not consistent in their location (Mellars et al. 1996).

Until recent years, the Mousterian was commonly thought to represent a static culture. However, a re-dating of Mousterian sites has shown changes with time in regional industries from lithic raw material Europe and the Near East. Additionally, reanalysis of some Mousterian sites has shown technological responses to climatic changes (Shea and Brooks 2000; Facchini and Belcastro 2009). Some ‘‘transitional’’ Middle-Upper Paleolithic industries, like the Châstelperronian industry in France, the Uluzzian in Italy, and the Szeletian in East Central Europe, also show strong affinities However, a redating of Mousterian sites has shown changes with time in regional industries from lithic raw material Europe and the Near East. These were originally thought to have been made by early modern humans, also generally considered responsible for the Aurignacian industry.

Recently, however, the Châstelperronian has been found associated with Neanderthal skeletal remains in two sites in France, St. Césaire, and Arcy-sur-Cure, both dated to approximately 35 ka, suggesting that at least some of these transitional industries were produced by late Neanderthal populations (Facchini and Belcastro 2009). This discovery has prompted intense debate over the identity of the makers of these industries, the possibility of Neanderthal acculturation by, or trade with, early modern humans, and the cognitive capacities and ability for symbolic thought in Neanderthals.

In this regard, it is possible to recognize two different types of symbolism in Neanderthals:

1-Functional symbolism: it can 'be identified in the technology, organization of the territory and hunting. According to Wynn et al. (2002), it reveals a double-sided symmetrical shape thought. The organization of space at home or for hunting purpose can not be explained without a human activity cognitive.

2-Spiritual symbolism: demonstrations expressing abstract intelligence do not appear related to livelihood strategies (Hublin et al. 1996; Facchini and Belcastro 2009). Also, in two sites of the Neanderthal-associated Middle Paleolithic of Iberia, dated to as early as approximately 50,000 years ago, Zilhao et al. (2010) discovered perforated and pigment-stained marine shells. At Cueva de los Aviones, three umbo-perforated valves of Acanthocardia and Glycymeris were found alongside lumps of yellow and red colorants, and residues preserved inside a Spondylus shell (Figure 14) consist of a red lepidocrocite base mixed with ground, dark red-to-black fragments of hematite and pyrite. A perforated Pecten shell, (Figure 15) painted on its external, white side with an orange mix of goethite and hematite. Comparable early modern human-associated material from Africa and the Near East is widely accepted as evidence for body ornamentation, implying behavioural modernity. The Iberian finds show that European Neanderthals were no different from coeval Africans.

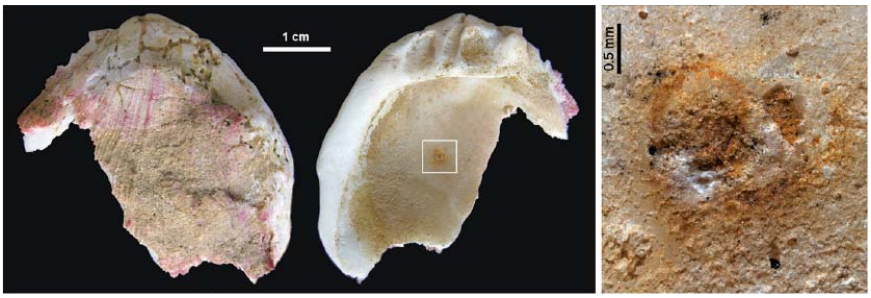


Figure 14: (Left) Excavation-broken upper valve of Spondylus gaederopus from Cueva de los Aviones (the pigment residue is indicated); (Right) close-up view of the pigment residue (Zilhão et al. 2010).

In this regard, countering genetic/cognitive explanations for the emergence of symbolism and strengthening demographic/social ones.

Figure 15: A perforated upper half-valve of Pecten maximus from level I-k at Cueva Antón. The internal, naturally red side (Left) and the external, whitish side that was painted with an orange colorant made of goethite and hematite (Right) (Zilhão et al. 2010).

1.1.4. Language

About the Neanderthal ability in the language there is no uniformity of views. In the opinion of Nobel and Davidson (1991) the origin of language would be between 70,000 and 100,000 years ago. Lieberman (1985) believes that the full development of language is a recent acquisition and the morphology of the Neanderthal larynx was not fit for that purpose. In favor of the hypothesis of language in Neanderthal there is the discovery of hyoid bone in Kebara fossil (Israel, 60 Ka) (Arensburg et al. 1989; 1990; Arensburg 1991) (Figure 16) and in El Sidron SDR-034 fossil (Spain, 38 Ka) (Martinez et al. 2003). Both Neanderthal specimens are modern-human-like in size and shape (Arensburg et al. 1989, Martinez et al. 2003; Martinez et al. 2008) (Figure 17) and based on the Kebara 2 hyoid, it was proposed that the position, form, size, and relationship of this bone with the Neanderthal larynx was also modern-human-like and that, consequently, these fossil humans had the capacity for spoken language (Arensburg et al. 1989; 1990; Martinez et al. 2008).

Other researchers have contested this conclusion, arguing that neither the dimensions nor the morphology of the hyoid bone were direct indicators of the position of the larynx in the throat (Lieberman et al. 1992). On the other hands, the hyoid bone presence in *Australopithecus. afarensis* indicates that the derived modern human morphology emerged at some point during the course of human evolutionary history (Martinez et al. 2008) so, the presence of a human-like hyoid in the Neanderthal specimens Kebara 2 and SDR-034 suggests that this derived condition was

also present in the last common ancestor of the Neanderthal and modern human evolutionary lineages (Martinez et al. 2008).

To try to shed light on this interesting aspect of the Neanderthals, molecular data can help to understanding the origin and evolution of language, in fact Krause et al. (2007) analyzed the evolutionary changes in *FOXP2*, a gene that has been implicated in the development of speech and language, shows that Neanderthals, share with modern humans two changes in this gene. In this case selective sweep started before the divergence of the ancestral populations of Neanderthals and modern humans around 300,000–400,000 years ago.

Even if Neanderthals could not pronounce the totally of the sounds, the complexity of their way of life required a language ability to evoke the hunting, gathering fruits, the diversity of the instruments used. Probably also among Neanderthals, language and culture are developed together (Facchini et al. 1996).

Figure 16 : Kebara hyoid bone (Arensburg et al. 1989, 1990).

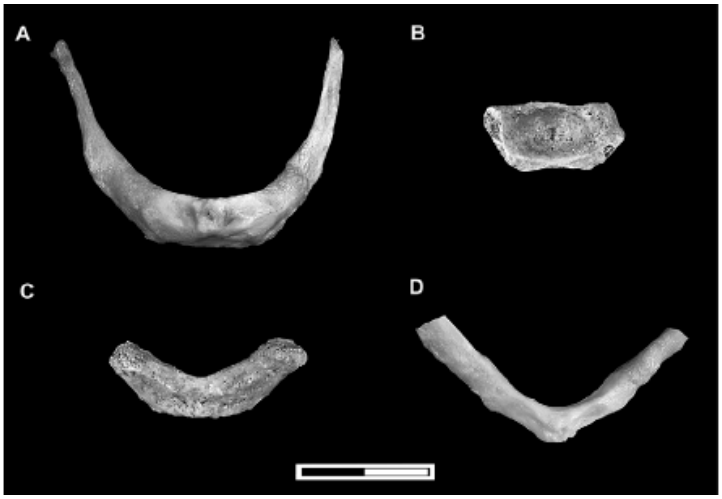


Figure 17: Anatomical variation in modern human hyoid bones. (A) Ventral view showing a well-developed anterior tubercle (Individual 1 from cadaveric collection). (B) Dorsal view showing a well-delimited round fossa (Individual 197 from San Pablo). (C) Superior view showing a V-shaped superior contour (Individual 64 from San Pablo). (D) Superior view showing a V-shaped superior contour (Individual 1 from cadaveric collection). Scale bar ¼ 2 cm (Martinez et al 2008).

1.2. Mitochondrial DNA

Deoxyribonucleic acid or DNA, is a nucleic acid that contains the genetic instructions used in the development and functioning of all known living organisms. The main role of DNA molecules is the long-term storage of information. The DNA segments that carry this genetic information are called genes. DNA consists of two long polymers of simple units called nucleotides, with backbones made of sugars and phosphate groups joined by ester bonds. These two strands run in opposite directions to each other and are therefore anti-parallel. Attached to each sugar is one of four types of molecules called bases. It is the sequence of these four bases along the backbone that encodes information. This information is read using the genetic code, which specifies the sequence of the amino acids within proteins. The code is read by copying stretches of DNA into the related nucleic acid RNA, in a process called transcription. Within cells, DNA is organized into long structures called chromosomes. These chromosomes are duplicated before cells divide, in a process called DNA replication. Eukaryotic organisms (animals, plants, fungi, and protists) store most of their DNA inside the cell nucleus. In contrast, prokaryotes (bacteria and archaea) store their DNA only in the cytoplasm. Within the chromosomes, chromatin proteins such as histones compact and organize DNA. These compact structures guide the interactions between DNA and other proteins, helping control which parts of the DNA are transcribed (Russel 1998).

Mitochondrial DNA (mtDNA) (Figure 18) is contained in the mitochondria, cytoplasm organelles members of the process of cellular respiration. In 1981, Anderson and colleagues at the MRC Laboratory of Molecular Biology in Cambridge, obtained the first complete sequence of human mitochondrial DNA. This sequence is now universally used by convention as a reference sequence (Cambridge Reference Sequence-CRS) (Anderson et al. 1981). Human mitochondrial DNA is a circular molecule containing 16,569 base pairs. The two strands that constitute it are called H (heavy = heavy) and L (light = light) for their different nucleotide constitution: the H strand contains primarily purine bases A and G (of higher molecular weight), the filament bases L pyrimidine bases C and T (smaller molecular weight). This results in a different separation of the filaments in a gradient of cesium chloride. Based on the functions of the mitochondrial DNA molecule can be divided into two main regions: the coding region and non-coding control region. The coding region, which represents the largest fraction, contains several genes for RNA and 13 genes coding for proteins involved in the process of cellular respiration. The control region is about 1100 bp long and its interior can be divided into two sections, called hypervariable region 1 and 2 (hyper-variable region I and II, and HVR HVR-I-II), which are characterized by a high rate mutation. It plays a regulatory function and contains the sites of initiation of heavy strand replication and transcription promoters of both strands. The region is also called D-loop (displacement loop) for the presence of a dormant stretch of RNA that prevents the coupling of the complementary filaments (Russel 1998)

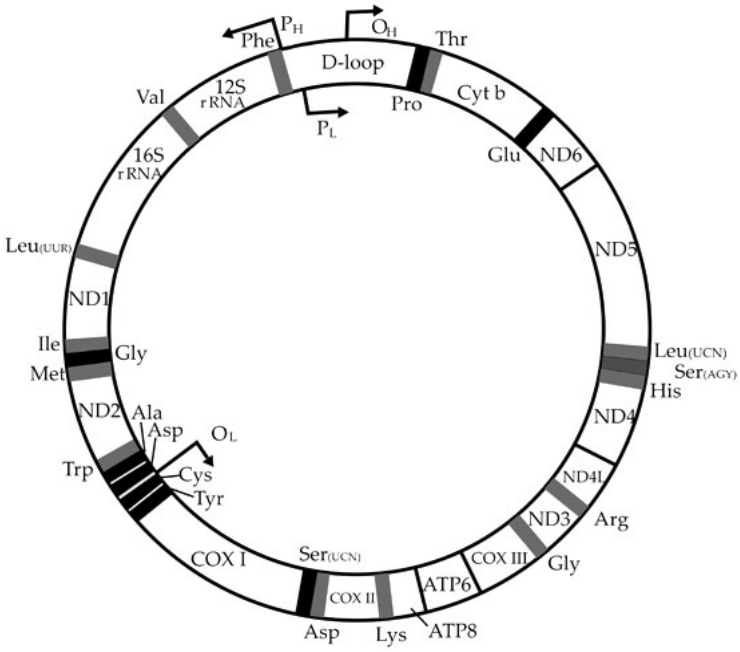


Figure 18: Human Mitochondrial DNA. White part represent genes coding proteins and ribosomial RNA; Black and grey part represent genes coding Transfer RNA (*www.mitomap.org)*

The most interesting features of mitochondrial DNA for the study of ancient samples are: the maternal inheritance and the consequent absence of recombination, the high rate of mutation, the presence of a number of multiple copies of the mitochondrial genome in each cell. Unlike the nuclear genome, which has a type of biparental inheritance in which both parents contribute equally to represent the genetic heritage of the offspring, the inheritance of the mitochondrial genome is uniparental maternal type. This occurs because the cytoplasm of the zygote, which is contained within the mitochondria, comes almost exclusively from the egg and the sperm's contribution is almost irrelevant. According to this non-

Mendelian mechanism of inheritance, mitochondrial DNA molecules do not recombine with each other (Giles et al. 1980).

Another important feature of the mitochondrial genome is the high frequency of mutations, consisting mainly of nucleotide substitutions and length mutations (deletions and insertions). A high mutation rate is proportional to a high rate of evolution. The rate of nucleotide substitution in mitochondrial DNA is 5 to 10 times higher than that of single-copy nuclear genes there are approximately 10 * 10-9 nucleotide substitutions per nucleotide site per year (Brown et al. 1979). The mutation rate is not constant throughout the sequence of mitochondrial DNA. The region of the D-loop evolves 10 times faster than the entire mitochondrial genome (Greenberg et al. 1983); also the hypervariable region I contains a variable twice the hyper-variable region II (Vigilant et al 1990). The ratio of transitions (ie purine-purine substitutions or pyrimidine-pyrimidine) and transversions (ie substitutions purine-pyrimidine) is quite high, resulting in an average of 5 to 1, with some variations depending on the region (2.5:1 in coding region, Cann et al. 1984; 24:1 in non-coding region, Greenberg et al. 1983).

Among the factors responsible for the rapid evolution of mammalian mitochondrial DNA, a primary role is attributed to the non efficiency of the repair system faults (Cann et al. 1984; Wilson et al. 1985) with the addition of a high incidence of oxidative damage, due to the high flux of radicals within the mitochondrial respiratory chain (Wilson et al. 1985).

Another important factor is considered the relaxation of the mechanisms that regulate protein translation apparatus that is tied to the type of proteins that are encoded by mitochondrial DNA. The mitochondrial genome, in fact, does not code for proteins directly involved in the replication, transcription and translation of the molecule and is then able to tolerate a lower accuracy in the translation.

The peculiarities of the mitochondrial DNA so far described are very useful for investigating the ancestor descendant relationships between species or populations because using it as a molecular marker it should not be taken into account the phenomena of recombination and segregation affecting nuclear genes, and its high mutation rate, also makes it particularly suitable for studies on human populations that are characterized by short evolution time. Finally, another feature that made the mitochondrial DNA molecular marker further investigated in studies of ancient DNA is undoubtedly the large number of copies in each cell. This feature greatly increases the probability of finding, even in highly degraded samples such as those derived from ancient remains, a sufficient number of copies to be able to perform the analysis.

1.3. History of ancient DNA

In the last decades, DNA analysis has become very versatile and accessible to a wider community of researchers. Recent technological developments in the study of DNA have provided a new dimension in search of our origins. In fact beside the study of bone remains and lithic industries, it is now possible to analyze the DNA of our ancestors.

The term Ancient DNA (aDNA) refers to DNA from a dead body or from ancient skeletal remains and can therefore be held to consider all traces of ancient DNA subjected to autolysis or degradation process (Rollo 1999).

The history of ancient DNA starts twenty-seven years ago when Higuchi et al.(1984) extracted and sequenced, by means of molecular cloning, DNA fragments from a museum specimen of the quagga (*Equus quagga quagga,* an equid from South Africa that became extinct in the XIX century). This finding revolutionised the field of molecular biology because it showed that it was possible to retrieve DNA from an organism from the distant past. The improvement of aDNA search is associated with the advent of Polymerase Chain Reaction (PCR) developed by Karry Mullis (Saiki et al. 1985; Mullis and Faloona 1987). This method is useful to amplify a single or a few copies of a fragment of DNA across several orders of magnitude, generating thousands to millions of copies of a particular DNA sequence. Bone and teeth were quickly found to be better sources of aDNA than mummified soft tissues (Hagelberg et al. 1989) and this meant that museums suddenly

became recognized as storehouses of preserved genetic information from the past.

Several studies have revealed the phylogenetic relationships of extinct animals, generally on the basis of mtDNA, such as Quagga (Higuchi et al. 1984), Mammuth (Hagelberg et al. 1994; Hoss et al. 1994; Yang et al. 1996; Ozawa et al. 1997; Noro et al. 1998), Moas (Cooper et al. 1992), Cave bear (Hanni et al. 1994) and *Myotragus balearicus* (Lalueza-Fox et al. 2000; Ramirez et al. 2009). These and other studies in progress show that it is now possible to begin to reconstruct the population history of extinct animals, going back in time to well before the last Ice Age (Hofreiter et al. 2001).

Since the discovery of PCR, the list of publications related to aDNA increased exponentially. Some journals started publishing articles related to aDNA from specimens that were millions of years old which could be successfully extract and sequenced (Golenberg et al. 1990; de Salle et al. 1992; Soltis et al. 1992; Cano et al 1993; Cano and Borucki 1995; Vreeland et al. 2000; Fish et al. 2002; Willerslev et al.2004a). Examples of this period are the retrieval of DNA sequences surviving in amber inclusion (de Salle et al 1992; Cano et al 1993), plant leaves dated from Miocene (Golenberg et al. 1990; Soltis et al. 1992), the retrieval of *cytocrome b* mitochondrial gene from a dinosaur bone dated to over 80 million years ago (Woodward et al 1994) and the recovery of bacterial DNA and cells from amber and halite that are claimed to be many millions of years old (Cano and Borucki 1995; Vreeland et al. 2000; Fish et al. 2002; Willerslev et al. 2004a). Later, due to numerous technical problems, all these extraordinary discoveries turned out

to be false and severely criticised (Pääbo and Wilson 1991; Lindhal 1993; Gibbons et al.1994).

One of the most interesting achievements in the aDNA field was produced by the team supervised by Svante Pääbo who recovered 380 bp of the hipervariable region mtDNA from a Neanderthal specimen (Krings et al. 1997). This result has been subsequently corroborated by more mtDNA sequences from the same individual (Krings et al. 1999) and from sequences of other Neanderthal specimens spread all over Europe (Krings at al. 2000; Ovchinnikov et al. 2000; Schmitz et al. 2002; Serre et al. 2004; Beauval et al 2005; Lalueza-Fox et al 2005; Caramelli et al 2006; Lalueza-Fox et al 2006; Orlando et al. 2006; Briggs et al. 2009; Lalueza-Fox et al. 2011).

However, ancient DNA studies are also subject to problems that are peculiar to the field, which manifest themselves as difficulties in generating sufficient authentic DNA sequences to make a study conclusive. These arise as a result of the post mortem degradation of DNA, either through the generation of miscoding lesions, which can lead to sequence errors, or through the physical destruction of the molecule, which increases the risk of preferentially amplifying an exogenous contaminant sequence. The field has predominantly attempted to deal with these problems through the publication of guidelines that are designed to ensure the quality of ancient DNA data and conclusions. Starting from a few relatively simple suggestions as, the use of negative controls; analyses of multiple extracts per sample; the observation of an inverse correlation between amplification efficiency and size of the amplification product, reflecting the degradation and damage in the ancient DNA

template (Pääbo 1989). The guidelines have evolved over time into a more detailed and extensive list of requirements, resulting in the nine key criteria of Cooper and Poinar (appendix A). In particular, these authors argue that, in the absence of full compliance to all nine criteria, the reliability and authenticity of results are uncertain. These criteria were intended to assist in determining the authenticity of a study (Gilbert at al. 2005).

Other important applications of aDNA have been the study of the evolution of current pathogens extracted from infected human bones, such the bacteria *Mycobacterium tubercolosis* (Mays et al. 2002; Zink et al. 2003; 2005; Matheson et al. 2009) and *Yersinia pestis* (Drancourt et al.1998; Drancourt et al. 2007; Bianucci et al. 2008).

Ancient DNA studies can also contribute to the understanding of more recent history, in what can be called “historical forensic” In fact recently some works have tried to shed new light on historical mysteries. These include the analysis of the remains of the Romanov family (Gill et al. 1994; Coble et al. 2009), the putative evangelist Luke (Vernesi et al. 2001), the American outlaw Jesse James (Stone et al. 2001), the heart of Louis XVII, (Jehaes et al. 2001), the Italian poet Francesco Petrarca (Caramelli et al. 2007), the astronomer Nicolaus Copernicus (Bogdanowicz et al. 2009) and a piece of a handkerchief with the blood of Louis XVI, stored in a pumpkin (Lalueza-Fox et al. 2010).

1.3.1. Neanderthal and Denisova genomes: the future of molecular anthropology

Neanderthals are the sister group of all present-day humans. Thus, comparisons of the human genome to the genomes of Neanderthals and apes allow identify the changes that have become fixed or have risen to high frequency in modern humans during the last few hundred thousand years and should be informative for identifying genes affected by positive selection since humans diverged from Neanderthals (Green et al. 2010).

The question of whether Neanderthals interbred with anatomically modern humans is a controversial point (Green et al. 2010). Researcher are not agree on the information from the morphological analysis (Trinkaus et al. 2003; Bailey et al. 2009) The part of the genome that has been examined from multiple Neanderthals, the mitochondrial DNA (mtDNA) genome, consistently falls outside the variation found in present-day humans and thus provides no evidence for interbreeding (Krings et al. 1997, Ovchinnikov et al. 2000; Orlando et al. 2006; Briggs et al. 2009; Lalueza-Fox et al. 2011). However, this observation does not preclude some amount of interbreeding between them (Serre et al. 2004).

Instead, the nuclear genome, thanks to its recombining capacity, is able to clarify the picture of the relationship between Neanderthals and present-day humans (Green et al. 2010). In the past years four short gene sequences have been determined by PCR (Krause et al. 2007; Lalueza-Fox et al. 2007; 2008; 2009).

Recently Green et al. (2010), through the use of high-throughput DNA sequencing technologies, have retrieved a large part of Neanderthal nuclear genome. The protocol used is very complex. In the first phase they analyzed 21 Neanderthal bones from Vindija Cave (Croatia). All samples were screened for the presence of Neanderthal mtDNA by PCR and Vi33.16, Vi33.25 and Vi33.26, were selected for further analysis. A total of nine DNA extracts were used to construct Roche/454 sequencing libraries that carry the project-specific tag sequence 5′-TGAC-3′ in their 3′-ends. The second phase was estimate the percentage of endogenous Neanderthal DNA in the extracts. They carried out sequencing of Vi33,16 and Vi33,26 runs using the 454 Life Sciences GS FLX platform and mapped the reads against the human, chimpanzee, rhesus, and mouse genomes as well as all nucleotide sequences in GenBank. DNA sequences with a significantly better match to the primate genomes than to any of the other sources of sequences were further analyzed. Taking into account that between 95 and 99% of the DNA sequenced in the libraries was derived from presumably microbes to improve the ratio of Neanderthal to microbial DNA they identified restriction enzymes that preferentially cut bacterial DNA sequences in the libraries. In the third phase they carried out production sequencing on the Illumina/Solexa GAII platform from the bone Vi33,16, Vi33,25 and Vi33,26 (1.2 Gb, 1.3 Gb and 1.5 Gb respectively). Each molecule was sequenced from both ends and bases were called with the machine learning algorithm Ibis (Kircher et al. 2009).

In total they generated, with 1.3-fold coverage, 5.3 Gb of Neanderthal DNA sequence from about 400 mg of bone powder. After the carry on vary type of analyses focusing mainly on the sequences that house amino-acid changes that are foxed in present-day humans but ancestral in Neanderthal.

Results revealed that Neanderthal DNA sequences and those of present-day humans share common ancestors on average about 800,000 years ago and that the population split of Neanderthal and modern human ancestors occurred 270,000–440,000 years ago. It also showed that Neanderthals shared more genetic variants with present-day humans in Eurasia than with present-day humans in sub-Saharan Africa, indicating that gene flow from Neanderthals into the ancestors of non-Africans occurred to an extent that 1–4% (2.5% of average) of the genomes of people outside Africa are derived from Neanderthals (Figure 19). Also point to a number of genomic regions and genes as candidates for positive selection early in modern human history, for example, those involved in cognitive abilities and cranial morphology.

Recently, another important finding was made and it shed new light on the evolutionary history of genus *Homo*. Indeed Krause et al. (2010) described the distal manual phalanx of a juvenile hominine excavated at Denisova Cave (Altai Mountains, southern Siberia). The phalanx was found in layer 11, which has been dated to 50,000 to 30,000 years ago. This layer contains micro blades and body ornaments of polished stone typical of the ‘Upper Palaeolithic industry’ generally thought to be associated with modern humans, but also stone tools that are more characteristic of the earlier

Middle Palaeolithic, such as side scrapers and Levallois blanks (Derevianko et al. 2008).

Capture approach (Briggs et al. 2009; Kircher et al. 2009) in combination with high-throughput sequencing (Marguiles et al 2005; Kircher et al. 2009) were used to determine a complete mtDNA genome from the Denisova phalanx.

They generated 1,178,300 merged sequences, of which 93,349 (7.9%) aligned to the revised Cambridge Reference Sequence (rCRS) (Andrew et al.1999) using an iterative mapping assembler, MIA (Briggs et al. 2009). Fragments with identical start and end coordinates were merged to single sequences where at each position the base with the highest quality score was used. A total of 30,443 such distinct sequences were then used to assemble mtDNA sequence. Coverage across the mtDNA was high (mean 5156-fold, lowest 2-fold, highest 602-fold).

Surprisingly, this mtDNA diverged from the common lineage leading to modern human and Neanderthal mtDNAs about one million of years ago (Krause et al. 2010) and it shows that it was distinct from *H. erectus* that first left Africa 1.9 Myr ago, and also from the taxon *H. heidelbergensis*, if the latter is the direct ancestor of Neanderthal (Hublin 2009).

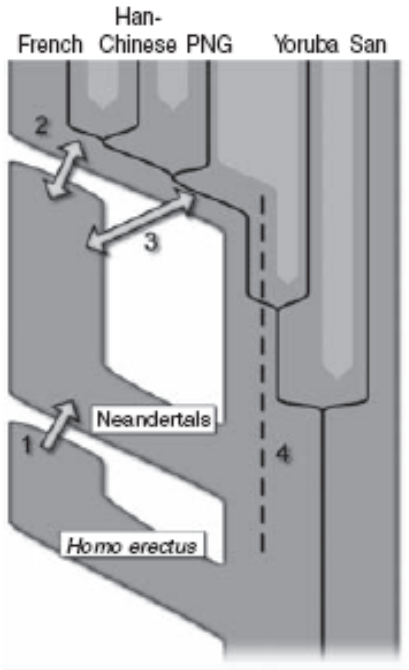


Figure 19: Four possible scenarios of genetic mixture involving Neanderthals. **Scenario 1** represents gene flow into Neanderthal from other archaic hominins, here collectively referred to as *Homo erectus*. This would manifest itself as segments of the Neanderthal genome with unexpectedly high divergence from present-day humans. **Scenario 2** represents gene flow between late Neanderthals and early modern humans in Europe and/or western Asia. No evidence of this because Neanderthals are equally distantly related to all non-Africans. However, such gene flow may have taken place without leaving traces in the present-day gene pool. **Scenario 3** represents gene flow between Neanderthals and the ancestors of all non-Africans. This is the most parsimonious explanation of our observation. Although was detect gene flow only from Neanderthals into modern humans, gene flow in the reverse direction may also have occurred. **Scenario 4** represents old substructure in Africa that persisted from the origin of Neanderthals until the ancestors of non-Africans left Africa. This scenario is also compatible with the current data (Green et al. 2010).

To clarify the relationship of the Denisova individual to other hominines groups, Reich et al. (2010) sequenced the Denisova nuclear genome and analyzed its genomic relationships to Neanderthals and presents-day humans.

Two independent sequencing libraries were created from the DNA, using a modified Illumina protocol. The libraries were sequenced on the Illumina Genome Analyser IIx platform (Kircher et al. 2009) and bases were called using the machine-learning algorithm Ibis (Kircher et al. 2009). A total of 82,227,320 sequences mapped uniquely (mapping quality>30) to the human genome, yielding about 5.2 gigabases of DNA sequences (1.9-fold genomic coverage), and 72,304,848 sequences mapped uniquely to the chimpanzee genome. When the substitutions inferred to have occurred on the Denisova and the present-day human lineages were compared, the relative numbers of different classes of nucleotide substitutions are remarkably similar, and the excess number of candidate substitutions on the Denisova lineage relative to the present-day human lineage is only 1.7-fold.

The results show that the Denisova individual belongs to a hominine group, called “Denisovans” that shares a common ancestor with Neanderthals but had a distinct population history to both Neanderthal and modern human (Figure 20). Perhaps Denisovans, but not Neanderthals, contributed genes to ancestor of present-day Melanesians but not to present-day populations which currently live much closer to Altai region such as Han Chinese or Mongolians and this leave open the interesting question about how widespread Denisova were. One of the major importance of this study is that for the first time a population,

“Denisovans” is characterized primarily by its DNA. In fact, at present, excluding the phalanx and a tooth retrieved during the archaeological excavation, no other physical characteristics of this population are known.

With the new sequencing technologies, complete genome of extinct hominins are now available, within the time frame of DNA preservation, thus opening new prospection of the study of human evolution.

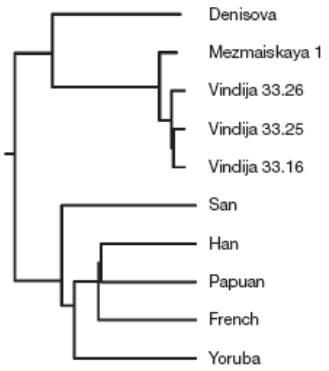


Figure 20: A neighbour-joining tree based on pairwise autosomal DNA sequence divergences for five ancient and five present-day hominins. (Reich et al. 2010).

1.4. Ancient DNA problems

1.4.1. Ancient DNA damage

DNA is a polymer. The monomer units of DNA are nucleotides, and the polymer is known as a "polynucleotide." Each nucleotide consists of a 5-carbon sugar (deoxyribose), a nitrogen containing base attached to the sugar, and a phosphate group. There are four different types of nucleotides found in DNA, differing only in the nitrogenous base, adenine (A), guanine (G), cytosine (C) and thymine (T) linked to phosphorilated sugar by means of glycosilic bonds. The phosphorelated sugars are 2' deoxirribose units linked to each other by means of phosphodiester bonds shaping like a double helix The sugar and phosphate backbone are on the outside of the helix, whereas the bases are placed inside the DNA molecule (Caramelli and Lari 2004). However, the DNA is unstable molecules and as consequence it is labile and prone to many forms of damage. Mainly two types of damage are likely to affect the DNA.

When the chemical bond between a DNA base and its respective deoxyribose, is subject to a cleavage by a water molecule, the process is known as spontaneous **hydrolysis** (Figure 21). The DNA is prone to hydrolytic damage due to some reasons. First, the removal of the ribose sugar's 2'-OH group in ribose units, makes the phosphodiester bond quite labile and subject to quick hydrolytic cleavage geneting single stranded nick in the double helix. It has been estimated that direct cleavage of the phosphate backbone is probably the most frequent type of hydrolytic damage the DNA

(Lindhal et al. 1993, Poinar et al. 2002). Second, the glycosidic bond is also subject to direct hydrolytic attack. Base protonation will cause the cleavage of the glycosidic bond, termed depurination, and form what is known as a baseless or an apurinic site (AP site). The rate of depurination is increased and dependent upon the temperature, the ionic strength, the pH, and heavy metal ion chelation (Poinar et al. 2002). Third, DNA bases with secondary amino groups such as adenine, cytosine, 5-methylcytosine, and guanine can undergo deamination, (the hydrolytic cleavage of their amino groups), resulting in hypoxanthine, uracil, thymine, and xanthine, respectively. Deamination has recently been shown to be a prominent component of some fossil DNA remains (Hofreiter et al 2001, Poinar et al. 2002).

Another important type of DNA damage is the **oxidative** (Figure 21) damage that occurs through the action of free radicals such as peroxide radicals (O_2), hydrogen peroxide (H_2O_2) and hydroxy radicals (OH) (Lindhal 1993) and are likely to play an important role in limiting the life of DNA in the fossil record. These radicals are endogenously generated by the cell but they may also derive from exogenous sources such as ionising radiation, UV light that cause also pyrimidine dimmers in the DNA double helix and cellular process during bacterial and fungal degradation (Poinar et al. 2002).

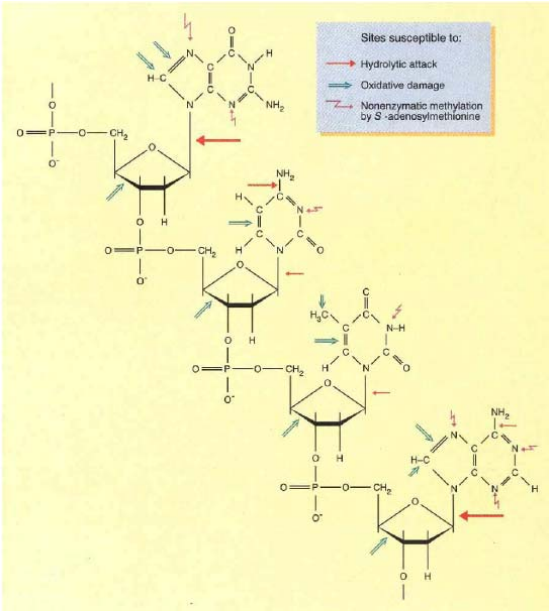


Figure 21: Sites susceptible to hydrolytic attack are indicated by solid red arrows, oxidative damage by open green arrows and non enzymatic methylation by S. adenosylmethionine as zig-zagged purple arrows. Hydrolytic and oxidative damage, but not methylatec residues would accumulate in fossil specimens DNA (Lindahl et al. 1993).

Traces of genetic material preserved within ancient specimens can provide a unique and important real-time record of the past.

However, this record is compromised because ancient DNA is invariably damaged and degraded. Hydrolytic and oxidative reactions can fragment the DNA backbone producing short fragment length, thus resulting in much shorter PCR products than those obtained from modern DNA. hydrolysis, can also produce a depirination after cleavage of phospodiester backbone causing the formation of abases sites and can induce deamination. Deamination is the main cause of miscoding lesion in ancient DNA

(Briggs et al. 2007). Infact deamination of cytosine and adenine to uracile and hypoxanthine, respectively, will result in the incorporation of erroneous bases during the amplification (Pääbo 1989; Lindhal 1993; Willerslev and Cooper 2005).

Also a series of factors that determine the degradation of DNA and then the appearance of damage exsist:

Microbial activity: After death, the organism is decomposed by microbial activity (Rollo 1999; Caramelli and Lari 2004)

pH: Some studies showed that in an acid environment, DNA is subjected to a more rapid deamination (Mitchell et al. 2005) while the alkaline pH accelerates the breakdown of the molecule at the level of apurinic and apirimidinic sites (Lindhal 1993).

Temperature: low temperature inhibits the enzymatic activity and this benefits the conservation of DNA. Some example are, the mummy of Similaun (Frozen Frizt or Otzi) case, dated 3300 years BP preserved thanks to the special climatic condition in the glacier (Handt et al. 1994; Rollo et al. 1994;1995;1999), the retrieval from permafrost setting of a specimen of mammoth (Hoss et al 1994; Gilbert et al 2007; Poinar et al. 2006), bison (Gilbert et al 2004a; Shapiro et al 2004) and old plant chloroplast (cpDNA) (Willerslev et al 2003a, 2004a) In contrast the high temperature impairs the Van der Waals forces and hydrogen bonds and also promotes the growth of bacteria, mould and other microorganisms that may contaminate the DNA (Rollo 1999; Carmelli and Lari 2004; Handt et al 1994; Gilbert et al 2007).

UV: the UV light can fragment the nucleid acid creating pirimidinic dimmers (Lindhal 1993; Rollo1999; Caramelli and Lari 2004).

1.4.2. Inhibitors

The advent of PCR has greatly aided investigations in ancient DNA field. Infact minute amounts of short DNA fragments retrieved in fossil specimens can be selectively amplified (Pääbo et al. 1989). One of the problems with hypersensitive PCR technology, is that several factors can block the *Taq polimeras*i activity, preventing the DNA amplification (Pääbo 1989, Hanni 1995, Poinar 2002).

Inhibitory substances come from an external source, rather than being a component of the sample (Scholzet al. 1998; Poinar et al; 2002).These can be: Maillard reaction products (reaction of carbonil groups on reducing sugar such as those that belong to the backbone skeletal of DNA double helix) (Sholz et al. 1998; Poinar et al. 2002; Willerslev and Cooper 2005), some types of oxidative damage that may block the progression of *Taq* polymerase (Mitchell et al. 2005; Sholz et al.1998), non target DNA mainly derived from bacteria (Pääbo 1989), and soil components like fulvic and humic acids, tannins, or complexing ions like iron (Fe_{21}) (Pääbo 1989; Hagelberg et al. 1991; Poinar et al 1992; Scholz et al 1998).

Various approaches have been employed to overcome inhibition of PCR, including, the use of molecules, such as BSA (Bovine Serum Albumin) that binds the inhibitors; making serial dilutions of the

extract until the dilution for which the PCR is successful is found, PTB (N-phenacyltiazolium bromide) To try to remove the cross link caused by Maillard reactions, allow DNA sequences to be amplified from ancient remains that otherwise could not be amplified (Poinar et al. 2002).

1.4.3. Miscoding lesion

DNA decays rapidly after death in biological samples and the ensuing damage is manifested in many forms (See chapter 1.4.1) but a small proportion of damage events do not hinder replication but generate miscoding lesions (Pääbo 1989). These are manifested as base modifications in the amplified sequence, changing the appearance of a DNA template and potentially generating misleading haplotype analyses (Gilbert et al. 2003a).

Several studies (Pääbo 1989; Hofreither et al. 2001; Gilbert et al. 2003a; 2003b) underline that the majority of changes arise from the deamination of C to uracil (U), an analogue of T, or the deamination of A to hypoxanthine (HX), an analogue of G (Lindhal 1993; Hansen et al. 2001; Hofreiter et al. 2001; Gilbert et al. 2003b). However, because either of the complementary DNA strands can be sequenced after amplification, when using Sanger technology, each of these transitions can produce two observable phenotypes. The C to T degradation may simply be observed as C to T, but, if the complementary strand is sequenced, then it will be read as a G to A transition. Similarly, an A to G degradation may be observed as either A to G or as a T to C transition (Hansen et al. 2001; Hofreiter et al. 2001; Gilbert et al. 2003b). Hansen et al.

(2001) called each set of miscoding lesion as “type 1” (A to G/T to C) or “type 2” (C to T/G to A) transitions, respectively (Figure 22).

The recent development of the sequencing-by-synthesis technology offer the possibility of going deeper into the nature of those miscoding lesions (Gilbert et al 2007) The key benefit of this approach is that each final sequence reaction is generated from a single single-stranded DNA molecule opening a window into any damage derived miscoding lesions (Margulies et al. 2005).

Gilbert et al (2007) demonstrated that type 2 damage (C to T / G to A) represent the majority of damage derived from miscoding lesions. Also they argued that underpinning a significant portion of Type 2 damage there is a guanine-adenine analogue modification. Brotherton et al. (2007) using a new technique called SPEX,

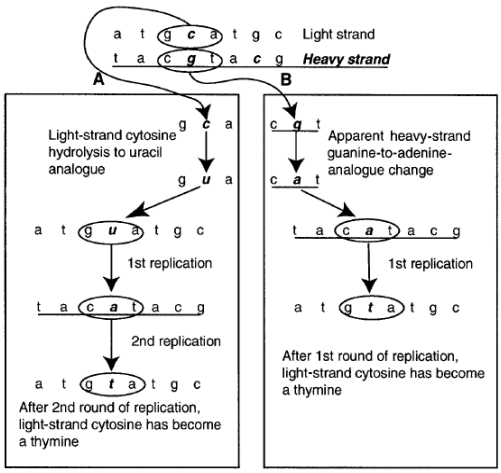


Figure 22: Determination of a strand of origin for post-mortem-DNA-damage events by using type 2 (C to T/G to A) transitions as an example. *A,* L-strand C to T transitions after two cycles of amplifications, resulting in a permanent L-strand change. *B,* A theoretical H-strand G to A change, producing the L-strand phenotype of C to T change following one cycle of amplification. However, since a direct G to A post-mortem modification is chemically impossible, the example depicted in this panel is not possible. Thus, all Cr to T changes observed on the L strand must have occurred as L-strand C to T post-mortem damage, and all G to A changes on the L strand must have occurred as H-strand C to T post-mortem damage (Gilbert et al. 2003).

and in contrast with previous aDNA studies, (Pääbo 1989; Hofreiter et al. 2001; Gilbert et al. 2003a; 2003b; Binladen et al. 2006) provided strong quantitative evidence that C - U type base modifications are the sole cause of authentic endogenous damage derived miscoding lesions. The Type 1 damage represents non-endogenous PCR-generated sequence artefacts and it disappears once alternative techniques (eg. Digestion with uracil-N-glycosylase) are employed (Stiller et al. 2006; Gilbert et al. 2007; Brotherton et al. 2007).

Another important consideration is that miscoding lesions are not randomly distributed across the mtDNA genome but are concentrated in *hot spots* where certain nucleotide positions appear to mutate at significantly higher rates than others (Gilbert et al. 2003a; Willerslev and Cooper 2005). This could be associated to the clustering of several C in particular regions (Vives et al 2008).

1.4.4. Fragmentation

The most obvious type of post-mortem damage is its fragmentation into small sequences <500 bp (Pääbo 1989) That type of degradation is due to the action of endonucleases that occur shortly after death or to the action of micro organisms and the hydrolytic or oxidative damage (Lindhal 1993) that disrupt the double helix causing the fragmentations and the consequent short PCR products.

The advent of next generation sequences (Margulies et al. 2005; Blow et al. 2008) and new system of captures (Briggs et al. 2009; Burbano et al. 2010; Maricic et al. 2010), has shown that more than half of DNA fragments extracted from, ancient remains, are of a size less than 50 nt and therefore not retrievable by PCR (Figure 23).

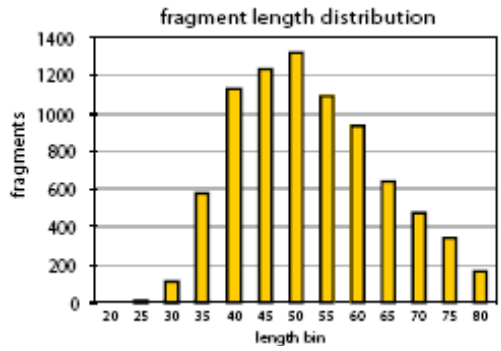


Figure 23: Length distributions fragments of endogenous DNA (Krause et al. 2010)

1.4.5. Contamination

Contamination is one if the most common and important problems in ancient DNA studies, especially those affecting ancient Human DNA works. As the retrieval of tiny amounts of DNA from ancient remains is a multi-step process, contaminants can enter in multiple stages:

1-Sample Handling: This type of contamination is extremely important in the study of ancient human remains, because usually it cannot be adequately monitored or controlled. Unprotected handling of remains may impregnate the samples with the handler's sweat or skin cells, and so exogenous DNA could penetrate into remains (Gilbert et al. 2005a). Bone and teeth remains are extremely porous and probably very susceptible to

contamination by handling. (Gilbert et al. 2005a). It is possible to distinguish different potential sources of handling contaminants:

a- **Archaeologist and anthropologist**: they are responsible for recovering, washing and macroscopically analysing the fossil remains and interact directly with the fossil material (Sampietro et al. 2006).

b- **Molecular anthropologist**: When the sample arrives in the DNA laboratory the molecular anthropologist team must decontaminate the sample using bleach or UV light before starting the DNA extraction. In this step it is compulsory to use gloves, face mask and coverall inside the aDNA laboratory in order to avoid contamination by handling (Caramelli and Lari 2004).

2- **Extraction Procedures**: Contaminants can be introduced in the aDNA extract during the preparation of all the reagents that are going to be used in the extraction process as well as in the place where the extraction is going to be performed. In order to avoid this kind of contaminations, it is very important that the place where the aDNA extraction is carried out is physically separated from the main molecular laboratory with positive air pressure, UV light at night and continuous bleach cleaning of the bench surface. In addition, the manipulation of all reagents must be done in a flow cabinet and the use of coverall, gloves, facemask and sterile filtered tips is compulsory (Cooper and Poinar 2000).

3- **PCR setting up**: Contaminants can be introduced at this stage by two main ways:

First: due to the very low amount of endogenous DNA, often, unspecific conditions (high number of cycles and low temperature) in the PCR are used. For this reason, it could be that the primers can also bind to exogenous sequences that are present in the extract and therefore amplify them.

Second: contaminants can be introduced while researchers are setting up the PCR reaction. As in the extract step, contaminants can be in the used reagents and in the environment where researchers are working. All the precautions described in the previous step are compulsory here as well. The setting up of the PCR reaction must be done in the ancient laboratory and then carried to the main laboratory. Material interchange between the aDNA laboratory and the main laboratory should be strictly forbidden because the risk of contamination is extremely high due to large amounts of amplicons that are generated in the molecular laboratory. After going into the main lab never go back to aDNA lab the same day. Furthermore, a study of Champlot et al. (2010) developed a new method to eliminate contaminating DNA contained in PCR reagents. It consists in a combination of treatment (Y-and UV-irradiation and mutant recombinant heat-labile double-strand specific DNAse from Antarctic shrimp *Pandalus borealis)* adapted to different reagent categories. Another way to minimize this problem is described in Krause et al. (2007) where the researcher and colleagues used a Shrimp Nuclease in order to degrade double-stranded DNA that may contaminate reagents.

1.5. aDNA Authenticity

Ancient DNA studies are subject to problems that are peculiar to the field which manifest themselves as difficulties in generating sufficient authentic DNA (Gilbert et al. 2005). Cooper and Poinar (2000) released guidelines that are designed to ensure the quality of ancient DNA data and conclusions. It is a more detailed and extensive list of requirements that, in nine criteria, summarizes the characteristic that an aDNA study must produces reliable and authentic results (see appendix A) (Gilbert et al. 2005). However, if the samples were contaminated before the analysis, the criteria provide no way to spot the human contaminants (Gilbert et al. 2005). To overcome this problem other criteria were suggested.

Some studies (Pääbo 1989; Malmstrong et al. 2007; Green et al. 2007; Krause et al. 2010), have underlined that authentic aDNA showed a more rapid increase in yield with decreased fragment size than contaminating DNA and the cause of this reduction in size is the post-mortem damage, characteristic of ancient DNA study (see chapter 1.4.1.) (Figure 24).

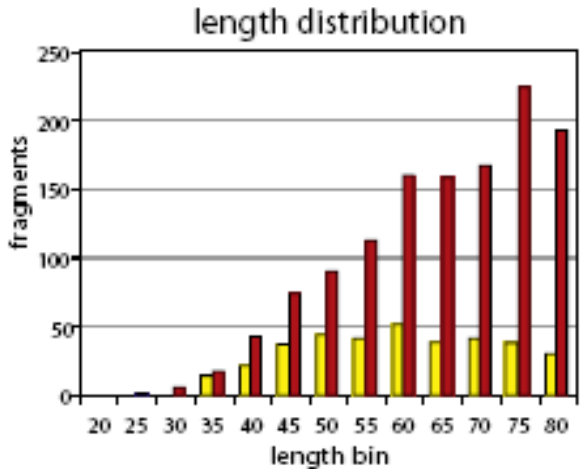


Figure 24: Length Distribution of Neanderthal DNA (in yellow) and Contaminating DNA (in red) (Krause et al. 2010).

Recently, high-throughput sequencing techniques have become available and allow large numbers of DNA sequences to be determined (Margulies et al. 2005; Bentley et al. 2008). The advent of this new approach to ancient DNA analysis makes it important to revisit the criteria of authenticity (Green et al. 2009).

Briggs et al. (2007) investigated the DNA sequence context around strand breaks in ancient DNA. This has not been previously possible, because using PCR technique, primers that target particular DNA sequences are generally used, and thus the ends of the ancient DNA molecules are not revealed. The 454 sequencing process are of crucial importance for this analyzes. The high frequency of C to T misincorporations at the 5'-ends of ancient DNA sequences and the correspondingly high frequency of G to A misincorporations at the 3'-ends imply that deamination of cytosine residues is significantly elevated at the 5'-ends of ancient DNA molecules (Figure 25). This could be caused either by a tendency

of cytosine residues at the ends of molecules to undergo deamination or a tendency of strand breaks to occur near deaminated cytosine residues. Therefore, the author proposes that cytosine residues close to the ends of ancient DNA molecules are more susceptible to deamination than cytosine residues more internal in the molecule (Briggs et al. 2007).

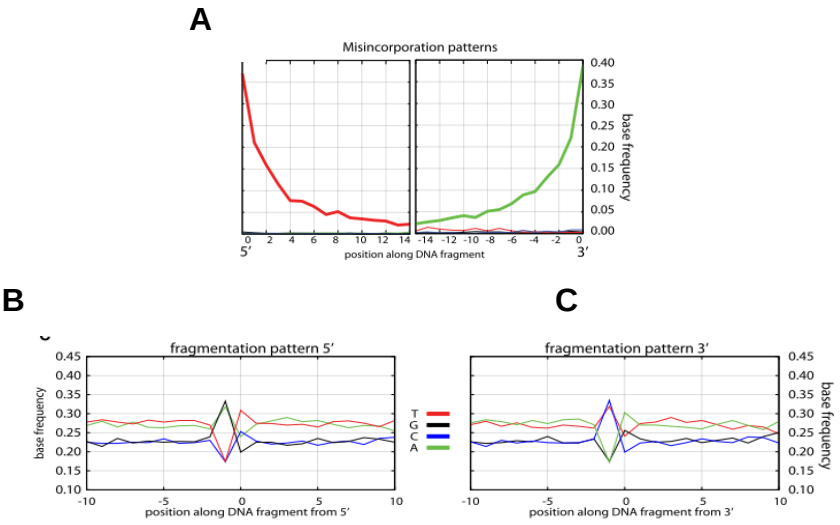


Figure 25: Cytosine residues in the consensus sequence appear as thymine residues at 5'-ends of sequences, and guanidine residues appear as adenine residues at 3'-ends (A). Furthermore, purines are substantially elevated at the base 5' to the mtDNA sequences and pyrimidines are elevated at the base 3' to sequences (B and C), indicating preferential breakage at purines (Krause et al. 2010).

This pattern consists in a preferential fragmentation of ancient DNA at purine bases, and in contrast, the human contaminants show no increase in frequency of purines or pyrimidines on either side of the fragment. In conclusion, fragment length, deamination-induced sequence errors at the ends of molecules, and purine-associated fragmentation represent features by which endogenous and contaminating populations of DNA molecules can be distinguished in at least some late Pleistocene specimens (Briggs et al. 2007).

1.6. Technological approaches

1.6.1. Polymerase Chain Reaction

The **polymerase chain reaction** (**PCR**), developed by Kary Mullis (Mullis and Faloona 1987) in 1983, is a scientific technique in molecular biology to amplify a single or a few copies of a piece of DNA across several orders of magnitude, generating thousands to millions of copies of a particular DNA sequence (Saiki et al. 1988; Pääbo et al.1989).

The method relies on thermal cycling, consisting of cycles of repeated heating and cooling of the reaction for DNA melting and enzymatic replication of the DNA. Primers containing sequences complementary to the target region along with a DNA polymerase are key components to enable selective and repeated amplification. Almost all PCR applications employ a heat-stable DNA polymerase, such as Taq polymerase, an enzyme originally isolated from the bacterium *Thermus aquaticus* (Saiki et al. 1988).

A basic PCR set up requires several components and reagents. These components include:

- *DNA template* that contains the DNA region (target) to be amplified.
- Two *primers* that are complementary to the 3' (three prime) ends of each of the sense and anti-sense strand of the DNA target.
- *Taq polymerase* or another DNA polymerase with a temperature optimum at around 70 °C.

- *Deoxynucleotide triphosphates* (dNTPs), the building blocks from which the DNA polymerases synthesizes a new DNA strand.
- *Buffer solution,* providing a suitable chemical environment for optimum activity and stability of the DNA polymerase.
- Divalent cations, magnesium or manganese ions; generally Mg^{2+} is used.

This DNA polymerase enzymatically assembles a new DNA strand from DNA nucleotides, by using single-stranded DNA as a template and primers, which are required for initiation of DNA synthesis.

The vast majority of PCR methods use thermal cycling. These thermal cycling steps are necessary first to physically separate the two strands in a DNA double helix at a high temperature in a process called DNA melting. At a lower temperature, each strand is then used as the template in DNA synthesis by the DNA polymerase to selectively amplify the target DNA. The selectivity of PCR results from the use of primers that are complementary to the DNA region targeted for amplification under specific thermal cycling conditions (Caramelli and Lari 2004).

The main steps of Polymerase Chain Reaction are (Figure 26):

Initialization step: This step consists of heating the reaction to a temperature of 94–96°C It is only required for DNA polymerases that require heat activation by hot-start PCR.

Denaturing step: This step is the first regular cycling event and consists of heating the reaction to 94–98°C for 20–30 seconds. It causes DNA melting of the DNA template by

disrupting the hydrogen bonds between complementary bases, yielding single-stranded DNA molecules.

Annealing step: The reaction temperature is lowered to 50–65°C for 20–40 seconds allowing annealing of the primers to the single-stranded DNA template. Typically the annealing temperature is about 3-5 degrees Celsius below the Tm of the primers used. Stable DNA-DNA hydrogen bonds are only formed when the primer sequence very closely matches the template sequence. The polymerase binds to the primer-template hybrid and begins DNA synthesis (Rychlik et al. 1990).

Extension/elongation step: The temperature at this step depends on the DNA polymerase used; Taq polymerase has its optimum activity temperature at 72–80°C, and commonly a temperature of 72°C is used with this enzyme. At this step the DNA polymerase synthesizes a new DNA strand complementary to the DNA template strand by adding dNTPs that are complementary to the template in 5' to 3' direction, condensing the 5'-phosphate group of the dNTPs with the 3'-hydroxyl group at the end of the nascent DNA strand. Under optimum conditions, at each extension step, the amount of DNA target is doubled, leading to exponential amplification of the specific DNA fragment.

Final elongation: This single step is occasionally performed at a temperature of 70–74°C for 5–20 minutes after the last PCR cycle to ensure that any remaining single-stranded DNA is fully extended.

Final hold: This step at 4–15°C for an indefinite time may be employed for short-term storage of the reaction (Caramelli and Lari 2004).

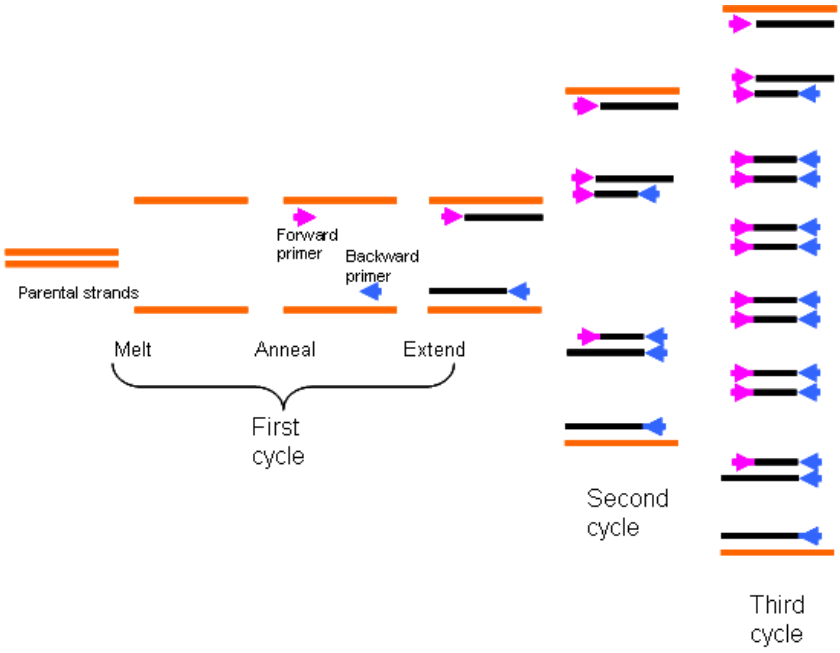


Figure 26: Polymerase Chain Reaction (PCR). The double-strand DNA is denatured, primers bind to single-strand DNA target (in pink arrow forward primer and in blue arrow backward primer). Taq polymerase, using the target DNA as template, synthesized the new strand of DNA. The number of new strands of DNA is exponential (http://www.google.es).

The PCR is commonly carried out in a reaction volume of 10–200 µl in small reaction tubes (0.2–0.5ml volumes) in a thermal cycler. Typically, PCR consists of a series of 20-60 cycles.

To check whether the PCR generated the DNA fragment (amplicon), agarose gel electrophoresis is employed for size separation of the PCR products. The size of PCR products is determined by comparison with a DNA ladder (a molecular weight marker), which contains DNA fragments of known size, run on the gel alongside the PCR products (Caramelli and Lari 2004).

In the analysis of ancient remains it can be useful to use the modified PCR protocols:

Quantitative PCR (Q-PCR): used to measure the quantity of a PCR product (commonly in real-time). It quantitatively measures starting amounts of DNA, cDNA or RNA. Q-PCR is used to determine whether a DNA sequence is present in a sample and the number of its copies in the sample. *Quantitative real-time PCR* has a very high degree of precision (Hodges et al. 2009).

Multiplex-PCR: consists of multiple primer sets within a single PCR mixture to produce amplicons of varying sizes that are specific to different DNA sequences (Krause et al. 2006; Lalueza-Fox et al. 2008).

Despite its great usefulness, in ancient DNA study, PCR presents two main limitations: the small quantity and short length of ancient DNA fragments due to the degradation of the original template molecules and the difficulty in distinguishing between endogenous and exogenous DNA sequences (Lalueza-Fox et al. 2008).

1.6.2. Next Generation Sequencing (NGS)

It is widely recognised that recent advances in DNA sequencing technology and the development of downstream genomics tools, are changing the face of most areas of biology (Stapley et al. 2010). The recent development of the Next Generation Sequences (NGS) (Genome sequencer GS20; Roche Applied Science, Illumina/Solexa and ABI SOLID) offers a solution to complicated problems. Specifically, the nature of the data-generation process is such that DNA sequence data can be unambiguously assigned to individual, original single-stranded molecules (Gilbert et al. 2007).

Three are the main sequencing-by-synthesis technology machines:

Illumina/Solexa: The Illumina Genome Analyzer is based on parallel, fluorescence-based readout of millions of immobilized sequences that are iteratively sequenced using reversible terminator chemistry (Bendey et al. 2008). In brief, up to eight DNA libraries are hybridized to an eight-lane flow cell. In each of the lanes, single-stranded library molecules hybridize to complementary oligos that are covalently bound to the flow cell surface. Using the double stranded duplex, the reverse strand of each library molecule is synthesized and the now covalently bound molecule is then further amplified in a process called bridge amplification. This generates clusters each containing more than 1,000 copies of the starting molecule. One strand is then selectively removed, free ends are subsequently blocked and a sequencing primer is annealed onto the adapter sequences of the

cluster molecules. Starting from the sequencing primers, 3' terminated and fluorescence-labeled nucleotides are incorporated using a modified polymerase. Base incorporation ceases after the addition of a single base due to the 3' termination of the incorporated nucleotides. The fluorophores attached to the nucleotides are illuminated using a red and a green laser, and imaged through different filters, yielding four images per tile.

The number of tiles varies; for Genome Analyzer I it is typically 300 tiles per lane, for Genome Analyzer II it is 100 tiles per lane. After an imaging cycle, the fluorescent labels as well as the 3' terminators are chemically removed and the next incorporation cycle is started. Incorporation and imaging cycles are repeated up to a designated number of cycles, defining the read length for all clusters. After sequencing, images are analyzed and intensities extracted for each cluster (Kircher et al. 2009).

Applied Biosystem SOLID: This platform generates 6 or more Gb per run, in 25-35 nt reads, and uses a unique ligation-mediated sequencing strategy that is less prone to some of the problems that have been associated with high throughput sequencing-by-synthesis strategies, such as inaccurate recording of homopolymer sequences (Shendure et al. 2005). In addition, the SOLiD system uses a two-base encoding scheme in which each data point represents two adjacent bases, and each base is interrogated twice, which helps in discriminating between sequencing errors and true polymorphisms. Collectively, these attributes make the SOLiD sequencing system particularly well suited to a variety of functional genomics applications (Ondov et al. 2008).

In contrast to other sequencing systems, SOLiD data are not collected directly as DNA sequences, but instead are recorded in 'colour space', in which the individual values (colours) within a read provide information about two adjacent bases. Without a decoding step, in which colour data are converted to sequence data, they cannot be mapped to a reference genome using conventional alignment tools. Direct conversion of colour data to sequence data, however, has a significant drawback-reads that contain sequencing errors cannot be converted accurately (in translating a color space string, all bases after a sequencing error will be translated incorrectly) (Ondov et al. 2008).

Given this, there is a clear incentive to map sequence reads to a reference genome within colour space, and there have been several software tools developed recently to perform this task:

- MAQ (http://maq.sourceforge.net/)
- Shrimp (http://compbio.cs.toronto.edu/shrimp/)
- Mosaik (http://bioinformatics.bc.edu/marthlab/Mosaik)

The main works, show in this thesis, were made using 454 Life Sciences GS FLX technologies (Figure 27), for this reason we decided to report an accurate description of its protocol.

454 Life Sciences GS FLX: this approach uses a large-scale parallel pyrosequencing system capable of sequencing roughly 400-600 megabases of DNA per 10-hour run**.** It can sequence any double-stranded DNA and enables a variety of applications including *de novo* whole genome sequencing, re-sequencing of

whole genomes and target DNA regions, metagenomics and RNA analysis.

The main steps of ultrasequencing technology are three:

Preparing a DNA Library: the DNA sample must be transformed into a library of DNA fragments appropriate for sequencing with the Genome Sequencer FLX System. The method for preparing this library varies according to the type of sample and the objective of the experiment. The first step is the nebulization of DNA yields in some fragments with frayed ends. These ends are made blunt and ready for ligation to adaptors by the action of T4 DNA polymerase and T4 polynucleotide kinase (T4 PNK). The 5'→ 3' polymerase activity of T4 DNA polymerase fi lls in 3'–recessed ends (5'–overhangs) of DNA, while its single-stranded 3'→5' exonuclease activity removes 3'–overhang ends. The kinase activity of T4 PNK adds phosphate groups to the polished 5'–hydroxyl termini. Following fragmentation and polishing of the DNA library, primer sequences termed "Adaptors" are ligated to the ends of each sample DNA fragment. These Adaptors are a pair of double-stranded oligonucleotides (Adaptors "A" and "B") that provide priming regions to support both amplifi cation and nucleotide sequencing. They also provide a unique 4-base non palindromic sequencing key used by the system's software for base calling and to recognize legitimate library reads. Adaptor B also contains a biotin tag on its 5'-strand. The ligation mixture is immobilized onto magnetic streptavidin coated beads, via the biotin moiety of Adaptor B. The ligation reaction produces a mixture of molecular species including: sample DNA fragments with ligated adaptors on either or both ends, unligated or self-ligated (circularized) sample

DNA fragments, unbound single Adaptors, and Adaptor dimers. Because the DNA oligonucleotides used for the Adaptors are not phosphorylated, gaps will be present at their 3'-junctions with the DNA fragments. These two gaps or nicks are repaired using a strand-displacing DNA polymerase, which recognizes the nicks, displaces the nicked strands (starting from the free 3'–end of each Adaptor), and fills in the single-stranded gap, producing full-length dsDNA. The next step is to isolate the single-stranded moieties (single-stranded template DNA, or sstDNA) by melting off the non-biotinylated strand of each dsDNA fragment. The adapted dsDNA fragments are immobilized via the biotin tag of Adaptor B. Therefore, fragments without an Adaptor B have already been washed away, and fragments with Adaptor B at both ends will not release either strand. The procedure is thus designed to generate a library comprising only sstDNA molecules flanked with Adaptor A at the 5'–end and Adaptor B at the 3'–end. sstDNA Library.

Amplifying the Library: The fragments that make up the DNA library are immobilized onto DNA Capture Beads, to allow for their segregation in the emulsion. The goal is to achieve one effective molecule (a functional clone) of DNA per bead; and one bead per aqueous microreactor (micelle), insulated from other beads by the surrounding oil. It is the segregation of each bead into its own microreactor (with amplification reagents) that will maintain clonality during the amplifi cation step. The captured DNA library is resuspended in the amplification mix and oil, and emulsified to form a water-in-oil mixture. This step involves vigorous mechanical shaking under tightly controlled conditions. The products of the emulsification step are aqueous phase "microreactors" 50 to 100 μm in diameter, each containing all the components of the

amplification mix and no more than a single bead. The emulsified beads are subjected to PCR to clonally amplify each template DNA molecule. The DNA templates are hybridized to bead-bound oligonucleotide primers. These capture primers double as PCR primers, anchoring the newly synthesized, complementary strands to the beads. As the PCR reaction progresses, these bead-bound, complementary strands direct the synthesis of more first-strand moieties, which hybridize to an excess of bead-bound capture primers. Also, the second, soluble amplification primer is biotinylated; this allows for enrichment of the beads carrying amplified DNA later in the procedure. After amplification, typical immobilized template copy number ranges from 10 to 30 × 106 copies per bead. Bead Recovery: After the amplification step, the emulsion is broken chemically, and the beads carrying the amplified DNA library (double-stranded at this point) are recovered and washed by filtration. DNA Library Bead Enrichment: The procedure above generates a certain proportion of beads that carry no amplified DNA (null beads), either because they did not capture a molecule of template in the beginning or because the DNA template did not amplify properly. To reduce the percentage of beads without template, the sixth step of the procedure enriches the total bead population for amplified DNA-carrying beads. This enrichment step involves binding of the biotinylated amplification primer “A” to streptavidin-coated magnetic beads, and then separating the bound beads (carrying amplified DNA) from “null beads” with a magnetic particle collector. The DNA library beads are then separated from the magnetic beads by melting the double-stranded amplification products, leaving a population of bead-bound single-stranded template DNA fragments: the immobilized and amplified DNA library. Sequencing Primer Annealing: The final step in the

emPCR process is the annealing of the sequencing primer to the immobilized, amplified DNA templates.

Sequencing the Library: After amplification, the DNA-carrying beads are loaded into the wells of a PicoTiterPlate device (PTP) so that the wells contain single DNA beads. The loaded PTP is then inserted into the Genome Sequencer FLX Instrument, and sequencing reagents are sequentially flowed over the plate. The Genome Sequencer FLX Instrument automatically performs and monitors the sequencing reactions in all the wells of the PTP simultaneously (Margulies et al 2005).

The key benefit therefore is that each final sequence reaction is generated from a single single-stranded DNA molecule and, as such, provides a direct window into any damage-derived miscoding lesions that were present on the molecule, thus in an instant providing the critical information that has been lacking from previous aDNA damage studies (Gilbert et al. 2007).

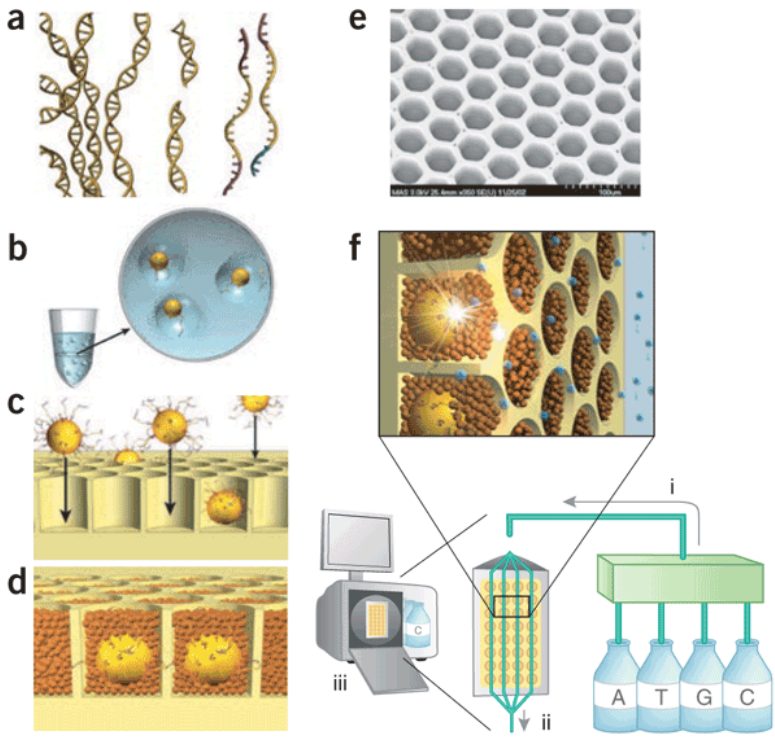


Figure 27: (**a**) Genomic DNA is isolated, fragmented, ligated to adapters and separated into single strands. (**b**) Fragments are bound to beads under conditions that favor one fragment per bead, the beads are isolated and compartmentalized in the droplets of a PCR-reaction-mixture-in-oil emulsion and PCR amplification occurs within each droplet, resulting in beads each carrying ten million copies of a unique DNA template. (**c**) The emulsion is broken, the DNA strands are denatured, and beads carrying single-stranded DNA templates are enriched (not shown) and deposited into wells of a fiber-optic slide. (**d**) Smaller beads carrying immobilized enzymes required for a solid phase pyrophosphate sequencing reaction are deposited into each well. (**e**) Scanning electron micrograph of a portion of a fiber-optic slide, showing fiber-optic cladding and wells before bead deposition. (**f**) The 454 sequencing instrument consists of the following major subsystems:a fluidic assembly, a flow cell that includes the well-containing fiber-optic slide, a CCD camera-based imaging assembly with its own fiber-optic bundle used to image the fiber-optic slide (part of object iii), and a computer that provides the necessary user interface and instrument control (Margulies et al. 2005).

1.6.3. Capture Systems

After the advent of shot gun sequences (Margulies et al. 2005), new methods to try to recover more sequences from ancient DNA samples, were developed (Briggs et al. 2009; Burbano et al. 2010; Maricic et al. 2010). Consider that the amount of shotgun sequencing sequences is prohibitive for most ancient bone specimens due to the high fraction of environmental DNA that they contain and a simple shotgun approach would require hundreds or thousands of high-throughput pyrosequencing. Also direct polymerase chain reaction (PCR) is poorly suited for retrieving complete Neanderthal mtDNA genomes, because DNA extracted from the fossils is so fragmented that hundreds of overlapping amplicons would be necessary, either requiring highly multiplexed primer mixes that present severe difficulties for avoiding modern human contamination, or many parallel amplification reactions that consume large amounts of precious ancient DNA extracts. For these reasons Briggs and co-workers developed a new method called primer extension capture (PEC) that directly isolates specific DNA sequences from complex libraries of highly degraded DNA (Figure 28).

PEC uses 5′- biotinylated oligonucleotide primers and a DNA polymerase to capture specific target sequences from an adaptor-ligated DNAlibrary. It combines the high specificity of PCR primers with the numerous advantages of a library sequencing approach, including immortalization through reamplification from adaptor priming sites, contamination control with project specific barcodes, access to very short fragments predominant in ancient extracts,

and quantification of the number of unique ancient DNA molecules which is necessary to identify nucleotide misincorporations (Briggs et al. 2009).

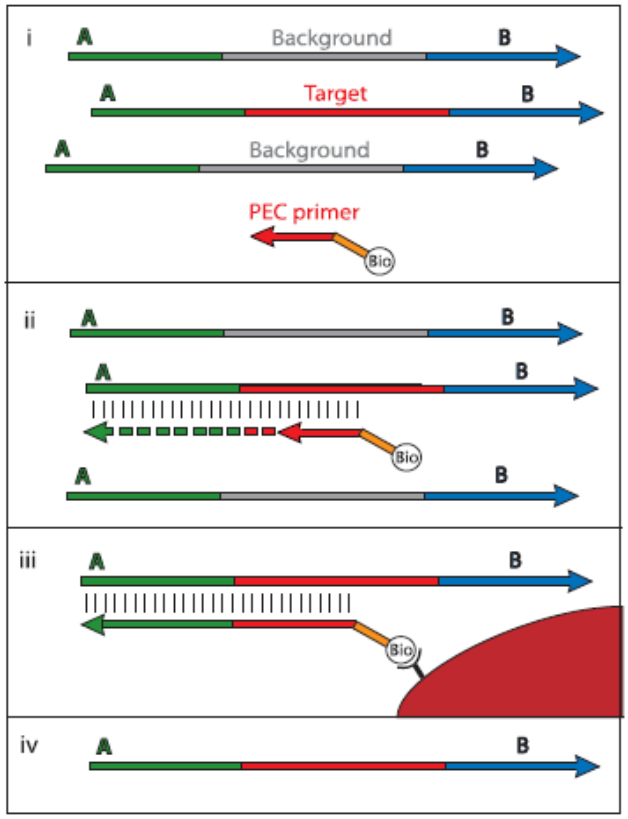


Figure 28: Primer extension capture (PEC). (**i**) 5′-Biotinylated oligonucleotide primers (PEC primers) are added to a 454 library [in which the A and B adaptor molecules carry a project-specific barcode] and are allowed to anneal to their respective target sequences. (**ii**) A single Taq DNA polymerase extension step is performed, resulting in a double-stranded association between primer and target that includes the 5′ adaptor sequence. (**iii**) Excess PEC primers are removed by spin column purification, and the biotinylated primer:target duplexes are captured by streptavidin-coated magnetic beads. The beads are washed stringently above the melting temperature of the PEC primers, to ensure that templates upon which extension occurred will preferentially remain associated with the primers. (**iv**) Captured and washed targets are eluted from the beads, amplified with adaptor priming sites, and subjected either to a second round of extension and capture or directly to 454 emulsion PCR (Briggs et al. 2009).

The PEC method is unlikely to be scalable up to megabase target regions, ruling out experiments such as the retrieval of large portions of genome. For this reason Burbano et al. (2010) investigate the use of massively parallel hybridization capture on glass slide microarrays (Hodges et al. 2009) on ancient DNA sample to study thousands of genomic positions where nucleotide substitutions changing amino acids have occurred on the human lineage since its split from chimpanzees.

The methodological approach is based on tiling arrays which consist of chemically synthesized and spatially immobilized oligonucleotides in which predefined sequences reflect regions of the genome at frequent and uniform intervals. Such arrays may be programmed to exclude repetitive elements, thereby optimizing available array capacity and performance. (Hodges et al 2009). As described in Hodges et al. (2009), the working steps five:

a- Array design: the array configuration is designed in a similar manner to each described in Hodges et al. (2007)
b- Library construction: the quality of the genomic DNA library strongly influences the success of the array capture. Usually the library building is based on the standard Illumina procedure (Bentley et al. 2008).
c- Blocking repeats: the arrays are processed under stringent hybridization and washing conditions, certain factors can influence the specificity of the capture experiments
d- Elution: Hodges and colleagues takes advantage of the standard Agilent slide gasket chamber system. This method comprises a steel chamber base and a rubber gasket-slide that, when assembled, forms a sandwich with the microarray and creates a

hybridization compartment for the printed surface of the array. The elution mixture is withdrawn from the chamber using a syringe. Subsequently, an additional qPCR step is carried out to accurately quantify the captured material.

e- Evaluating the results: Capture performance can be assessed either by qPCR for a defined set of intervals. qPCR is useful for acquiring an initial snapshot of success, and as an experimental quality control for determining whether sequencing will be informative.

However, qPCR does not provide a global sense of enrichment specificity and sensitivity. Therefore, massively parallel sequence analysis from single captured molecules is the only accurate and comprehensive way to estimate performance.

All these methods (Briggs et al. 2009; Hodges et al. 2009; Burbano et al. 2010) involve probes and/or equipments that are only available from commercial vendors and are not suitable for the targets of a few kilobases in length.

To try to make the technique of the capture available to a wider group of researchers, Marcic et al. (2010) (Figure 29) developed a new method in which custom made long-range PCR products are used to capture complete human mitochondrial genomes. The basic point of this technique is the production of the probes starting from PCR product. In this step the two overlapping long-range PCR products encompassing the whole mitochondrial genome, were sonicated producing fragments from 150 to 850 bases long. These fragments were biotinylated by ligating the Bio-T/B adapter,

purified, made single-stranded and immobilized on streptavidin-coated magetic beads. This approach is economical and particularly suitable for targets that can be amplified by PCR and do not contain highly repetitive sequences.

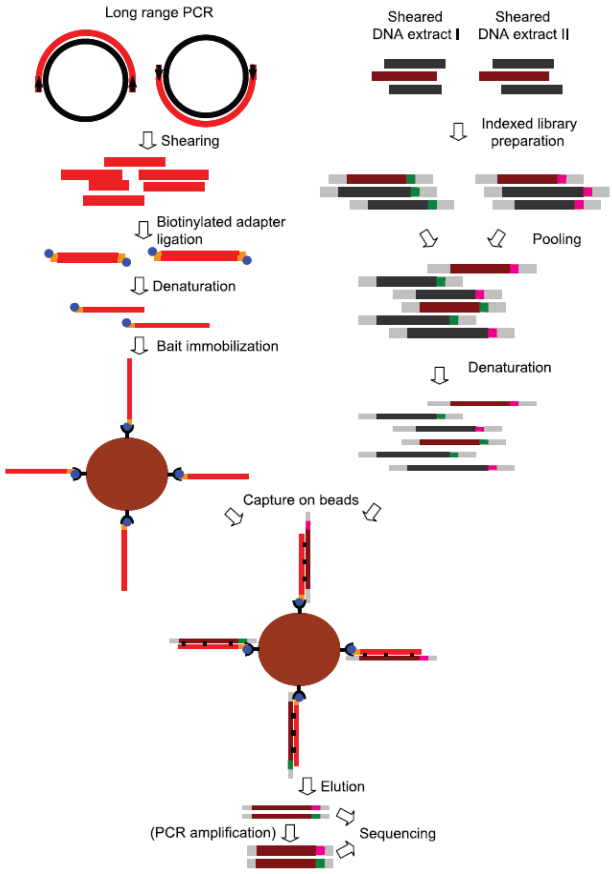


Figure 29: Overview of the capture-on-beads method. On the left the production of the immobilized bait from two long range PCR products is shown; on the right the production of a pool of indexed libraries which are used in the capture (bottom). The eluted molecules can either be sequenced directly or first amplified and then sequenced. The bait is light red, mitochondrial DNA in the libraries is dark, indices are shown in green and pink, adapters in gray. Thicker lines represent double stranded DNA while thinner lines represent single stranded DNA (Maricic et al. 2010).

1.7. Analysis of DNA sequences

1.7.1. Analysis of PCR products

To analyse the small products of PCR, electrophoresis technique is used. DNA **electrophoresis** is an analytical technique used to separate the molecules of DNA by size. DNA molecules which are to be analyzed are situated in the gel (made with agarose) at the level of the cathode, where an electric field induces the DNA to migrate toward the anode, due to the net negative charge of the sugar-phosphate backbone of the DNA chain. The separation of these fragments is accomplished by exploiting the mobility with which different sized molecules are able to pass through the gel (Figure 30). The DNA fragments of different lengths are visualized using a fluorescent dye specific for DNA, such as ethidium bromide.The gel shows bands corresponding to different DNA molecule populations with different molecular weight. Fragment size determination is typically done by comparison to commercially available DNA markers containing linear DNA fragments of known length (Russel 1998).

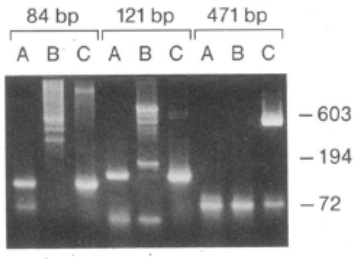


Figure 30: Agarose gel electrophoresis of amplification products. Lanes A contain the sample, lanes B control extract and lanes C contemporary human DNA. The migration positions of molecular size markers are given in base pair (Pääbo 1989).

The latest step after PCR and cloning is the **sequencing**. The DNA sequencing is a method for determining the order of nucleotides bases. The most known and used, for this purpose, is the technique developed by Frederik Sanger also called *chain-termination method* (Sanger et al. 1977). This approach (Figure 31) requires a single-stranded DNA template, a DNA primer, a DNA polymerase, normal deoxynucleotidephosphates (dNTPs), and modified nucleotides (dideoxyNTPs) that terminate DNA strand elongation. These ddNTPs will also be radioactively or fluorescently labelled for detection in automated sequencing machines. The DNA sample is divided into four separate sequencing reactions, containing all four of the standard deoxynucleotides (dATP, dGTP, dCTP and dTTP) and the DNA polymerase. To each reaction is added only one of the four dideoxynucleotides (ddATP, ddGTP, ddCTP, or ddTTP) which are the chain-terminating nucleotides, lacking a 3'-OH group required for the formation of a phosphodiester bond between two nucleotides, thus terminating DNA strand extension and resulting in DNA fragments of varying length.

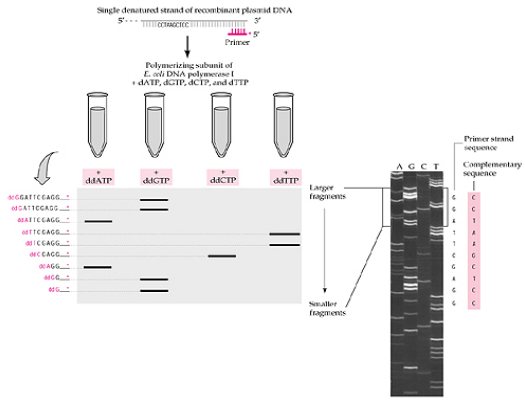


Figure 31: Sanger sequencing technique (www.google.es)

The newly synthesized and labelled DNA fragments are heat denatured, and separated by size by gel electrophoresis on a denaturing polyacrylamide-urea gel with each of the four reactions run in one of four individual lanes (lanes A, T, G, C); the DNA bands are then visualized by autoradiography or UV light, and the DNA sequence can be directly read off the X-ray film or gel image. In the image on the right, X-ray film was exposed to the gel, and the dark bands correspond to DNA fragments of different lengths. A dark band in a lane indicates a DNA fragment that is the result of chain termination after incorporation of a dideoxynucleotide (ddATP, ddGTP, ddCTP, or ddTTP). The relative positions of the different bands among the four lanes are then used to read (from bottom to top) the DNA sequence (Sanger et al. 1977).

Once that the sequences have been obtained, the next step is the **alignment**. This can be performed using some types of software building with different algorithms.

Very short or very similar sequences can be aligned by hand, but, because of the complexity of the sequence, this might not be enough. For this reason, human knowledge is applied in constructing algorithms to produce high-quality sequence alignments. Computational approaches to sequence alignment generally fall into two categories: *global alignments* and *local alignments.* Calculating a global alignment is a form of global optimization that "forces" the alignment to span the entire length of all query sequences. By contrast, local alignments identify regions of similarity within long sequences that are often widely divergent

overall. Local alignments are often preferable, but can be more difficult to calculate because of the additional challenge of identifying the regions of similarity. Hybrid methods, known as semiglobal or "glocal" methods, attempt to find the best possible alignment that includes the start and end of one or the other sequence. This can be especially useful when the downstream part of one sequence overlaps with the upstream part of the other sequence. In this case, neither global nor local alignment is entirely appropriate: a global alignment would attempt to force the alignment to extend beyond the region of overlap, while a local alignment might not fully cover the region of overlap.

Pairwise sequence alignment methods are used to find the best-matching piecewise (local) or global alignments of two query sequences. Pairwise alignments can only be used between two sequences at a time, but they are efficient to calculate and are often used for methods that do not require extreme precision (such as searching a database for sequences with high similarity to a query). Multiple sequence alignment is an extension of pairwise alignment to incorporate more than two sequences at a time. Multiple alignment methods try to align all of the sequences in a given query set. Multiple alignments are often used in identifying conserved sequence regions across a group of sequences hypothesized to be evolutionarily related. Such conserved sequence motifs can be used in conjunction with structural and mechanistic information to locate the catalytic active sites of enzymes. Alignments are also used to aid in establishing evolutionary relationships by constructing phylogenetic trees.

Some of the most used software to alignment the PCR product are summarized in the tables following (Tables 1 and 2).

Table 1: Database search only

Name	Description	Type*	Year	Authors
BLAST	local search with fast k-tuple heuristic (Basic Local Alignment Search Tool)	both	1990	Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ
FASTA	local search with fast *k*-tuple heuristic, slower but more sensitive than BLAST	both	1985	Lipman DJ Pearson WR

Type*- proteins or nucleotides

Table 2: Multiple sequence alignment

Name	Description	Type*	Type**	Year	Authors
CLUSTAL W	Progressive alignment	both	global or local	1994	Thompson *et al*
T-COFEE	More sensitive progressive alignment	both	global or local	2008 (Last version)	C. Notredame *et al.*
MUSCLE	Progressive/iterative alignment	both	global or local	2004	R. Edgar

Type*- proteins or nucleotides

Type**- Global or local aligment

1.7.2. Massively parallel sequencing analysis

The first step in the analysis of shotgun ancient DNA data is to **identify the target-species (endogenous) fragments**. The primary goal of this step is to reliably identify as many endogenous fragments as possible. If only microbial contamination is present, microbial sequences could be initially identified and then subtracted. But it is good to keep in mind that currently available microbial sequence data are too incomplete to detect the full

diversity naturally occurring in microbial communities (Prufer et al. 2010).

When contaminants are human exogenous sequences, the situation becomes more complex. To identify the endogenous sequences, it is possible to follow a different approach. One could be genotyping all individuals who have come into contact with a sample and excluding the sequences identical to those (Krause et al. 2010) but, it is not always possible to identify all the people (archeologist, paleontologist, etc) who worked with the sample.

Another way is to identify target- species DNA fragments by similarity between these and the sequence of a closely related species, (Green et al. 2010; Prufer et al. 2010) taking into account that, recently, by studying mtDNA in Neanderthal remains (Briggs et al. 2009; Lalueza-Fox et al. 2010), where contamination and endogenous DNA can be distinguished by sequence, Krause et al. (2010) observed that fragmentation patterns and nucleotide misincorporations can be used to gauge authenticity of ancient DNA sequences.

The second step is mapping the reads against the reference genome. To do this several software are available, some of which present particular adaptations to ancient DNA characteristics. Some examples are, MIA (Briggs et al. 2009), BWA (Reich et al. 2010), and Suffix array-based method (Rasmussen et al. 2010).

Not all sequences are taken in to account, usually, sequences with less of 30 nt, (Green et al. 2006) and reads with multiple matches or no matches were discarded (Rasmussen et al. 2010). Read

sequences from the same library that were mapped to the reference genome with same start and end positions were considered clonal, and were collapsed to single sequences with higher quality scores.

2. OBJECTIVES

OBJECTIVES

In the current thesis we are interested in different technological aspects of the aDNA research field and in adaptive traits that characterized the Neanderthal populations.

(i) The pre-laboratory modern human DNA contamination when working with ancient human remains is one of the main problems to solve in aDNA field. Some guidelines has been suggested to try to decrease the amount of exogenous sequences in the laboratory, but little is known regarding pre-laboratory contamination. We will develop an anti-contamination protocol using the Neanderthal samples retrieved in El Sidron (Spain) site. The main idea of this work is test its efficiency through DNA analysis of samples retrieved before and after the implementation of this protocol at the excavation.

(ii) In addition to, a central issue in ancient DNA field is the control of contamination in the laboratory itself, which is particularly important when contaminants and endogenous sequences are potentially indistinguishable. To control this problem, we will develop a new methodological approach based on

the incorporation of “blocking primers” designed to match contaminant sequences during the PCR.

(iii) The recent development of high-throughtput DNA sequencing technologies has allowed to show that the principal cause of fragmentation in ancient DNA is the depurination. In this context, our goal is to try to verify a posteriori endogenous and exogenous sequences by the study of fragmentation pattern analyzing the base composition at 5’ and 3’ –ends of the 454-generated sequences.

(iv) Analysis of genetic variation in *H. neanderthalensis* and *H. sapiens* is a powerful tool for understanding their specific evolutionary characteristics. Coalescent models can estimate the divergence time of different gene systems, but obviously a direct analysis can provide experimental evidence for the presence of some polymorphisms. We will analyze two nuclear genes, *TAS2R38* -associated to bitter taste perception- and *ABO* blood group system –associated to natural immunity-, that are highly polymorphic in modern humans. Our goal is to understand adaptation and gene divergence in Neanderthals in these genetic systems.

3. RESULTS

3.1. Excavation protocol of bone remains for Neandertal DNA analysis in El Sidrón Cave (Asturias, Spain)

Fortea J, de la Rasilla M, Garcia-Tabernero A, **Gigli E**, Lalueza-Fox C. (2008) J. Hum. Evol. 55(2): 353-7

3.2. An improved PCR method for endogenous DNA retrieval in contaminated Neandertal samples based on the use of blocking primers.

Elena Gigli*, Morten Rasmussen*, Sergi Civit, Antonio Rosas, Marco de la Rasilla, Javier Fortea, M. Thomas P. Gilbert, Eske Willerslev and Carles Lalueza-Fox. (2009) Journal of Archaeological Science 36:2676-79

* These authors equally contributed to this work

3.3. Fragmentation of contaminants and endogenous DNA in ancient samples determined by shotgun sequencing; prospects for human paleogenomics

Marc Garcia-Garcerà*, Elena Gigli*, Federico Sanchez-Quinto, Sergi Civit, Carles Lalueza-Fox

*These authors equally contributed to this work

In preparation

Fragmentation of contaminant and endogenous DNA in ancient samples determined by shotgun sequencing; prospects for human paleogenomics

Marc García-Garcerà[1]*, Elena Gigli[1]*, Federico Sánchez-Quinto[1], Sergi Civit[2], Carles Lalueza-Fox[1]

*These authors equally contributed to this work

[1]*Institut de Biologia Evolutiva, CSIC-UPF, Dr. Aiguader 88, 08003 Barcelona, Spain*

[2]*Department de Bioestadística, Universitat de Barcelona, 08028 Barcelona, Spain*

Corresponding author

Carles Lalueza-Fox
Institute of Evolutionary Biology (UPF-CSIC)
Dr. Aiguader 88
08003 Barcelona (Spain)
Phone: +34933160845
Email: carles.lalueza@upf.edu

Running head: Fragmentation of contaminants and ancient human DNA

Abstract

Despite the successful retrieval of genomes from past remains, the prospects for human paleogenomics remain unclear because of the difficulty for distinguishing contaminant sequences from the endogenous ones. Results from high throughput sequencing platforms indicate that the

fragmentation in ancient DNA sequences is primarily due to depurination processes that create abasic sites leading to DNA breaks. To test if this pattern is present in ancient remains from a temperate environment, we have subjected to GS20-FLX 454 pyrosequencing different samples dated between 5,500 and 49,000 years ago: a bone from an extinct goat (*Myotragus balearicus*) that was treated with a depurinating agent (bleach), an Iberian lynx bone not subjected to any treatment, a human Neolithic sample from Barcelona (Spain) and also a Neanderthal sample from El Sidrón site (Asturias, Spain). We have used the non human samples to identify human sequences (0.35 and 1.4%, respectively), that we positively know are contaminants. We found that the bleach treatment creates a depurinating-associated fragmentation pattern undistinguishable from the endogenous one, caused by depurinations. The pattern observed among contaminants is much more complex and similar to that previously found in other contaminanted samples. Unless endogenous and contaminant sequences can be distinguished, human paleogenomics is going to be a problematic field.

Keywords: Paleogenomics, 454, human, contamination

Introduction

Ancient DNA (aDNA) analysis on extinct human populations can potentially provide information on past human migrations and evolutionary processes. Several laboratory practices were widely adopted to authenticate ancient DNA results (Cooper and Poinar 2000). They included, for instance, physical separation of the ancient and modern DNA laboratory, frequent bleaching and UV irradiation on the working surfaces, sterile clothing, cloning of PCR products and independent replication of the results in a second laboratory. However, all these authentication criteria have proven to be ineffective while working on ancient human specimens, because modern contaminant DNA can be mistaken for endogenous DNA (Hofreiter et al. 2001a). This is due to the apparent impossibility of distinguishing both types of sequences and the fact that not all the contamination can be controlled at the laboratory, even under the most stringent precautions. Several studies pointed to the existence of a pervasive pre-laboratory contamination, as showed by the presence of human DNA in non-human samples, such as pigs, foxes, cave bear, dogs or Neandertals (Hofreiter et al. 2001b; Mälmstrom et al. 2007). Therefore, independent replication of the results in another laboratory will not eliminate the problem.

This contamination can be to some extent monitorized in present-day excavations, in which the sequences of the people that

contemporaneously handle and manipulate the remains can be genotyped. Studies that have followed this approach have identified significant amounts of these traceable contaminating sequences, mixed with the putative -and usually prevalent- endogenous ones. It has been suggested that ancient human DNA specimens destined to future paleogenetic studies should be based on material extracted under controlled conditions (e.g., with sterile lab gear and with the immediate freezing of the samples) (Fortea et al. 2009). However, a vast majority of human archaeological material derives from old excavations in which these precautions were not implemented. An additional concern is that, at least in rather old contaminants (e.g., more than ten years old), and probably depending on the storage conditions, these extraneous DNA can be chemically degraded to a similar level than the endogenous one (Sampietro et al. 2006), thus rendering impossible to distinguish both types of sequences in the light of the postmortem damage (Helgason et al. 2007; Vives et al. 2008).

Another authentication criterion proposed, that of the "appropriate molecular behavior" refers to potential difference in length between the endogenous and the contaminant DNA. As time passes after the death of the organism, the surviving DNA is subjected to enzymatic hidrolytic processes that break the double strand and degrades it into progressively smaller fragments. Since any contaminant DNA is significantly more recent than the ancient sample, it is assumed that the former will often be

of larger average size than the original DNA. Mälmstrom et al. (2007) analysed dog samples contaminated with human DNA and found more authentic DNA in shorter than in longer fragments retrieved. Thus, ancient samples show an increase in authentic DNA yield with decreased fragment size than the contaminating DNA. This assymetrical behaviour could be explained if the contaminant sequences are on average longer than the endogenous ones (Mälmstrom et al. 2007).

In the last years, new high-throughput sequencing techniques, such as the 454 pyrosequencing (Life Sciences-Roche) or the Illumina-Solexa platform have allowed the generation of large number of ancient DNA sequences and even complete extinct genomes (Hofreiter 2008; Rasmussen et al. 2010; Green et al. 2010). The advent of these techniques has involved the re-examination of the authenticity criteria in the light of the new data generated. For instance, it is now possible to directly observe the distribution length of the endogenous and contaminant sequences (with the only limit of the sequencing technique and the analytical cut-off) in Neandertals (Green et al. 2008). However, while the endogenous fragments have similar distributions (although differing on their mode and average length, depending on the sample), the contaminant sequences show rather different lengths among samples. For instance, in one Neandertal specimen (El Sidrón 1253), all contaminants are between 30-

60 bp in length, while in others (Feldhofer 1 and Vindija 33.16), they range from 30 up to >200 bp (Green et al. 2009).

On the other hand, the examination of the base composition at the 5' and 3'-ends of the 454-generated sequences, showed a fragmentation pattern not previously detected by conventional methods such as the PCR. In the endogenous Neandertal sequences, guanine and adenine residues are elevated relative to cytosine and thymine residues, immediately before the strand breaks (Briggs et al. 2007). Also, at the 5'-most sequenced positions, T shows a frequency above the average, while at the 3'-most sequenced positions, A is increased. This has been interpreted as the result of depurination gaps resulting in either thymines or cytosines in the complementary strand, being a significant proportion of the later posteriorly deaminated to uracils. During the blunt-end repairing at the genomic library preparation step, miscoding lesions at the 3'-overhanging ends are removed by the T4 DNA polymerase whereas miscoding lesions at the 5'-overhanging ends generate misincorporations at the end of the complementary strand (Briggs et al. 2007). If this interpretation is correct, then the primary cause of the fragmentation pattern observed is the DNA depurination.

These observations have been made in samples that share some traits, such as being Neandertals or coetaneous, having a low or negligible degree of human contamination and allowing the phylogenetic

discrimination of endogenous and contaminant sequences (Briggs et al. 2009; Green et al. 2009; Krause et al. 2010).

To test if the DNA depurination is the primary mechanism underlaying the detected fragmentation pattern, we have first analysed modern human contaminant sequences in an ancient non-human sample previously treated with bleach, a known depurinating agent. Second, we have explored the fragmentation pattern of recent human contaminants in a non-treated ancient non-human sample. Third, we sequenced two ancient human samples from different ages and taphonomic conditions: a recently found Neandertal sample and a Neolithic human sample to see if the observed pattern is consistent.

Material and methods

To identify human contaminants, two different non-human samples were subjected to the GS20-FLX 454 pyrosequencing platform at the Centre de Regulació Genòmica (CRG) in Barcelona: a *Myotragus* bone and an ancient Iberian lynx mandible. *Myotragus balearicus* was an extinct endemic caprine from the Balearic Islands (Lalueza-Fox et al. 2005). A *Myotragus* radius bone (IMEDEA 43619) excavated from Cova Estreta (Pollença, Mallorca) in 1996 and radiocarbon dated to about 6,300-5,700 years ago was analyzed (Ramirez et al. 2009). The Iberian

lynx (*Lynx pardinus*) is a critically endangered carnivore currently restricted to two isolated populations in the south of the Iberian Peninsula. However, until historic times its distribution was larger, reaching the North East of the Iberian Peninsula. A lynx mandible from Cova del Toll (Barcelona) dated to about 11,420 years ago, was analyzed.

To explore the pattern of human sequences found in an ancient modern human specimen, we chose a Neolithic tooth (inventory number CCG94-E33) belonging to the site of "Camí de Can Grau" (Granollers, Barcelona, Spain). This is a necropolis excavated in 1994, that comprised 23 tombs radiocarbon dated to between 5,500-5,000 years ago and that was previously studied by conventional, PCR-based methods (Sampietro et al. 2007). We knew that these samples, that were washed and handled without special precautions by the excavators, had a certain contamination background that was possible to quantify in about 17.1% on average of all the PCR-produced mtDNA sequences, as deduced by the genotyping of all the people involved in the manipulation of the specimens (Sampietro et al. 2006). However, the contaminating sequences seem to be lower (~5%) than the average figure in this particular sample (unpublished data).

Additionally, we have included a Neanderthal bone fragment from the El Sidrón site (Asturias, Spain), dated to about 49,000 years ago and labelled SD 1504. Previously, two mitochondrial DNA fragments were determined by PCR from this sample, and no contaminant sequences were

found among 115 clones generated (data not shown). Therefore, little or no contamination is suspected for this sample and thus, it can be taken as a positive control for the previously described fragmentation pattern.

To observe the possible fragmentation pattern the following protocol was used: starting with 454 raw files of the lynx, *Myotragus*, and the Neanderthal/Neolithic human shotgun sequencing, sequence reads were analysed to check the presence or absence of both (5' and 3') adaptor tags. Only the reads containing both adaptor sequences were keeped to this analysis, to avoid possible missassignment of the sequence ends. The remaining reads were aligned using NCBI Blast with the megablast algorithm, against the non-redundant nucleotide database (nt). Only blast hits with a coverage over 98% were keeped, and, in case of uncertanty in the blast assignment, the higher bit score hit was selected. The blast output was assigned to a specific taxonomic level, using the NCBI taxonomy database. TaxID assigned blast outputs were backtracked to the order level, and filtered keeping only the sequences that matched to Primate taxa. To avoid uncertainties in finding the fragmentation point in the sample sequence, the remaining 454 reads were extended 10 nt up- and downstream, using the GI reference sequence and the extreme up-/downstream 21 nucleotides to analyze the fragmentation pattern. Nucleotide frequencies for the resulting 21 positions were calculated.

Results

Human contaminant sequences in animal bones

Among the 96,357 sequences obtained from *Myotragus*, 337 (corresponding to 0,35% of the total) could be identified as human sequences (more than the endogenous ones, accounting for only 260 reads or 0.27%). These human sequences obviously corresponding to paleontologists that washed and cleaned the remains after its discovery in 1996, inadvertently contaminating the specimen. The average length of these sequences is 85 nucleotides, and they range from 30 bp (determined by the length cut-off in the analysis) to 300 bp (limited by the GS-FLX technology) (Figure 1). The distribution length of the contaminants is similar in shape to other ancient DNA distributions produced by shotgun methods (Green et al. 2008; Briggs et al. 2009; Krause et al. 2010). The bone powder was treated with bleach prior to DNA extraction as a decontaminating procedure (Ramírez et al. 2009). Consequently, the fragmentation pattern in both the endogenous and the human contaminant sequences is concordant with the suggested depurination pattern (Briggs et al. 2007) (Figure 2).

Among the 361,151 sequences obtained from the lynx sample, 5,078 (1,4% of the total) are human contaminants, again a figure significantly higher than the endogenous Carnivore sequences, that

account for only 414 (0.12%) of the total reads. Despite previous observations that recent contaminants on Neandertals show no increase in frequency of purines or pyrimides in either side of the DNA fragment (Krause et al. 2010), contaminants in the lynx sample do not seem to be fragmented at random. Moreover, they show a pattern difficult to interpret but clearly discernible (Figure 3). In fact, the 5' end pattern is very similar to that observed for Feldhofer 2 and Teshik Tash contaminants (Krause et al. 2010, supplementary information), in which there is an increase of G ratio at the breaking point, an increase of A at +1 position and an increase of T at +2 position. In the lynx human contaminants, there are other additional features, such as an increase of T prior to the breaking point (Figure 3).

Neolithic sequences

In the Neolithic sample, 168,998 reads were generated, from which 1,117 (0.66%) were human sequences. The efficiency of DNA retrieval is low (<1%), but is coherent with that obtained from *Myotragus*, a sample from a similar thermal environment and age (Ramírez et al. 2009). The figure is also similar to ratios observed in older specimens, including some Neanderthals (Green et al. 2008; Briggs et al. 2009).

In the human Neolithic sequences, cytosine residues at the 5'-ends appear as thymines in an increased frecuency (Fisher's exact test, $P<0.05$),

while guanines at the 3'-ends appear as adenines in most of the cases ($P<0.05$). Immediately before to the 5'-end strand break, purine bases are elevated, while pyrimidines are decreased. The opposite pattern is observed after the 3'-end strand break, that is, pyrimidines are increased, while purines are depressed. It must be remembered that the 3'-ends correspond to the terminal 5'-position in the complementary strand, and that the whole pattern is consistent with fragmentation at purine sites (Briggs et al. 2007).

Neandertal sample

In the El Sidrón Neandertal sample, 155,676 reads were generated, from which 503 (0.32%) were primate sequences. The efficiency of DNA retrieval is again very low (<1%), but it is almost identical (0.27%) to that obtained with another sample from the same site (Briggs et al. 2009). The purine fragmentation pattern previously described in ancient sequences can be seen, although the sample size is lower than that analysed in previous neandertal studies.

Discussion

The general efficiency ratios obtained for these four samples from a temperate environment and different ages (ranging between 49,000 and 5,500 years ago) and heterogeneous taphonomic conditions are similar

one each other and generally lower (<1%) than those published in samples from colder environments (Noonan et al. 2005; Poinar et al. 2006; Blow et al. 2008; Green et al. 2010). Even so, our results confirm the possibility of obtaining extinct genomes even in places thermally unfavourable to DNA preservation (Ramírez et al. 2009). We note that the non-human samples show a significant presence of human contaminant sequences (0,35 in *Myotragus* and 1,4% in the lynx) that in both cases outnumber the endogenous ones, pointing out the difficulties of working with ancient human specimens, specially those without a detailed handling history. It is probably better to work only with newly excavated remains, extracted under controlled conditions. Even with this limitation, the potentiality of high-throughput DNA sequencing platforms still makes possible the retrieval of ancient human genomes, even from temperate environments.

The analysis of the distribution lengths indicates that endogenous and contaminant sequences can display significantly different distributions (Figure 1). We have applied the Kruskal–Wallis test to test the difference between multiple samples and to know wether the underlying population distributions are non-normal or unknown. The null hypothesis (H_0) according to which *k* independent samples (*Myotragus*, Neandertal, Lynx, and Neolithic) were drawn from the same population (or identical populations), was tested against the alternative hypothesis (H_1), according to which these samples were drawn from populations

sharing the same shape but with different central tendencies (medians). The results of the Kruskal-Wallis chi-squared test was 1095.290 (degrees of freedom = 3) with a p-value < 2.2e-16. Therefore, the length distribution of the sequences is different among the samples considered. Of course, the pattern of size reduction could be related to factors such as the age of the sample or contamination ratios. However, previous suggestion that ancient DNA is fragmented to smaller lengths than contaminants is not a helpful authentication criterion *per se*, due to the wide overlapping of the lengths' distribution.

The human sequences from the neolithic and neandertal sample, as well as the caprine and carnivorous sequences from the *Myotragus* and the lynx, respectively, do show the purine-associated fragmentation pattern previously described as a feature of endogenous Neandertal sequences. There are some fluctuations in the frequencies beyond the fragmentation point due to the small number of reads available, specially in the lynx. Also, the pattern is always stronger at the 5' end, because sequences without the adaptor at the 3' end were discarded in the analysis. However, the pooled endogenous sequences, both human and non-human, clearly display the described depurination pattern, specially at the 5' end.

The human contaminant sequences show a fragmentation pattern different to that observed in endogenous sequences and similar to that found in other contaminanted samples (Krause et al. 2010). We speculate

that this complex pattern could be the result of the contaminant sequences being fragmented by enzymatic processes associated to bacterial action. The combination of several, most frequent bacterial endonucleases, could generate this particular pattern. Further research in this issue is needed, specially in the comparison between recent and old (>10 years) contaminants in a given sample.

We provide evidence that recent contaminants in bleached samples (a usual decontaminating procedure in ancient DNA) are fragmented following depurinating processes identical to those suffered by ancient sequences. Therefore, it is strongly recommended not to bleach human bones assigned to be subjected to shotgun sequencing, because the resulting contaminant sequences will imitate the ancient ones in its fragmentation pattern.

In any case, our results show that the fragmentation pattern is not exclusive of Neandertals or other Late Pleistocene samples and that it could be used to authenticate ancient human DNA sequences. It also confirms that contaminants do not show a simple depurinating-associated pattern. However, it is not clear if both sets of sequences (endogenous and contaminants) can be distinguished based on the fragmentation pattern alone. Further research on these two patterns is needed, since it is crucial for the possibility of working on ancient human samples even if no proper anticontamination measures have been taken during the excavation.

Acknowledgments

We are indebted to Josep Antoni Alcover (IMEDEA), to Antonio Rosas (MNCN), to Marco de la Rasilla (Universidad de Oviedo) and to Roser Pou and Miquel Martí (Museu d'Arqueologia de Granollers) for providing us with the *Myotragus*, the Neandertal and the Neolithic samples, respectively. C.L.-F. and S.C. are supported by a grant from the Ministerio de Innovación y Ciencia from Spain (BFU2009-06974).

Literature Cited

Blow MJ, Zhang T, Woyke T, Spelle CF, Krivoshapkin A, Yang D, Derevianko A, Rubin EM (2008) Identification of the source of ancient remains through genomic sequencing. *Gen Res* 18(8):1347-1353.

Briggs AW, Stenzel U, Johnson PL, Green RE, Kelso J, Prüfer K, Meyer M, Krause J, Ronan MT, Lachmann M, Pääbo S. 2007. Patterns of damage in genomic DNA sequences from a Neandertal. Proc. Natl. Acad. Sci. USA 104(37): 14616-14621.

Briggs AW, Good JM, Green RE, et al. (18 co-Authors). 2009. Targeted retrieval and analysis of five Neandertal mtDNA genomes. Science 325(5938): 318-321.

Cooper A, Poinar HN. 2000. Ancient DNA: do it right or not at all. Science 289: 1139.

Fortea J, de la Rasilla M, García-Tabernero A, Gigli E, Rosas A, Lalueza-Fox C. 2008. Excavation protocol of bone remains for Neandertal DNA analysis in El Sidrón Cave (Asturias, Spain). J. Hum. Evol. 55(2): 353-357.

Green RE, Malaspinas AS, Krause J, et al. (25 co-Authors) 2008. A complete Neandertal mitochondrial genome sequence determined by high-throughput sequencing. Cell 134(3): 416-426.

Green RE, Briggs AW, Krause J, Prüfer K, Burbano HA, Siebauer M, Lachmann M, Pääbo S. 2009.The Neandertal genome and ancient DNA authenticity. EMBO J. 28(17): 494-502.

Green, R.E., Krause, J., Briggs, A.W., Maricic, T., Stenzel, U., Kircher, M., Patterson, N., Li, H., Zhai, W., Fritz, M.H., Hansen, N.F., Durand, E.Y., Malaspinas, A.S., Jensen, J.D., Marques-Bonet, T., Alkan, C., Prüfer, K., Meyer, M., Burbano, H.A., Good, J.M., Schultz, R., Aximu-Petri, A., Butthof, A., Höber, B., Höffner, B., Siegemund, M., Weihmann, A., Nusbaum, C., Lander, E.S., Russ, C., Novod, N., Affourtit, J., Egholm, M., Verna, C., Rudan, P., Brajkovic, D., Kucan, Z., Gusic, I., Doronichev, V.B., Golovanova, L.V., Lalueza-Fox, C., de la Rasilla, M., Fortea, J., Rosas, A., Schmitz, R.W., Johnson, P.L., Eichler, E.E., Falush, D., Birney, E., Mullikin, J.C., Slatkin, M., Nielsen, R., Kelso, J., Lachmann, M., Reich, D., Pääbo, S. (2010). A draft sequence of the Neandertal genome. *Science* 328: 710-722.

Helgason A, Pálsson S, Lalueza-Fox C, Ghosh S, Sigurðardóttir S, Baker A, Hrafnkelsson B, Árnadóttir L, Þorsteinsdóttir U, Stefánsson K. 2007. A statistical approach to identify ancient template DNA. J. Mol. Evol. 65: 92-102.

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

Hofreiter M, Serre D, Poinar HN, Kuch M, Pääbo S. 2001b. Ancient DNA. Nat. Rev. Genet. 2(5): 353-359.

Hofreiter M. 2008. Paleogenomics. C. R. Palevol.7: 113-124.

Krause J, Briggs AW, Kircher M, Maricic T, Zwyns N, Derevianko A, Pääbo S. 2010. A complete mtDNA genome of an early modern human from Kostenki, Russia. Curr Biol. 20(3): 231-237.

Lalueza-Fox C, Castresana J, Sampietro L, Marqués-Bonet T, Alcover JA, Bertranpetit J. 2005. Molecular dating of caprines using ancient DNA sequences of *Myotragus balearicus*, en extinct endemic Balearic mammal. BMC Evol. Biol. 5: 70.

Malmström H, Svensson EM, Gilbert MTP, Willerslev E, Götherström A, Holmlund G. 2007. More on contamination: the use of asymmetric molecular behavior to identify authentic ancient human DNA. Mol. Biol. Evol. 24(4): 998-1004.

Noonan JP, Hofreiter M, Smith D, Priest JR, Rohland N, Rabeder G, Krause J, Detter JC, Pääbo S, Rubin EM (2005) Genomic sequencing of Pleistocene cave bears. *Science* 309(5734):597-599.

Poinar HN, Schwarz C, Qi J, Shapiro B, Macphee RD, Buigues B, Tikhonov A, Huson DH, Tomsho LP, Auch A, Rampp M, Miller W, Schuster SC (2006) Metagenomics to paleogenomics: large-scale sequencing of mammoth DNA. *Science* 311(5759):392-394.

Ramírez O, Gigli E, Bover P, Alcover JA, Bertranpetit J, Castresana J, Lalueza-Fox. C. 2009. Paleogenomics in a temperate environment : shotgun sequencing from an extinct Mediterranean caprine. PloS One. 22;4(5): e5670.

Sampietro ML, Gilbert MPT, Lao O, Caramelli D, Lari M, Bertranpetit J, Lalueza-Fox C. 2006. Tracking down Human Contamination in Ancient Human Teeth. Mol. Bio. Evol. 23(9): 1801-1807.

Sampietro ML, Lao O, Caramelli D, Lari M, Pou R, Marti M, Bertranpetit J, Lalueza-Fox C. 2007. Palaeogenetic evidence supports a

dual model of Neolithic spreading into Europe. Proc. R. Soc. B. 274(1622): 2161-2167.

Vives S, Gilbert MT, Arenas C, Gigli E, Lao O, Lalueza-Fox C. 2008. Statistical analysis of *post mortem* DNA damage-derived miscoding lesions in Neandertal mitochondrial DNA. BMC Res. Notes 1: 40.

Figure 1: Sequence length distribution of the sequences obtained from a bone sample of *Myotragus balearicus*, a Neandertal, an ancient Lynx and a Neolithic human by 454 pyrosequencing.

Figure 2: Nucleotide base frequencies at the end of the *Myotragus* human contaminants, treated with a depurianting agent, bleach. The base composition is plotted as a function of distance from the 5'-ends and 3'-ends.

Figure 3: Nucleotide base frequencies at the 5' end of the human contaminant sequences in the *Lynx* sample. The breaking point position is at nucleotide 10.

Figure 1

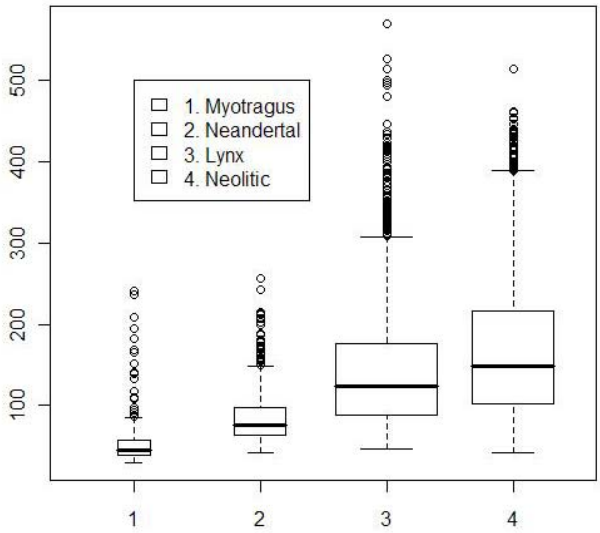


Figure 2

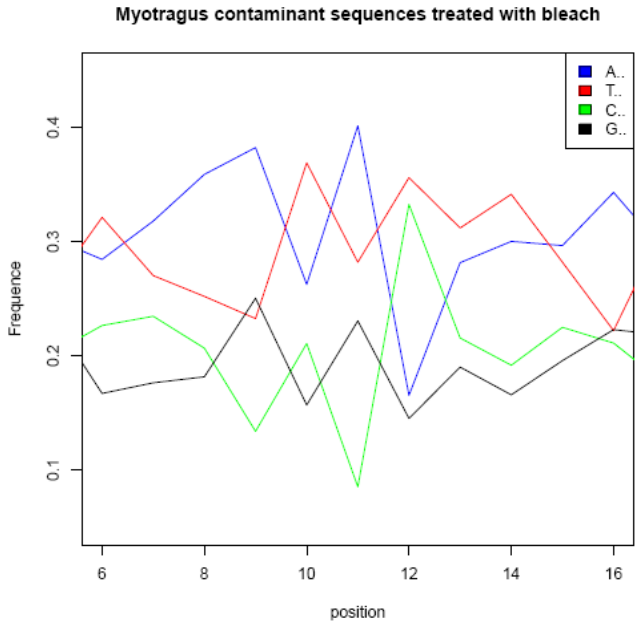


Figure 3

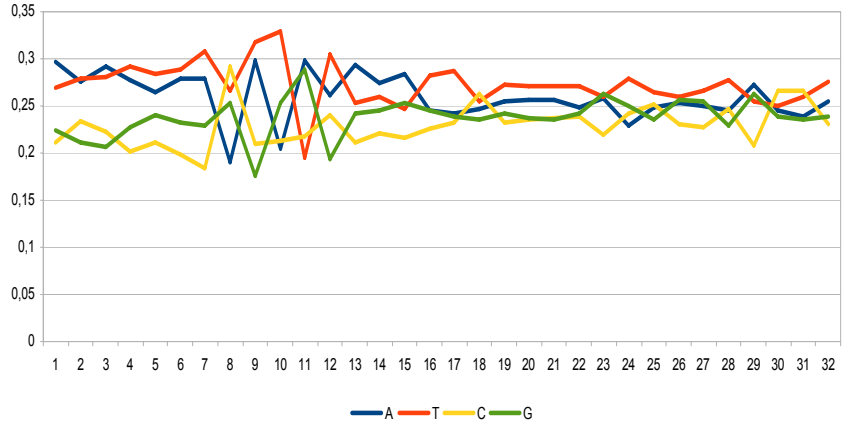

0,35
0,3
0,25
0,2
0,15
0,1
0,05
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32
A
T
C
G

Table 1: Specimens subjected to GS20-FLX 454 pyrosequencing, number of reads and ratio of endogenous and contaminant sequences obtained. *: in the case of human samples, it is impossible to discern *a priori* which sequences are endogenous and which human contaminants. However, we have mitochondrial DNA estimates that rate the maximum potential contamination <1% in the neandertal sample and <5% in the neolithic sample.

Specimen	**N reads**	**% endogenous seq**	**% human seq**
Neandertal	155,676	0.32% (N=503)	*
Neolithic	168,998	0.66% (N=1,117)	*
Lynx	361,151	0.11% (N=414)	1.4% (N=5,078)
Myotragus	96,357	0.27% (N=260)	0.35% (N=337)

3.4. Genetic characterization of the ABO blood group in Neandertals.

Carles Lalueza-Fox, **Elena Gigli**, Marco de la Rasilla, Javier Fortea, Antonio Rosas, Jaume Bertranpetit and Johannes Krause. (2008) BMC Evolutionary Biology 8:342

3.5. Bitter taste perception in Neanderthals through the analysis of the TAS2R38 gene.

Carles Lalueza-Fox, **Elena Gigli**, Marco de la Rasilla, Javier Fortea and Antonio Rosas. (2009) Biology Letters 5:809-11

4. DISCUSSION

In the current thesis we have addressed two different, related topics. First we have focussed on the development of new strategies for trying to avoid pre-laboratory and laboratory-derived contaminants, respectively. Second, we have analyzed two nuclear genes involved in adaptive traits of the Neanderthals.

4.1. Anticontamination procedures

The presence of DNA in ancient remains was initially shown by staining of DNA in histological samples of Egyptian mummies (Pääbo 1984) and by extraction and cloning in bacterial plasmids of DNA from the extinct quagga and mummies (Higuchi et al. 1984; Pääbo 1985). However, it was only the advent of the polymerase chain reaction (PCR) (Mullis and Faloona 1987) that made it possible to reproduce results. Through repeated PCRs it became also possible to ensure that the sequence determined did not contain errors (Pääbo and Wilson 1988).

However, the field is still marred by an underestimation of the extent of contamination within laboratories and samples themselves (Willerslev and Cooper 2005). As a remedy, a list of 'criteria of authenticity' for the study of ancient DNA were suggested (Pääbo et al. 1989) and a set of laboratory practices to prevent contamination have been developed over time (Cooper and Poinar 2000; Hofreiter et al. 2001b; Pääbo et al. 2004).

The aDNA studies demonstrated that the genetic material surviving in ancient specimens was often microbial or fungal in origin (Green et al. 2010), and that endogenous sequences were limited to very low concentrations, depending on the sample (Higuchi et al. 1984;

Pääbo et al. 1985, 1989). The main problem of aDNA studies highlighted by using the PCR (Pääbo 1989) and high-throughput DNA sequencing technologies (Green et al. 2010) is contamination and in the case of hominins, that is, modern humans and their close relatives such as Cro-Magnons, Neanderthals, “Denisovans” and *Homo floresiensis* (Brown et al. 2004), the issue of authenticity is particularly acute because they can be expected to be identical to current humans for much or almost all of their genome (Green et al. 2009; 2010).

In this regard, a problematic and harder issue to resolve is when the analyzed specimen has been contaminated prior to preparation for genetic analysis (Pääbo et al. 2004; Bandelt et al. 2005; Willerslev and Cooper 2005; Sampietro et al. 2006; Gilbert et al. 2006). Usually, human archaeological material is handled, washed, and labeled during excavation and it seems that at this stage there is a strong contamination with exogenous DNA. Also is very likely that pre-laboratory contamination can occur at any stage prior to genetic analyses, including the anthropological study (Sampietro et al. 2006).

Probably, contaminants enter the specimens by direct handling and washing, presumably due to DNA derived from the handler permeating throughout dentinal tubules into the pulp cavity (in teeth) and the Haversian system (in bone) (Rudbeck et al. 2005; Sampietro et al. 2006), although possibly not permeating as far as the osteocytes (Malmstrom et al. 2005). Therefore, washing the specimens is a critical step for contaminating the samples (Gilbert et al. 2006). The exact mechanism of contamination by handling is still poorly understood, but the transport of DNA molecules from the

exogenous source to the interior of the specimen seems to be related to sample moisture content. In fact, it has been argued that once the remains are excavated and dried, mechanisms such as collapse of collagen bundles within bones and teeth or sedimentation and precipitation of minerals in water might block further entry of waterborne contaminant DNA deep within the sample (Gilbert et al. 2006, Sampietro et al. 2006; Rudbeck, et al. 2005).

Human pre-laboratory contaminant has been reported in some studies (Hofreiter et al. 2001; Malmstrom et al. 2005). In addition, this type of contaminant DNA sequences have been identified in Neanderthal remains to belong to some of the archaeologist that previously had studied the samples (Serre et al. 2004; Lalueza-Fox et al. 2005). However the biggest problem is when endogenous and exogenous sequences are indistinguishable (Green et al. 2009, Krause et al. 2010). This could be the case when working with ancient human remains that has been excavated by archaeologist under uncontrolled condition (Sampietro et al. 2006).

In the first study of this PhD work, we have tried to solve this issue through the developed an anti-contamination protocol for the Neanderthal site of El Sidron (Spain) (Fortea et al. 2003; Rosas et al. 2006). The El Sidron is a karstic system located in the region of Asturias, northern Spain, a significant archaeo-paleontological record with >1,900 Neanderthal fossil remains (labeled SD or in some case SDR) and 358 lithic artifacts has been recovered at the site since 2000 (Fortea et al. 2003; Rosas et al. 2006).

The main purpose of our work was to test the anti-contamination protocol effectiveness through DNA analysis of samples retrieved before and after the protocol was implemented. For the analysis we selected three fragments of femur from at least two different individuals (SD 441, SD 1161, SD-1252). These small bone fragments look macroscopically well-preserved but they are morphologically uninformative. The samples were extracted under the anti-contamination protocol that is done in full lab gear, which includes a coverall (Tyvek), head cap (Tyvek), shoe covers (Tyvek), medical face mask, face shield (Hucoa-Erloss), and sterile latex gloves. The excavation blades and brushes are regularly bleached with 30% sodium hypochlorite solution prior to and after use. The excavated remains are not washed; instead, they are put into plastic bags and immediately frozen at -20°C until they arrive at the laboratory. Taking into account that a large proportion of contamination may results from samples handling prior to arriving in laboratory, we extracted and typed mtDNA from blood drops from the archaeologists, paleontologists, and geologists involved in the excavation of El Sidron, as well as from the people involved in the laboratory analyses. To indirectly test the effectiveness of our anti-contamination protocol we compared the genetic data obtained from these ‘‘clean’’ specimens to that obtained from remains, such as SD-441 and SD-1161, retrieved before the protocol was implemented (Lalueza-Fox et al. 2005). Two are the main results that we obtained:

First: the majority of the contaminant sequences found in the ‘unclean’ specimens (SD-441 and SD-1161) and in SD-1252, extracted under controlled conditions, were consensus reference sequence (CRS: Cambridge Reference Sequence) and cannot be

attributed with certainty to any of the researchers involved. In the only two cases in which it was possible to attribute the contaminant sequences to particular researchers, both were the archaeologists that handled and washed the remains at the site prior to the implementation of the excavation protocols. This result corroborates previous findings that washing and cleaning are critical steps for contamination (Sampietro et al. 2006).

Second: the application of the anti-contamination protocol appears to have helped to increase the endogenous versus contaminants ratio; the value of the pre-protocol SD-441 is 5.17 % whereas the post-protocol SD-1252 ratio is 69.6 %.

There are few protocols that draw guidelines on action to be taken before the specimens arrive at the laboratory. Sampietro et al. (2006) suggested to monitor archaeologists, anthropologist and geneticists who had access to the samples. Unfortunately this is not always possible. In fact at the best, the samples handled by only a limited and well-known people. Very often, especially for samples excavated from many years, the genetic information of archaeologists cannot be traced back and therefore it is impossible to identify a possible source of contamination. Pruvost et al. (2006) analyzed samples which had been excavated several years before and had been subsequently brushed with water, dried, and stored under light exclusion conditions in collections at room temperature and fresh samples originated from the same archeological site but had been recently excavated and not handling. They recommend not to treat or wash the samples and store them immediately to -20°C for a better preservation of the surviving DNA. We generally agree, but we consider that not washing the samples could

complicate the molecular analyses. In fact not eliminating waste of land, fulvic and humic acids contained in it, could block the action of *Taq polymerase* causing a failure of PCR resulting in a false negative. A compromising solution would be to wash the samples with sterile water once at the ancient DNA laboratory.

Also, is clear that DNA preservation in ancient bones must be influenced by many different parameters such as pH and temperature. It is well known that temperature increase and decrease in pH induce further degradation of DNA. Taphonomic conditions may vary within archaeological sites determining different levels of DNA preservation. To assess the possible presence of DNA in the samples, a good method is to evaluate amino-acid racemization values for aspartic acid. In SD-441 amino-acid racemization values were higher than for SD-1252 suggesting that the former was subjected to worse biochemicall conditions. Also, as stressesed by Sampietro et al. (2006) different types of samples could show different levels of contamination. In fact human teeth remains seem to be more susceptible to contamination during the excavation and washing respect to bone samples. In this light the observed difference could not be the exclusive results of our protocol.

The results of our work suggest an important guideline to control pre-laboratory contaminants. Through the use of coverall, head cap, shoe covers, medical face mask, sterile gloves, regular bleaching with sodium hypochlorite solution of blades and brushes and the use of plastic bags to recover the samples before to have them frozen at -20°C, is possible to decrease the amount of contaminants. The application of this protocol will be of

fundamental importance especially when endogenous and exogenous sequences are expected to be difficult to discriminate.

Pre-laboratory contamination is not the unique problem. In fact, historically, the field of aDNA has also focused on the problems associated with laboratory-derived contamination. A number of guidelines have been suggested to help deal with this issue. Examples include independent replication of results in a second laboratory, separation of pre- and post-PCR laboratories, adherence to sterile techniques, blank controls in amplifications and extractions, cloning of PCR products, quantification of the number of DNA templates, irradiation with UV light and wiping of objects and equipment with hypochlorite solution (bleach) (Handt et al. 1996; Cooper and Poinar 2000; Sampietro 2006; Champlot et al. 2010). But despite these precautions, to work with ancient humans, where contaminants and endogenous are often indistinguishable, can be complicated (Caramelli et al. 2003). As evidenced by Green et al. (2009) and Rasmussen et al. (2010), high-throughput sequencing studies of ancient genomes are not exempt from this problem (Champlot et al. 2010).

In the second study of this PhD work, we tried to develop a new methodological approach, to decrease the laboratory contaminations. This technique is an alternate PCR-based method, based in the incorporation of “blocking primers” (Vestheim and Jarman 2008). These primers are designed to preferentially bind to contaminant sequences, and subsequently block their amplification during the PCR step. For the analysis we have selected four El Sidron Neanderthal specimens that have been excavated between 2004 and 2006. The samples are: dentine root fragments from and

adult incisor (SD 441) (Lalueza-Fox et al. 2005), an adult tooth (SD 1161) and two adult female femur fragments (SD 1253 and SD 1351c) (Krause et al. 2007; Lalueza-Fox et al. 2008; 2009).

We designed two Neanderthal primer pairs that amplify DNA segments of different length. The first Neanderthal specific shorter amplicon (16,230–16,262) produces a 70 bp product (including primers) and the primers contain 3 and 4 mismatches relative to the human reference sequence in the L and H primers, respectively. The second Neanderthal specific primer set (16,244–16,301) amplifies a longer product of 100 bp (including primers) and the primers bind with less specificity (3 and 2 mismatches in the L and H primers, respectively). The blocking primers are human primers that have been modified at the 3' end with a C3 spacer to prevent the *Taq polymerase* from extending it once they are annealed to the targeted DNA. After the extraction (Lalueza-Fox et al. 2005; 2006), in the PCR reaction mix, a pair of Neanderthal specific primers and human-specific blocking primers, were mixed together. The Neanderthal primers are regular PCR primers that anneal and extend normally under the appropriate PCR conditions. We expected that the blocking primers will bind to the human contaminants preventing their amplification and that specific Neanderthal primers normally amplified the Neanderthal sequences. The same samples analyzed in this work were previously used to develop the anti-contamination protocol so the contamination level was, a priori, known. Our result shows that:

First: the blocking primers approach is useful to eliminate a large proportion of known contaminant sequences from ancient samples. In fact, analysis of the cloned DNA sequences indicates that the

efficiency of Neanderthal sequence recovery for the short (16,230–16,262) amplicon was increased (from 5% to 75.3% for the SD 441 sample and from 69.6% to 100% in the highly contaminated SD 1161 sample). Also the cloned sequences of the longer (16,244–16,301) fragment showed an increase (from 16% to 95.7% in the SD 1253 sample and from 10.3% to 89.7% in the SD 1351c). It could be also possible to improve the efficiency of the Blocking Primers through titration, using End-Point Titration PCR (ET-PCR) or flow-cytometry-based titration method, adjusting the conditions for each particular specimen.

Second: with the blocking primers, we have been able not only to successfully retrieve non-contaminated Neanderthal sequences, but also to amplify longer products than with the traditional PCR approach by eliminating the competing contaminant background in the PCR. Briggs et al. (2009) showed that the long PCR targets a fragment that is longer than the average template DNA in most Neanderthal samples. The same authors analyzed the sample SD-1253, determining by pyrosequencing that the average fragment length is only 51.3 bp (Briggs et al. 2009). Also, some works (Green et al. 2007; Krause et al. 2010) point out that contaminant sequences are longer that endogenous sequences. This means that amplifying >100bp fragment from an ancient sample, the probability of obtaining only human contaminants is very high. Thus, blocking the amplification of contaminant sequences can allow the retrieval of longer endogenous sequences.

During their taphonomic history, an ancient sample is subjected to contamination. Contaminants may come into contact with the sample during handling phases such as archaeological excavation,

cleaning and morphological study. In addition, other source of contamination may be the equipment and reagents used for laboratory analyses. In our work, SD-1253 showed low level of contamination that could reflect favourable condition of preservation, but also played a key role the application of anti-contamination protocol (5.1) that allowed a minor contamination due to handling during the pre-laboratory phases. In this sense, the only two contaminant sequences retrieved in the SD 1253 amplification with blocking primers, display the 16,294T–16,296T haplotype that is carried by one of the laboratory researchers (C.L.-F.). Thus, we have to exclude the handling contamination during excavation. We can guess that in this case the contaminants probably derived from the laboratory.

The application of Blocking Primers techniques could be useful to reanalyze ancient samples that previously have resulted in partial sequences or negative results. Also, as in El Sidron large fossil record, even partial sequences can be useful for attributing different bones to the same individual.

The developement of high-throughtput DNA sequencing associated with capture systems has allowed the analysis of large amounts of DNA and recently, entire genomes from Pleistocene specimens. In any case, the presence of human contaminants remains one major problem. In a future prospective it could be interesting to apply the Blocking Primers approach to Next Generation Sequencing platforms as a way to decrease the amount of exogenous sequences and to improve the authenticity of the results.

In the last years, high-throughput sequencing techniques (Margulies et al. 2005; Ondov et al. 2008; Bentley et al. 2008; Kircher et al. 2009) have allowed the generation of large number of ancient DNA sequences and complete extinct genomes (Rasmussen et al. 2010; Green et al. 2010). This new techniques led to a revision of the authenticity criteria in the light of new data generated (Green et al. 2009).

In fact is now possible to observe that the size distribution can help to make a first distinction between endogenous and exogenous sequences (Briggs et al. 2009). The mode of the size distributions of contaminating molecules seems to be higher than the endogenous DNA fragments and the longest sequences are routinely suppressed in the analyses (Krause et al. 2010). Also, as described in Krause et al. (2010) for the endogenous fragments an elevation of G and A in frequency was found at the 5'-end, whereas C and T were found elevated at the 3'-end; this is consistent with a preferential fragmentation of ancient DNA at purine bases (Briggs et al. 2007). In contrast, the human contaminants show no increase in frequency of purines or pyrimidines on either side of the fragment (Krause et al. 2010).

In the third work of this thesis we intend, first, to analyse modern human contaminant sequences in an ancient non-human sample previously treated with bleach; second, to explore the fragmentation pattern of recent human contaminants in a non-treated ancient non-human sample; and third, to sequence a Neanderthal and a Neolithic human sample from different ages and taphonomic conditions, to analyze if the fragmentation pattern is consistent between them.

To identify human contaminants we used two different non-human samples, a Myotragus (*Myotragus balearicus)* bone, dated 6,300-5,700 years ago and an Iberian Lynx (*Lynx pardinus*) mandible. To explore the pattern of human sequences found in ancient modern human specimen, we chose a Neolithic tooth belonging to the site of "Camí de Can Grau" (Granollers, Barcelona, Spain) dated to 5,500-5,000 years ago. During the recovery of these samples no anti-contamination protocol (5.1) was applied. For this reason it was previously possible to quantify the amount of contamination in about 17.1% on average of all the PCR-produced mtDNAs. Also, we have included a Neanderthal bone fragment from the El Sidrón site (Asturias, Spain), dated to about 49,000 years ago. This sample was previously analyzed and little or no contamination was detected and thus, it can be taken as a positive control for enodgenous fragmentation pattern.

To study the fragmentation pattern, lynx, *Myotragus*, Neanderthal and Neolithic samples were subjected to 454 (Life Science) massively parallel sequencing, and sequence reads were analysed to check the presence or absence of both (5' and 3') adaptor tags. Only the reads containing both adaptor sequences were taken for subsequent analyses. Later, reads were aligned using NCBI Blast with the megablast algorithm, against the non-redundant nucleotide database (nt). Only blast hits with sequence coverage over 98% were analyzed. Using the NCBI taxonomy database TaxID, assigned blast outputs were backtracked to the order level. At the end, to avoid uncertainties in finding the fragmentation point in the sample sequence, the remaining 454 reads were extended 10 nt up- and downstream, using the GI reference sequence and the

extreme up-/downstream 21 nucleotides to analyze the fragmentation pattern. Nucleotide frequencies for the resulting 21 positions were calculated.

The results show that human contaminants retrieved correspond to 0,35% and 1,4%, of the *Myotragus* and the lynx reads, respectively. The large amount of human contaminants in animal samples (outnumbering in both cases the endogenous sequences) once again demonstrates the importance of a proper sample handling and the application of anti-contaminant protocols to reduce the contaminants in pre-laboratory phases. Keeping in mind that, as shown by recent studies, the average length of exogenous sequences is longer than that of endogenous sequences, for the *Myotragus* specimen the distribution length of contaminants is similar to that of ancient DNA distributions produced by shotgun approach. This is probably due to the effect of the treatment with bleach, a strong depurinating agent that may be responsible for DNA fragmentation, both endogenous and exogenous.

As proposed by other authors, length distribution of the sequences can't be considered by itself a reliable criterion to discriminate contaminants from endogenous sequences. This is primarily because the use of aggressive substances, such as bleach, but also because the enzymatic action of microorganisms that could fragment the contaminant sequences along time. In any case, our results show that distribution length of human contaminants and endogenous sequences can display a significantly different distribution, although with a large overlapping that prevents the attribution of any particular sequence to one group or another.

Thanks to developed of high-throughput DNA sequencing technologies Briggs et al. (2007) discovered that substitutions resulting from miscoding cytosine residues were overrepresented in the DNA sequences and grouped in the ends of the molecules. The data showed that in the endogenous DNA molecules cytosine residues at the 5'-ends of DNA fragments appear as thymine residues and G-to-A substitutions were also seen at the 3'-ends. This fragmentation pattern is primarily the consequence of depurination processes that create abasic sites where the complementary DNA strand is posteriorly broken. Contaminant sequences apparently did not show this distribution. Therefore, the authors suggested that fragmentation pattern could be a helpful way to discriminate between endogenous and exogenous sequences. The results of the human contaminants in the Myotragus bone show an almost perfect depurinating-based fragmentation pattern, clearly produced by the use of bleach in this sample. The bleach is a product commonly used for decontaminating purposes in ancient DNA studies. Treating the contaminant DNA with bleach mimicks the fragmentation pattern of ancient DNA; thus exogenous and endogenous sequences can't be discriminated against based on the pattern present in the ends of the reads. In our opinion, it could be useful to substitute the bleach as a decontaminating procedure for alternative methods such as UV irradiation.

In the lynx sample -not treated with bleach- the human contaminant sequences show a vey complex fragmentation pattern that is different to that observed in endogenous sequences. It could be due to the combined action of different restriction enzymes present

in bacteria acting upon the bones. In any case, it seems difficult that particular sequences could be distinguished by the fragmentation pattern alone, although with the future analysis of more samples, both contaminated and uncontaminated, it could be possible to attribute a group of sequences to an endogenous or a contaminating origin with a given probability.

Another interesting result of this work was to explore the ratio of efficiency of DNA retrieval by shotgun sequencing in a temperate environment. Temperature is a crucial factor in the DNA preservation. Previous studies had stressesed that constantly low temperatures play a central role in the longevity of aDNA molecules while the high temperatures induce a faster DNA degradation resulting in a lower probability of finding genetic material in ancient specimens. We show that under favorable taphonomic conditions the discovery of endogenous DNA sequences in temperate environment is possible, although extremely inefficient. Furthermore, it should be noted that our study is structured on the basis of three components: climate (temperate, Mediterranean environment), age (ranging between 49,000 and 5,000 years ago) and heterogeneous taphonomic conditions (caves or burials). The results show that the general efficiency ratio obtained for the samples are similar one each other and lower (<1%) then those published in samples from colder environments.

In conclusion, the third work of this thesis highlights how the development of high-throughput DNA sequencing technologies has allowed identify a correlation between fragmentation pattern at 5'-ends and 3'-ends and endogenous sequences, not only in Neanderthals but also in more recent samples such as those from

the Neolithic. This approach could be helpful for identifying endogenous sequences *a posteriori* of the sequencing process, although further research is needed.

4.2. Adaptive traits

Since the discovery of the first Neanderthal specimen in 1856 in Kleine Feldhofer grotto (Germany) (King 1864; Krings et al. 1997), this group of Homininae has played a central role in scientific discussion of the later phase of human evolution.

The typical morphological features of Neanderthals first appear in the European fossil record about 400.000 years ago (Hublin 2009). Progressively more distinctive Neanderthal forms subsequently evolved until Neanderthals disappeared from the fossil record about 30.000-28.000 years ago (Finlayson et al. 2006). They lived in Europe and Western Asia as far east as Southern of Siberia (Krause et al. 2007) in a time period ranging principally from OIS 7 to OIS 3 (Hublin 2009).

One of the key questions is whether Neanderthals interbred with anatomically modern humans (Green et al. 2010). Information about the separation of lineages leading to *H. sapiens* and *H. neanderthalensis* can help to clarify the evolutionary relationship between them. However it is important to stress that the genetic coalescent, population separation and phenotypic differentiation of the two taxa can represent chronologically distinct events (Hublin 2009).

In recent years, thanks also to major technological improvements, (Margulies et al. 2005; Briggs et al. 2009; Burbano et al. 2010) extensive researches have been done on mitochondrial DNA of Neanderthal specimens (Green et al. 2008; Briggs et al. 2009;

Lalueza-Fox et al. 2010), The results of these works, strengthen the hypothesis of no interbreeding between *H. neandethalensis* and *H. sapiens*, but is not possible to exclude the possibility that Neanderthals contributed other parts of their genomes to present day humans (Green et al. 2010).

Considering that mitochondrial DNA is inherited maternally and recombination is either rare or absent (William 2008), is possible back up to the most recent common matrilineal ancestor (MRCA) (Gibbson 1998; Endicott et al. 2009) who the two Homininae groups have shared. Green et al. (2008) in order to estimate the date of the divergence of the Neanderthal and extant human mtDNAs, compared the Neanderthal mtDNA to 10 divergent human mtDNAs. They assumed, based on the fossil record, that humans and chimpanzees diverged between 6 (Galik et al. 2004) to 8 million years ago (Brunet et al. 2002; Lebatard et al. 2008). The results show an estimate of the mean divergence time between Neanderthal and extant human mtDNAs of 660,000 with a 95% credibility interval of 520,000 to 800,000 years ago, if no gene flow occurred

In contrast to mitochondrial DNA inherited, the nuclear genome is composed to of tens of thousands of recombining and hence independently evolving, DNA segments that provide an opportunity to obtain a clear picture of the relationship between Neanderthals and present-day humans. The genetic divergence (the proportion of nucleotides differing between representative individuals of the two species) time between two species varies substantially across the genome, conveying important information about the timing and process of speciation (Patterson et al. 2006). Green et al. (2010)

analyzed a large part of Neanderthal nuclear genome. Assuming an average DNA divergence of 6.5 million years between the human and chimpanzee genome the average divergence of Neanderthal and modern human autosomal DNA sequences is 825,000 years, and Neanderthal and present-day human populations separated between 270,000 and 440,000 years ago. The majority of the Neanderthal genetic diversity overlap with those of the humans reflecting the fact that Neanderthal fall inside of the variation of present-day humans sharing derived single-nucleotide polymorphism (SNP) alleles with present-day humans (Green et al. 2010).

A plethora of hypotheses are open to try to explain the genetic mixture involving Neanderthals. The more plausible models and compatible with Green et al. (2010) data are two: gene flow between Neanderthals and the ancestor of all non-Africans, probably occurred before the divergence of Europeans, East Asia and Papuan populations or an old structure in Africa that persisted from the origin of Neanderthals until the ancestors of non-Africans left Africa (Green et al. 2010).

Two works in this thesis intend to make their contribution to the analysis of Neanderthal nuclear genome. For this specific nuclear retrieval we have chosen single genetic variants in genes of interest that can provide specific information on aspects of the phenotype, immunity and perception of the Neandertals.

First: "Genetic Characterization of the ABO blood group in Neanderthals". In this study we analysed two male Neanderthals (SD 1253, SD 1351c) that were retrieved under controlled

conditions at El Sidron site (Asturias, Spain).Thanks to the application of anti-contamination protocol (5.1) they seem to have no modern human DNA contamination. We used a PCR approach to detect the presence of single informative substitutions in the ABO gene.

The ABO blood group systems consist in three alleles: two co-dominant A and B alleles, and one silent and recessive O allele (Mourant 1954). The system is controlled by a single gene at the ABO locus. This gene encodes a glycosyltransferase enzyme that adds a sugar residue to a carbohydrate structure known as antigen H that is present in the membrane of red cells. The allele A codes for an enzyme that adds a N-acetyl galactosamine to the H antigen, while B allele codes for an enzyme that adds a D-galactose. The O allele carries a human-specific inactivating mutation which produces a non-functional enzyme. In this case H antigen remains without modification on the surface of the cell. The combination of the three ABO alleles results in four major phenotypes, named A, B, AB, O, that carrying different antigens, react to specific antibodies (Mourant 1954).

The main A and B alleles (hereafter A101 and B101, respectively) differ by four amino acid changes at positions 176, 235, 266 and 268, with the two later being determining the A and B specificity of the enzyme. The most frequent among all human ABO alleles is the O allele that show a single G nucleotide deletion in position 261 of exon 6 (Δ261) that creates a premature stop codon resulting in a truncated protein with no enzymatic activity. The O allele can present two haplotypes O01 and O02. The O01 haplotypes differ from the O02 haplotypes by at least six nucleotide substitutions,

including position 297 in exon 6, which is A in the former and G in the latter.

Some studies stresses that the B101 allele, which gave rise to all the B variants, derived from the ancestral A101 allele around 3.5 million years ago (Mya), whereas the youngest and most common human allele O01, carrying the Δ261 deletion, derived from the A allele only about 1.15 ± 0.2 Mya determined by coalescence analysis.

Our results indicated that SD 1253 and SD 1351c showed the Δ261 deletion, which defines the O alleles, and the sequences display an A at position 297 on exon 6, which is compatible with the most frequent O haplotype, the O01.

We concluded that with all probability, the two El Sidrón Neanderthal individuals were most likely homozygous for the O01 allele and the data suggest the presence of the human O01 allele already in the common ancestor of Neanderthals and modern humans. Therefore, it is confirmed an emergence of the O01 allele more than 1 Mya predating the divergence of the modern human and Neanderthal populations.

In addition to, it seems to exist an association between the ABO alleles and a variety of infectious diseases. It has been suggested, for instance, that the O allele protects against severe malaria. At the same time, it can be more sensitive to *Helicobacter pylori* infections and to severe forms of cholera. The complex pattern of putative selective agents favouring or acting against different alleles could explain the maintenance of the high ABO

polymorphism, as evidenced by the signal of balancing selection detected on the gene. We don't know which diseases could be affecting the Neanderthal populations in the past, but in any case, this is the first analysis were information on the immunity is provided, thus establishing a connection between the genotype and the environment that in the future can furnish more information about Neanderthal paleodiseases.

Second: "Bitter taste perception in Neanderthals through the analysis the TAS2R38 gene". For our study we used a Neanderthal sample labelled SD-1253. It was retrieve under controlled conditions, applying anti-contamination protocol (5.1), in El Sidron site (Asturias). After the its extraction and the two step multiplex PCR (Krause et al. 2006), the PCR product was massively parallel sequenced using the 454 GS-FLX platform (Margulies et al. 2005), instead of subjected to cloning. The scope of this work is to investigate the presence/absence of the perception of bitter taste in the SD 1253 Neanderthal specimen, through the analysis of one allele at the *TAS2R38* gene.

Bitter taste is mediated by G-protein-coupled receptors expressed in taste cells on the surface of the tongue. These proteins are encoded by the TAS2R gene family. The *TAS2R38* gene has three amino-acid changes in high frequencies that determine only five main haplotypes. These polymorphisms are found at position 49 (encoding proline or alanine), 262 (alanine or valine) and 296 (valine or isoleucine) of the gene. The two most common haplotypes are proline–alanine–valine (PAV) and alanine–valine–isoleucine (AVI). PAV is the major taster haplotyp and AVI is the major non-taster haplotype

Our results show that a total of 4307 sequences were generated for the *TAS2R38* gene F142-R166 fragment. Of the total, 2391 (55.51%) showed a C in nucleotide position 145, corresponding to a proline amino acid (taster haplotype), and 1916 showed a G (44.49%), corresponding to an alanine amino acid (non-taster haplotype). Because almost half of the *TAS2R38* clones showed either one or another nucleotide, it is most probable that the El Sidron 1253 individual was A49P heterozygous. Humans and chimpanzees shared the balanced polymorphism associated with PTC tasting ability, but has recently been stresses by some studies, that this trait probably is controlled in chimpanzees by two alleles at the *TAS2R38* locus that are not present in modern humans. The non-taster alleles have evolved at least twice during hominid evolution.

From modern sequence data, the divergence time of the two common *TAS2R38* haplotypes has been estimated to be approximately 1.5 million years (Wooding et al. 2004). The most reasonable explanation for the observation of a Neanderthal A49P heterozygote is that this polymorphism effectively pre-dates the split of Neanderthal and modern human lineages.

Considering that the 25% of world population is tasteless despite the fact that the sense of bitter taste protects from ingesting toxic substances. Balancing natural selection seems to be maintaining divergent alleles, perhaps favouring heterozygotes. We can speculate that the presence of non-bitter taste allele could be indicative that non-genetic factors may have more influence on dietary choice than genetics (Gorovic et al. 2011). Also, the

presence of polymorphism in this gene may be related to the vegetable intake in Neanderthals' diet that has been traditionally described as highly carnivorous. Thus, this study suggests that Neanderthals had a more varied and complex dietary influences that previously assumed.

The times when species split and the recombination rates are particularly high when gene genealogies vary along the genome. A gene genealogy may be different from the species phylogeny (i.e., gene tree differ from the species tree). This phenomenon is called incomplete lineage sorting (ILS) (Hobolth et al. 2011). Genomic analysis of the ILS patterns are used to infer demographic population processes in the ancestral species. A region under purifying selection can cause reduction in polymorphisms in a linked neutral region and it results in a smaller amount of ILS. Positive directional selection on a gene can be a similar but broader effect. Balancing selection, on the other hand is expected to generate a higher amount of ILS (Holbolth et al. 2011). The effective population size also influences these patterns and we know from mtDNA variation that the female effective population size was really small in Neanderthals (Briggs et al. 2009). A small population size can involve the extinction of particular alleles in either the Neanderthal or the modern human lineages. In the case of the ABO blood group, we don't know yet if other alleles (O02, A, B and AB) were present or not among Neandertals. The sequences of Vindija are quite incomplete in this region and the 261 and 297 diagnostic positions are not present in the draft, However, other markers in LD with the SNPs defining O01 seem to be absent in the reads generated. Thus, it is likely that additional alleles were also present among these humans. In the case of *TAS2R38* gene,

only reads with the taster haplotype have been found at Vindija. This point out the interest of genotyping different Neanderthal individuals, as a way to have an insight into their intra-specific genomic diversity. For this reason, at least a second Neanderthal genome draft would be desirable in the future.

The result of our work shows that both non-taster allele and O allele were present in the ancestral human population from which both Neanderthals and modern human diverged, thus predating the split of the two populations. Different nuclear markers explain different stories related to the differences in the gene trees. Some genetic markers occurred before the split of Neanderthals and modern humans from a common ancestor and thus are shared in both lineages (e.g. *FOXP2*) (Krause et al. 2007). Some genetic systems were polymorphic at that time and then, different alleles are present in both lineages (e.g. *ABO* and *TAS2R38*). Finally, some changes took place after the divergence of Neanderthals and then, some mutations are either present in one lineage or another (e.g. MC1R) (Lalueza-Fox et al. 2007).

Future researche will focus on those genes with more chances to have been subjected to positive selection, either in Neanderthals or in modern humans. In fact we are interested in identifying specific polymorphisms that may have evolved in the Neanderthal lineage or in the modern human since their divergence from common ancestral population. The latter genes could be crucial for establishing a genetic definition of our species. These forthcoming genetic findings will help us to understand more deeply the evolutionary relationship between these two homininae lineages but also to explore the uniqueness that characterizes both of them.

REFERENCE

Agelarakis,A. (1993). The Shanidar Cave proto-Neolithic Human population: Aspect of demography and paleopathology. *Journal of Human Evolution* 235-253

Ambrose,S.H., Norr,L. (1993). Experimental evidence for relationship of the carbone isotope ratios of whole diet and dietary protein to those of bone collagene and carbonate. In Lambert, J,B. Groupe G (Eds). Prehistoric Human Bone: Archaeology of the molecular level. Springer-Verlang, Berlin pp. 1-37

Anderson,S.A.T., Bankier,B.G., et al. (1981). Sequence and organization of the human mitochondrial genome. *Nature* **290**:457–465

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

Arensburg, B. (1991). The vertebral column, thoracic cage and hyoid bone, in: Bar Yosef, O., Vandermeersch, B. (Eds.), Le squelette Mousterien de Kebara 2. Ed. CNRS, Paris, pp. 125–146.

Arensburg,B., Schepartz,L.A., Tillier,M.A., Vandermeersch,B., Rak,Y., (1990). A reappraisal of the anatomical basis for speech in Middle Paleo- lithic hominids. *Am. J. Phys. Anthropol*. **83**: 137e146

Arensburg,B., Tillier,A.M., Vandermeersch,B., Duday,H., Schepartz, L.A., Rak, Y. (1989). A Middle Palaeolithic human hyoid bone. *Nature* **338**:758e760

Bailey,S.E., Weaver,T.D., Hublin,J.J. (2009) Who made the Aurignacian and other early Upper Paleolithic industries? *J Hum Evol.* **57(1)**:11-26

Balzeau,A., & Rougier,H. (2010). Is the suprainiac fossa a Neandertal autapomorphy? A complementary external and internal investigation. *Journal of human evolution* 58(1)

Bandelt,H.J. (2005). Mosaics of ancient mitochondrial DNA: positive indicators of nonauthenticity. *J Hum Genet* **13**:1106–12

Bastir,M., Rosas,A., Tabernero,A.G., Peña-Melián,A., Estalrrich, A., Rasilla,M. de la, et al. (2010). Comparative morphology and morphometric assessment of the Neandertal occipital remains from the El Sidrón site (Asturias, Spain: years 2000-2008). *Journal of human evolution*, 58(1)

Beauval,C., Maureille,B., Lacrampe-Cuyaubère,F., Serre,D., Peressinotto,D., Bordes,J.-G., et al. (2005). A late Neandertal femur from Les Rochers-de-Villeneuve, France. *Proceedings of the National Academy of Sciences of the United States of America*, **102(20)**: 7085-90

Been,E., Peleg,S., Marom,A., & Barash, A. (2010). Morphology and function of the lumbar spine of the Kebara 2 Neandertal. *American journal of physical anthropology***, 142(4)**: 549-57.

Bentley,D.R. et al. (2008). Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* **456**: 53–59

Bermudez de Castro, J.M., Arsuaga, J.L., Carbonell, E., Rosas, A., Martinez, I., Mosquera, M., (1997). A hominid from the lower Pleistocene of Atapuerca, Spain: possible ancestor to Neandertals and modern humans. *Science* **276**:1392e1395

Bianucci,R., Rahalison,L., Massa,E.R., Peluso,A., Ferroglio,E., & Signoli, M. (2008). Technical note: a rapid diagnostic test detects plague in ancient human remains: an example of the interaction between archeological and biological approaches (southeastern France, 16th-18th centuries). *American journal of physical anthropology,* **136(3)**: 361-7

Binladen,J., Wiuf,C., Gilbert,M.T.P., Bunce,M., Barnett,R., Larson,G., et al. (2006). Assessing the fidelity of ancient DNA sequences amplified from nuclear genes. *Genetics*, **172(2)**:733-41

Bocherens,H., Drucker,D., Billiou, D., Patou-Mathis, M.,Vandermeersch, B., (2005). Isotopic evidence for diet and subsistence pattern of theSaint-Cesaire I Neanderthal: review and use of a multi-source mixing model. *J. Hum. Evol.* **49**:71e87

Bogdanowicz, W., Allen, M., Branicki, W., Lembring, M., Gajewska, M., Kupiec, T. (2009). Genetic identification of putative remains of the famous astronomer Nicolaus Copernicus. *Proc Natl Acad Sci U S A.* **106(30)**:12279-82

Briggs,A.W., Good,J.M., Green,R.E., Krause,J., Maricic,T., Stenzel, U., et al. (2009). Targeted retrieval and analysis of five Neandertal mtDNA genomes. *Science* , **325(5938)**: 318-21

Briggs,A.W., Stenzel,U., Johnson,P.L.F., Green,R.E., Kelso,J., Prüfer,K., et al. (2007). Patterns of damage in genomic DNA sequences from a Neandertal. *Proceedings of the National Academy of Sciences of the United States of America*, **104(37)**: 14616-21

Brotherton,P., Endicott,P., Sanchez,J.J., Beaumont,M., Barnett,R., Austin,J., et al. (2007). Novel high-resolution characterization of ancient DNA reveals C > U-type base modification events as the sole cause of post mortem miscoding lesions. *Nucleic acids research,* **35(17)**: 5717-28

Brown,W.M. George,M. Jr,, Wilson,A.C. (1979). Rapid evolution of animal mitochondrial DNA. *Proc Natl Acad Sci U S A*. **76(4)**:1967-71

Brunet,M., Fronty,P., Sapanet,M., de Bonis,L,. Viriot,L. (2002) Enamel hypoplasia in a pliocene hominid from Chad. *Connect Tissue Res* **43(2-3)**:94-7

Burbano, H.A., Hodges,E., Green,R.E., Briggs, A.W., Krause,J., Meyer,M., et al. (2010). Targeted Investigation of the Neandertal Genome by Array-Based Sequence Capture. *Science*, **328(5979)**: 723-725

Cann,R.L., Stoneking,M., Wilson,A.C. (1987). Mitochondrial and human evolution. *Nature* **325**:31-6

Cano, R.J., Borucki, M.K. (1995) Revival and identification of bacterial spores in 25- to 40-million-year-old Dominican amber. *Science*. **268(5215)**:1265

Cano,R.J., Poinar,H.N., Pieniazek, Acra,A., Poinar,G,O. (1993). Amplification and sequencing of DNA from 120-135-million years old weevil. *Nature* **363**: 536-8

Caramelli,D., Lalueza-Fox,C., Condemi,S., Longo,L., Milani,L., Manfredini,A., et al. (2006). A highly divergent mtDNA sequence in a Neandertal individual from Italy. *Current biology* **16(16)**: R630-2

Caramelli,D., Lalueza-Fox,C., Vernesi,C., Lari,M., Casoli,A., Mallegni,F., et al. (2003). Evidence for a genetic discontinuity between Neandertals and 24,000-year-old anatomically modern Europeans. *Proceedings of the National Academy of Sciences of the United States of America*, **100(11)**: 6593-7

Caramelli,D., Vernesi,C., Sanna,S., Sampietro,L., Lari,M., Castrì,L., et al. (2007). Genetic variation in prehistoric Sardinia. *Human genetics*, **122(3-4)**: 327-36

Caspari,R. (2005). The suprainiac fossa: the question of homology. *Anthropologie (Brno)* **43**: 229–239

Champlot,S., Berthelot,C., Pruvost,M., BennettE. A., Grange,T., & Geigl,E.-M. (2010). An efficient multistrategy DNA decontamination procedure of PCR reagents for hypersensitive PCR applications. *PloS one*, 5(9)

Coble,M.D., et al. (2009). Mystery solved: the identification of the two missing Romanov children using DNA analysis. *PLoS One.* **4(3)**:e4838

Cooper, A., Poinar, H. (2000). Ancient DNA: do it right or not at all. *Science* **289**:1139

Cooper,A. (1992). Removal of colourings, inhibitors of PCR, and the carrier effect of PCR contamination from ancient DNA samples. Anc. *DNA Newslett*. **1**:31–32

De Salle,R., Gatesy,J., Wheeler,W., Grimaldi,D. (1992). DNA sequence from a fossil termite in Oligo-Miocene amber and their phylogenetic implications. *Science* **257**: 1933-6

Drancourt,M., Signoli,M., Dang,L.V., Bizot,B., & Roux,V. (2007). Orientalis in Remains of Patients. Emerging *Infectious Diseases*, **13(2)**:1720-1721

Endicott,P., Ho,S.Y., Metspalu,M., Stringer,C. (2009) Evaluating the mitochondrial timescale of human evolution. *Trends Ecol Evol*. **24(9)**:515-21 Review.

Facchini,F., and Belcastro,G. (2009): La lunga storia dei Neanderthal. Biologia e comportamento. Jaka Book

Fortea,J., de la Rasilla,M., Martínez,E., Sánchez-Moral,S., Cañaveras,J.C., Cueva,S., Rosas,A., Soler,V., Julià,R., de Torres,T., et al. (2003). La Cueva de El Sidrón (Borines, Piloña, Asturias): Primeros resultados. *Estud. Geol*. **59**: 159–179.

Galik,K., Senut,B., Pickford,M., Gommery,D., Treil,J., Kuperavage,A.J., Eckhardt, R.B. (2004). External and internal morphology of the BAR 1002'00 Orrorin tugenensis femur. *Science*. **305(5689)**:1450-3

Gibbons,A. (1994). Possible dino DNA find is greeted with skepticism. *Science* **266**:1159

Gilbert,M.T. et al. (2007). Whole-genome shotgun sequencing of mitochondria from ancient hair shafts. *Science.* **28**:1927-30

Gilbert,M.T.P., Bandelt,H.-J., Hofreiter,M., & Barnes,I. (2005). Assessing ancient DNA studies. *Trends in ecology & evolution),* **20(10)**: 541-4

Gilbert, M.T.P., Hansen,A.J., Willerslev,E., Rudbeck L., Barnes,I., Lynnerup,N. et al. (2003). Characterization of genetic miscoding lesions caused by postmortem damage. *American journal of human genetics,* **72(1)**: 48-61

Gilbert, M.T.P., Cuccui,W. White,N. Lynnerup,R.W. Titball,A. Cooper,A., and PrenticeM.B. (2004a). Absence of Yersinia pestis-specific DNA in human teeth from five European excavations of putative plague victims. *Microbiology* **150**:341–354.

Gilbert,M.T.P., Janaway,R.C., Tobin,D.J., Cooper,A., & Wilson,A. S. (2006). Histological correlates of post mortem mitochondrial DNA damage in degraded hair. *Forensic science international*, **156(2-3)**: 201-7

Gilbert,M.T.P., Willerslev,E., Hansen,A.J., Barnes,I., Rudbeck,L., Lynnerup,N. et al. (2003). Distribution patterns of postmortem damage in human mitochondrial DNA. *American journal of human genetics*, **72(1)**: 32-47

Giles,R.E., Blanc,H., Cann,H.M., Wallace,D.C. (1980). Maternal inheritance of human mitochondrial DNA. *Proc Natl Acad Sci U S A.* **77(11)**:6715-9.

Gill,P., Hagelberg,E. (2004). Ongoing controversy over Romanov remains. *Science.* **306(5695)**:407-10

Golemberg,E.M., Giannasi,D.E., Clegg,M.T., Smiley,C.J., Durbin,M., Henderson, D., Zurawski A,D. (1990). Chloroplast DNA sequence from miocene Magnolia species. *Nature* **344**: 656-8

Gómez-Olivencia,A., Eaves-Johnson,K.L., Franciscus,R.G., Carretero,J.M., & Arsuaga,J.L. (2009). Kebara 2: new insights regarding the most complete Neandertal thorax. *Journal of human evolution*, **57(1)**: 75-90

Gorovic,N., Afza,lS., Tjønneland,A., Overvad,K., Vogel,U., Albrechtsen,C., Poulsen,H.E. (2011). Genetic vari.ation in the hTAS2R38 taste receptor and brassica vegetable intake. *Scand J Clin Lab Invest*. [Epub ahead of print]

Green,R.E. et al. (2009) The neandertal genome and ancient DNA authenticity. *EMBO J*. **28**: 2494–2502

Green,R.E., Briggs,A.W., Krause,J., Prüfer,K., Burbano,H.A, Siebauer,M., et al. (2009). The Neandertal genome and ancient DNA authenticity. *The EMBO journa*l, **28(17)**: 2494-502

Green,R.E., Krause,J., Briggs,A.W., Maricic,T., Stenzel,U., Kircher,M., et al. (2010). A draft sequence of the Neandertal genome. *Science*, **328(5979)**: 710-22

Green,R.E., Krause,J., Ptak,S. E., Briggs,A.W., Ronan,M.T., Simons,J F., et al. (2006). Analysis of one million base pairs of Neanderthal DNA. *Nature*, **444(7117)**: 330-6.

Green,R.E., Malaspinas,A.-S., Krause,J., Briggs,A.W., Johnson,P.L.F., Uhler,C., et al. (2008). A complete Neandertal mitochondrial genome sequence determined by high-throughput sequencing. *Cell*, **134(3)**: 416-26.

Greenberg,B.D., Newbold,J.E., Sugino,A. (1983) Intraspecific nucleotide sequence variability surrounding the origin of replication in human mitochondrial DNA. *Gene*. **21(1-2)**:33-49

Gunz,P., Harvati,K. (2007). The Neanderthal 'chignon': variation, integration and homology. *J. Hum. Evol*. **52**:262–74

Hagelberg,E., Bell,L.S., Allen,T., Boyde,A., Jones ,S.J, Clegg,J,B. (1991). Analysis of ancient bone DNA: techniques and applications. *Philos. Trans. R So.c Lond. B Biol Sci*. **333(1268)**:399-407.

Hagelberg,E., Sykes,B., Hedges,R. (1989). Ancient bone DNA amplified. *Nature*. **342(6249)**:485.

Hagelberg,E., Thomas,M.G., Cook,C.E.Jr., Sher,A.V., Baryshnikov,G.F., Lister,A.M. (1994). DNA from ancient mammoth bones. *Nature* **370**:333–34

Handt,O., Höss,M., Krings,M., Pääbo,S. (1994). Ancient DNA: methodological challenges. *Experientia* **50**:524–29

Hänni,C., Brousseau,T., Laudet,V., & Stehelin,D. (1995). Isopropanol precipitation removes PCR inhibitors from ancient bone extracts. *Nucleic acids research*, **23(5)**: 881-2

Hänni,C., Laudet,V., Stehelin,D., & Taberlet,P. (1994). Tracking the origins of the cave bear (Ursus spelaeus) by mitochondrial DNA sequencing. *Proceedings of the National Academy of Sciences of the United States of America*, **91(25)**:12336-40.

Hansen,A.J., WillerslevE., WiufC., Mourier,T., and Arctander.P. (2001). Statistical evidence for miscoding lesions in ancient DNA templates. *Mol. Biol. Evol*. **18**:262–265

Hansen,A.J., Willerslev,E., Wiuf,C., Mourier,T., & Arctander,P. (1997). Letter to the Editor Statistical Evidence for Miscoding Lesions in Ancient DNA Templates. *DNA Sequence*, 262-265.

Henry,A.G., Brooks,A.S., & Piperno,D.R. (2010). Microfossils in calculus demonstrate consumption of plants and cooked foods in Neanderthal diets. PNAS. 486:91

Higuchi, R., Bowman, B., Freiberger, M., Ryder, O.A., Wilson, A.C. (1984). DNA se- quences from the quagga, an extinctmember of the horse family. *Nature* **312**:282– 84

Hobolth,A., Dutheil,J.Y., Hawks,J., Schierup,M.H., Mailund,T. (2011).Incomplete lineage sorting patterns among human, chimpanzee, and orangutan suggest recent orangutan speciation and widespread selection. *Genome Res*. **21(3)**:349-56.

Hodges,E., Xuan,Z,. Balija,V., Kramer,M., Molla,M.N. et al. (2007) Genome-wide in situ exon capture for selective resequencing. *Nat Genet* **39**: 1522–1527.

Hodges,E., Rooks,M., Xuan,Z., Bhattacharjee,A., Benjamin,D., Brizuela,L., et al. (2010). *NIH Public Access. Leonardo*, **4(6)**: 960-974.

Hofreiter,M., Jaenicke,V., Serre,D., Haeseler,A.V., & Pääbo,S. (2001). DNA sequences from multiple amplifications reveal artifacts induced by cytosine deamination in ancient DNA. *Nucleic. Acids. Res.*, **29(23)**: 4793-9.

Hofreiter,M., Serre,D., Poinar,H.N., Kuch,M., & Pääbo,S. (2001). Ancient DNA. *Nature reviews Genetics* **2(5)**: 353-9.

Holton,N.E., & Franciscus,R.G. (2008). The paradox of a wide nasal aperture in cold-adapted Neandertals: a causal assessment. *Journal of Human Evolution* **55(6)** 942-51.

Hoss,M., Paabo,S., Vereshchagin,N.K. (1994). Mammoth DNA sequences. *Nature* **370**:333

Hotton, F. et al. (1976). Petrus bone of the Spy Neanderthal. *Radio-anatomic study. J. Belge Radiol*, **(59)**: 39-50

Hublin, J-J. (1998) Neandertals and Modern Humans in Western Asia, eds Akazawa T, Aoki K, Bar-Yosef O (Plenum, New York), pp 295–310

Hublin, J.J. (2009). Out of Africa: modern human origins special feature: the origin of Neandertals. *Proc Natl Acad Sci U S A*. **106(38)**:16022-7.

Hublin,J.J. (1980). La Chaise Suard, Engis 2 et La Quina H18: de´veloppement de la morphologie occipitale externe chez l'enfant pre´ne´andertalien et ne´an- dertalien. *C.R. Acad. Sci. Paris* **291**, 669–672.

Hublin,J.J., Spoor,F., Braun,M., Zonneveld,F., Condemi,S. (1996). A late Neanderthal from Arcy-sur-Cure associated with Upper Palaeolithic artefacts. *Nature* **381**: 224–226.

Kircher,M., Stenzel,U., & Kelso,J. (2009). Improved base calling for the Illumina Genome Analyzer using machine learning strategies. *Genome biology*, 10(8), R83.

Kramer,A., Crummett,T.L., Wolpoff,M.H., (2001). Out of Africa and into the Levant: replacement or admixture in Western Asia? *Quatern. Int.* **75**: 51–63

Krause,J., Briggs,A.W., Kircher,M., Maricic,T., Zwyns,N., Derevianko A. et al. (2010). A complete mtDNA genome of an early modern human from Kostenki, Russia. *Current biology : CB*, **20(3)**: 231-6.

Krause,J., Dear, PH., Pollack,J.L., Slatkin,M., Spriggs,H., Barnes,I., Lister,A.M, Ebersberger,I., Pääbo,S., Hofreiter,M. (2006). Multiplex amplification of the mammoth mitochondrial genome and the evolution of Elephantidae. *Nature* **439(7077)**:724-7.

Krause,J., Lalueza-Fox,C., Orlando,L., Enard,W., Green,R.E., Burbano,H.A, et al. (2007). The derived FOXP2 variant of modern humans was shared with Neandertals. *Current biology* **17(21)**: 1908-12.

Krings,M., Capelli,C., Tschentscher, F, Geisert,H.,Meyer,S., et al. (2000). Aviewof Neandertal genetic diversity. *Nat. Genet.* **26**:144–46

Krings, M., Stone, A., Schmitz, R.W., Krainitzk,i H., Stoneking, M., Pääbo, S. (1997). NeandertalDNAsequences and the origin of modern humans. *Cell* **90**:19–30

Krings,M., Geisert,H., Schmitz,R.W., Krainitzki,H., & Pääbo,S. (1999). DNA sequence of the mitochondrial hypervariable region II from the neandertal type specimen. *Proceedings of the National Academy of Sciences of the United States of America* **96(10)**: 5581-5.

Laitman,J.T. (1996). What the nose knows: New understandings of Neanderthal upper respiratory tract specializations. *Proceedings of the National Academy of Sciences* **93(20)**:10543-10545.

Lalueza,C., Pérez-Pérez,A., and Turbon,D. (1996). Dietary inferences through buccal microwear analysis of middle and upper Palaeolithic human fossils. *American Journal of Physical Anthropology* **100**:367-387.

Lalueza-Fox,C. et al. (2011). Genetic evidence for patrilocal mating behavior among Neandertal groups. *Proc Natl Acad Sci U S A*. **108(1)**:250-3.

Lalueza-Fox, C., Sampietro, M.L., Caramelli, D., et al. (2005). Neandertal evolutionary genetics: mitochondrial DNA data from the Iberian Peninsula. *Mol Biol Evol* **22**:1077–81

Lalueza-Fox,C. et al. (2007). A melanocortin 1 receptor allele suggests varying pigmentation among Neanderthals. Science **318(5855)**:1453-5

Lalueza-Fox,C., Castresana,J., Sampietro,L., Marquès-Bonet,T., Alcover,J.A., & Bertranpetit,J. (2005). Molecular dating of caprines using ancient DNA sequences of Myotragus balearicus, an extinct endemic Balearic mammal. *BMC evolutionary biology*, 5, 70.

Lalueza-Fox,C., Gigli,E., Bini,C., Calafell, F., Luiselli,D., Pelotti,S., Pettener,D. (2010). Genetic analysis of the presumptive blood from Louis XVI, king of France. *Forensic Sci Int Genet*.

Lalueza-Fox,C., Römpler,H., Caramelli,D., Stäubert,C., Catalano,G., Hughes,D., et al. (2007). A melanocortin 1 receptor allele suggests varying pigmentation among Neanderthals *Science* **318(5855)**: 1453-5

Lalueza-Fox,C., Rosas,, Estalrrich,A., Gigli,E., Campos,P.F., Garcia-Tabernero,A. et al. (2010). Genetic evidence for patrilocal mating behavior among Neandertal groups. *Proceedings of the National Academy of Sciences* **108(1)**:250-253.

Lebatard,A.E., et al. (2008). Cosmogenic nuclide dating of Sahelanthropus tchadensis andAustralopithecus bahrelghazali: Mio-Pliocene hominids from Chad. *Proc Natl Acad Sci U S* . **105(9)**:3226-31.

Lieberman,P. (1984) The Biology and Evolution of Language. Cambridge, MA: Harvard University Press

Lieberman,P., Laitman,J.T., Reidenberg J.S., Gannon,P.J., (1992). The anatomy, physiology, acoustics and perception of speech: essential elements in analysis of the evolution of human speech. *J. Hum. Evol*. **23**: 447e467

Lindahl,T. (1993).Recovery of antediluvian DNA. *Nature* **365**:700

Malmström,H., Storå,J., Dalén,L., Holmlund,G., & Götherström,A. (2005). Extensive human DNA contamination in extracts from ancient dog bones and teeth. *Molecular biology and evolution* **22(10)**: 2040-7.

Malmström,H., Svensson,E. M., Gilbert,M.T.P., Willerslev,E., Götherström,A., & Holmlund,G. (2007). More on contamination: the use of asymmetric molecular behavior to identify authentic ancient human DNA. *Molecular biology and evolution* **24(4)**: 998-1004.

Margulies, M. et al. (2005). Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437(7057)**:376-80.

Maricic,T., Whitten,M., & Pääbo,S. (2010). Multiplexed DNA Sequence Capture of Mitochondrial Genomes Using PCR *Products. PLoS ONE*, **5(11)**: e14004.

Martínez,I., Arsuaga,J.L., Quam,R., Carretero,J.M., Gracia,A, & Rodríguez,L. (2008). Human hyoid bones from the middle Pleistocene site of the Sima de los Huesos (Sierra de Atapuerca, Spain). *Journal of human evolution* **54(1)**: 118-24.

Martınez,I., Rosa,M., Arsuaga, J.L., Jarabo,P., Quam,R., Lorenzo,C., Gracia,A., Carretero,J.M., Bermudez de Castro,J., Carbonell,E. (2004). Auditory capacities in middle Pleistocene humans from the Sierra de Atapuerca in Spain. *Proc. Natl. Acad. Sci. U.S.A*. 101, 9976e9981.

Matheson,C.D., Vernon,K.K., Lahti,A., Fratpietro,R., Spigelman,M., Gibson,S., et al. (2009). Molecular exploration of the first-century Tomb of the Shroud in Akeldama, Jerusalem*. PloS one* **4(12)**: e8319.

Mays,S., Fysh,E., & Taylor,G.M. (2002). Investigation of the link between visceral surface rib lesions and tuberculosis in a Medieval skeletal series from England using ancient DNA. *American journal of physical anthropology* **119(1)**: 27-36.

McDermott,F., Grün,R., Stringer,C.B., and Howkeworth,C.J. (1993). Mass-spectrometric U-series dates for Israeli neanderthal/early modern human sites. *Nature* **363**: 193-255

Mellars,P., (1996). The Neanderthal Legacy. Princeton University Press, Princeton.

Mitchell,D., Willerslev,E., Hansen, A. (2005). Damage and repair of ancient DNA. *Mutat Res* **571(1-2)**:265-76.

Mourant AE, Kopec AC, Domaniewska-Sobczak K. (1978). Blood groups and diseases. Oxford: Oxford University Press.

Mullis,K.B. and Floona,F.A. (1987). Specific synthesis of DNA i n vitro via a polymerase - catalysedch a in reactIon. *Methods Enzymol*. **155**: 35-350

Noble,W., and Davidson,I. (1991). The evolutionary emergence of modern human behaviour: Language and i t s archaeology. *Man* **26**:223-253

Noro,M., Masuda,R., Dubrovo,I.A., Yoshida,M.C., Kato,M. (1998). Molecular phylogenetic inference of the woolly mammoth Mammuthus primigenius, based on complete sequences of mitochondrial cy- tochrome b and 12S ribosomal RNA genes. *J. Mol. Evol*. **46:**314–26

Ondov,B.D., Varadarajan,A., Passalacqua,K.D., & Bergman,N.H. (2008). Efficient mapping of Applied Biosystems SOLiD sequence data to a reference genome for functional genomic applications. *Bioinformatics* (Oxford, England), **24(23)**: 2776-7

Orlando,L., Darlu,P., Toussaint,M., Bonjean,D., Otte,M., and Hänni,C. (2006). Revisiting Neandertal diversity with a 100,000 year old mtDNA sequence. *Curr. Biol*. **16**: R400

Ovchinnikov, I.V., Gotherstrom, A., Romanova, G.P., Kharitonov, V.M., Liden, K., GoodwinW. (2000). Molecular analysis of NeanderthalDNAfrom the northern Caucasus. *Nature* **404**:490–93

Pääbo,S., Wilson,A.C. (1991). Miocene DNA sequences - a dream come true? *Curr Biol*. **1(1)**:45-6.

Pääbo,S. (1985). Molecular cloning of ancient Egyptian mummy DNA. *Nature* **314**: 644-45

Pääbo,S. (1989). Ancient DNA: extraction, characyerization, molecular cloning, and enzymatic amplification. *Proc. Natl. Acad. Sci. USA* **86**: 1939-43

Pääbo,S., Higuchi,R.G., & Wilson A.C. (1989). Ancient DNA and the polymerase chain reaction. The emerging field of molecular archaeology. *The Journal of biological chemistry*, **264(17)**: 9709-12

Pääbo,S., Poinar,H., Serre,D., Jaenicke-Despres,V., Hebler,J., Rohland,N. et al. (2004). Genetic analyses from ancient DNA. *Annual review of genetics*

Patterson,N., Richter,D.J., Gnerre,S., Lander,E.S., Reich,D. (2006). Genetic evidence for complex speciation of humans and chimpanzees. *Nature*. **441(7097):**1103-8.

Piperno,M and Giacobini,G. (1992). A taphonomic study of the paleosurface of Guattari Cave, 'Quaternaria Nova, Vol. I, 1990-1991, Proccedings of the International Symposium The fossil man of Monte Circeo: fifty years of studies on the neandertals in Latium, A. Bietti et G. Manzi Eds., (1992), pp. 143-161

Poinar,H.N. (2002). The genetic secrets some fossils hold. *Accounts of chemical research*, **35(8)**: 676-84

Ponce de Leon,M, et al. (2008). Neanderthal brain size at birth provides insights into the evolution ofhumanlife history. *Proc Natl Acad Sci USA* **105**:13764–13768

Prüfer,K., Stenzel,U., Hofreiter,M., Pääbo,S., Kelso,J., & Green,R.E. (2010). Computational challenges in the analysis of ancient DNA. *Genome biology*, **11(5)**: R47

Pruvost,M., Grange,T., & Geigl,E.-M. (2005). Minimizing DNA contamination by using UNG-coupled quantitative real-time PCR on degraded DNA samples: application to ancient DNA studies. *BioTechniques* **38(4)**:569-75.

Pruvost,M., Schwarz,R., Correia,V.B., Champlot,S., Braguier,S., Morel,N., et al. (2007). Freshly excavated fossil bones are best for amplification of ancient DNA. *Proceedings of the National Academy of Sciences of the United States of America* **104(3)**: 739-44

Ramírez,O., Gigli,E., Bover,P., Alcover,J.A., Bertranpetit,J., Castresana,J., et al. (2009). Paleogenomics in a temperate environment: shotgun sequencing from an extinct Mediterranean caprine. *PloS one* **4(5)**: e5670

Rasmussen,M., Li,Y., Lindgreen,S., Pedersen,J.S., Albrechtsen, A., Moltke,I., et al. (2010). Ancient human genome sequence of an extinct Palaeo-Eskimo. *Nature* **463(7282)**: 757-62

Reich,D., Green,R.E., Kircher,M., Krause,J., Patterson,N., Durand, E.Y., et al. (2010). Genetic history of an archaic hominin group from Denisova Cave in Siberia. *Nature*, **468(7327)**: 1053-1060

Richards,M.P., & Trinkaus,E. (2009). Isotopic evidence for the diets of European Neanderthals and early modern humans, 106(38).

Richards,M.P., Taylor,G., Steele,T., McPherron,S.P., Soressi,M., Jaubert,J., et al. (2008). Isotopic dietary analysis of a Neanderthal and associated fauna from the site of Jonzac (Charente-Maritime), France. *Journal of human evolution*, **55(1)**: 179-85

Rollo,F., Marota,I. (1999). How microbial ancient DNA, found in association with human remains, can be interpreted. *Philos Trans R Soc Lond B Biol Sci*. **345**:111-9

Rollo, F., Sassaroli, S., Ubaldi, M. (1995). Molecular phylogeny of the fungi of the Iceman's grass clothing. *Curr Genet*. **28(3)**:289-97.

Rosas,A., Martínez-Maza,C., Bastir,M., García-Tabernero,A., Lalueza-Fox,C., Huguet,R., et al. (2006). Paleobiology and comparative morphology of a late Neandertal sample from El Sidron, Asturias, Spain. *Proceedings of the National Academy of Sciences of the United States of America* **103(51)**: 19266-71

Ruff,C.B., (1994). Morphological adaptation to climate in modern and fossil hominids. *Phys. Anthrop* **37**: 65–107

Russel, P.J., (1998). Genetica.EdiSes

Saiki, R.K., Gelfand, D.H., Stoffe,l S., Scharf, S.J., Higuchi, R., et al. (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**:487–91

Saiki, R.K, Scharf, S., Faloona, F., Mullis, K.B., Horn, G.T., et al. (1985). Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* **230**:1350– 54

Sampietro,L., Gilbert,M.T.P., Lao,O., Caramelli,D., Lari,M., Bertranpetit,J., et al. (2006). Tracking down Human Contamination in Ancient Human Teeth. *Molecular Biology*. **23**:1801-07

Sanger,F., Nicklen,S., & Coulson,A.R. (1977). DNA sequencing with chain-terminating inhibitors. *Biotechnology* **24(12)**: 104-8

Schmitz,R.W., Serre,D., Bonani,G., Feine,S., Hillgruber,F., Krainitzki,H., Paabo,S., and Smith,F.H. (2002). The Neandertal type site revisited: Interdisciplinary investigations

of skeletal remains from the Neander Valley, Germany. *Proc. Natl. Acad. Sci*. USA **99**: 13342–13347

Scholz,M., Giddings,I., Pusch, C.M. (1998). A polymerase chain reaction inhibitor of ancient hard and soft tissue DNA extracts is determined as human collagen type I. *Anal Biochem.* **259(2)**:283-6

Schwartz,J.H., & Tattersall,I. (1996). Significance of some previously unrecognized apomorphies in the nasal region of Homo neanderthalensis. *Proceedings of the National Academy of Sciences of the United States of America,* **93(20)**: 10852-4

Serre,D., Hofreiter,M., & Pääbo,S. (2004). Mutations induced by ancient DNA extracts? *Molecular biology and evolution*, **21(8)**: 1463-7

Shapiro,B., et al. (2004) Rise and fall of the Beringian steppe bison. *Science*. **306(5701)**:1561-5

Shapiro,L. (1993). Functional morphology of the vertebral column in primates. In: Gebo D, editor. Postcranial adaptation in non-human primates. DeKalb: Northern Illinois University Press. p 121–149.

Shapiro,L. (2007). Morphological and functional differentiation in the lumbar spine of lorisids and galagids. *Am J Primatol* **39**:86–102

Shea, J.S., Brooks, A.S. (2000). Mousterian. In: Delson E, Tattersall I, Van Couvering JA, Brooks AS, editors. Encyclopedia of human evolution and prehistory. 2nd ed. New York: Garland; p. 453–60

Shendure, J., Porreca, G.J., Reppas, N.B., Lin, X., McCutcheon, J.P., Rosenbaum, A.M., Wang, M.D., Zhang, K,. Mitra, R.D., Church, G.M. (2005) Accurate multiplex polony sequencing of an evolved bacterial genome. *Science*. **309(5741)**:1728-32

Smith,T.M., Harvati,K., Olejniczak,A. J., Reid,D.J., Hublin,J.-J., & Panagopoulou,E. (2009). Brief communication: dental development and enamel thickness in the Lakonis Neanderthal molar. *American journal of physical anthropology*, **138(1)**:112-8

Smith,T.M., Toussaint,M., Reid,D.J., Olejniczak,A.J., & Hublin,J.-J. (2007). Rapid dental development in a Middle Paleolithic Belgian Neanderthal. *Proceedings of the National Academy*

of Sciences of the United States of America, **104(51)**: 20220-5

Spoor,C.F., (1993). The comparative morphology and phylogeny of the human bony labyrinth. Ph.D. dissertation, Utrecht University.

Spoor,C.F., Zonneveld,F.W., (1994). The bony labyrinth in Homo erectus, a preliminary report. *Cour. Forschungsinst. Senckenberg* **171**: 251–256

Spoor,F. (2003). The bony labyrinth of Neanderthals. *Journal of Human Evolution*, **44(2)**:141-165

Stewart,T.D. (1977). The Neanderthal skeletal remains from Shanidar Cave, Iraq: A summary of findings to date. *Proceeding of the American Philosophical Society*. **121**: 121-165

Stiller,M. et al. (2006). Patterns of nucleotide misincorporations during enzymatic amplification and direct large-scale sequencing of ancient DNA. *Proc Natl Acad Sci U S A.* **12**;103(37)

Stiner,M. (1992). The Guattari faunas then and now, 'Quaternaria Nova, Vol. I, 1990-1991, Proccedings of the International Symposium The fossil man of Monte Circeo : fifty years of studies on the neandertals in Latium, A. Bietti et G. Manzi Eds., (1992), pp. 163-192.

Stone, A.C., Starrs, J.E., Stonekin,G.M., (2001). Mitochondrial DNA analysis of the presumptive remains of Jesse James. J *Forensic Sci*. **46(1)**:173-6

Trinkaus, E. et al. (2003) An early modern human from the Peçstera cu Oase, Romania. *Proc Nat Acad Sci USA* **100**:11231–11236

Trinkaus, E. (1985). Pathology and the posture of the La Chapelle-aux-Saints Neandertal. *Am J Phys Anthropol*. **67(1)**:19-41.

Trinkaus, E., (2004). Eyasi 1 and the suprainiac fossa. *Am. J. Phys. Anthropol.* **124**: 28–32

Valladas, H., Joron, J.L., Valladas, G., Arensburg, B., Bar-Yosef, O., Belfer-Cohen, A., Goldberg, P., Laville, H., Meignen, L., Rak ,Y.et al. (1987). Thermoluminescence dates for the Neanderthal burial site at Kebara in Israel. *Nature* **330**:159-160

Vandermeersch,B. (1989). ESR dating of the Neanderthal site, Kebara Cave, Israel. Journal of Archaeological *Science* **16**:653-659

Vandermeersch,B., Trinkaus,E. (1995).The postacranial remains of the regoudon 1 Neanderthal-The shoulder and arm remains. *Journal of Human Evolution* **28**:439-476

Vernesi,C., Di BenedettoG., Caramelli,D., Secchieri,E., Simoni,L., Katti,E., Malaspina,P., Novelletto,A., Marin,V.T., Barbujani,G. (2001). Genetic characterization of the body attributed to the evangelist Luke. *Proc Natl Acad Sci U S A*. **98(23)**:13460-3

Vestheim,H., Jarman,S.N. (2008). Blocking primers to enhance PCR amplification of rare sequences in mixed samples a case study on prey DNA in Antarctic krillstomachs. *Front. Zool* **5**: 12

Vigilant,L., Pennington,R., Harpending,H., Kocher,T.D, Wilson,A.C. (1989). Mitochondrial DNA sequences in single hairs from a southern African population. *Proc Natl Acad Sci U S A.* **86(23)**:9350-4

Vives,S., Gilbert,M.T., Arenas,C., Gigli,E., Lao,O., & Lalueza-Fox,C. (2008). Statistical analysis of post mortem DNA damage-derived miscoding lesions in Neandertal mitochondrial DNA. *BMC research notes* **1: 40**

Vreeland, R.H., Rosenzweig, W.D., Powers, D.W. (2000) Isolation of a 250 million-year-old halotolerant bacterium from a primary salt crystal. *Nature.* **407(6806)**:897-900

Weaver,T.D., & Hublin,J.-J. (2009). Neandertal birth canal shape and the evolution of human childbirth. *Proceedings of the National Academy of Sciences of the United States of America,* **106(20)**: 8151-6

Weaver,T.D., Roseman,C.C., & Stringer,C.B. (2007). Were neandertal and modern human cranial differences produced by natural selection or genetic drift? *Journal of human evolution,* **53(2)**:135-45

Weber,J., & Pusch,C. M. (2008). The lumbar spine in Neanderthals shows natural kyphosis. European spine journal : official publication of the European Spine Society, *the European Spinal Deformity Society, and the European Section of the Cervical Spine Research Society*, **2**: S327-30

Weinstein,K.J. (2007). Thoracic Skeletal Morphology and High-Altitude Hypoxia in Andean Prehistory. *Chest* 36-49

Weinstein,K.J. (2008). Thoracic morphology in Near Eastern Neandertals and early modern humans compared with recent modern humans from high and low altitudes. *Journal of human evolution* **54(3)**: 287-95

Willerslev,E., & Cooper,A. (2005). Ancient DNA. *Proceedings. Biological sciences / The Royal Society*, **272(1558)**: 3-16

Willerslev,E., Hansen,A.J., Binladen,J., Brand,T.B., Gilbert,M.T., Shapiro,B., Bunce,M., Wiuf, C.,Gilichinsky,D.A., Cooper, A. (2003). Diverse plant and animal genetic records from Holocene and Pleistocene sediments. *Science* **300(5620):**791-5

Willerslev,E., Hansen,A.J., Rønn,R., Brand,T.B., Barnes,I., Wiuf,C., Gilichinsky,D., Mitchell,D., Cooper,A. (2004). Long-term persistence of bacterial DNA. *Curr Biol*. **14(1)**:R9-10

Wooding,S., Un-kyung,K., Bamshad,M.J., Larsen,L., Jorde,l.B., and Drayna,D. (2004). Natural Selection and Molecular Evolution in *PTC,* a Bitter-Taste Receptor Gene *Am. J. Hum. Genet.* **74**:637–646

Woodward,S.R., Weyand,N.J., Bunnel,M. (1994b). DNA sequence from Cretaceous period bone fragments. *Science* **266**:1229-32

Wynn,T. (2002). Archaeology and cognitive evolution. Behav. *Brain Sci*. **25**: 389–437

Zilhão,J., Angelucci,D.E., Badal-García,E., d'Errico,F., Daniel,F., Dayet,L., et al. (2010). Symbolic use of marine shells and mineral pigments by Iberian Neandertals. *Proceedings of the National Academy of Sciences of the United States of America*, **107(3)**:1023-8

Zink,A.R., Grabner,W., & Nerlich,A.G. (2005). Molecular identification of human tuberculosis in recent and historic bone tissue samples: The role of molecular techniques for the study of historic tuberculosis. *American journal of physical anthropology,* **126(1)**: 32-47

Zink,A.R., Sola,C., Reischl,U., Grabner,W., Rastogi,N., Wolf,H., et al. (2003). Characterization of Mycobacterium tuberculosis Complex DNAs from Egyptian Mummies by Spoligotyping. *Society* **41(1)**: 359-367

APPENDIX A. Golden Criteria

1.) **Physically isolated work areas**: the best case scenario is to have the aDNA laboratory and the main laboratory in different places

2.) **Negative control extraction and amplifications**: to detect sporadic contamination in the reagents

3.) **Appropriate molecular behaviour**: if shorter fragments are not easily amplified than longers ones, it is an indication that the source of DNA could be contaminats

4.) **Reproducibility**: multiple PCRs and extractions from the same sample are useful to detect contamination of a particular extraction or amplification

5.) **Cloning**: in order to detect post mortem damage and contamination in the PCR products

6.) **Independent replication**: separate samples of specimens are extracted and sequenced in indipendent laboratories

7.) **Biochemical preservation**: it may be used to test the general state preservation of DNA

8.) **Quantification of DNA templates**: using quantitative PCR or RT-PCR to estimate the number of DNA molecules

9.) **Associated remains**: the amplification of DNA target that survived in associated faunal remains can be an indicator of general DNA preservation

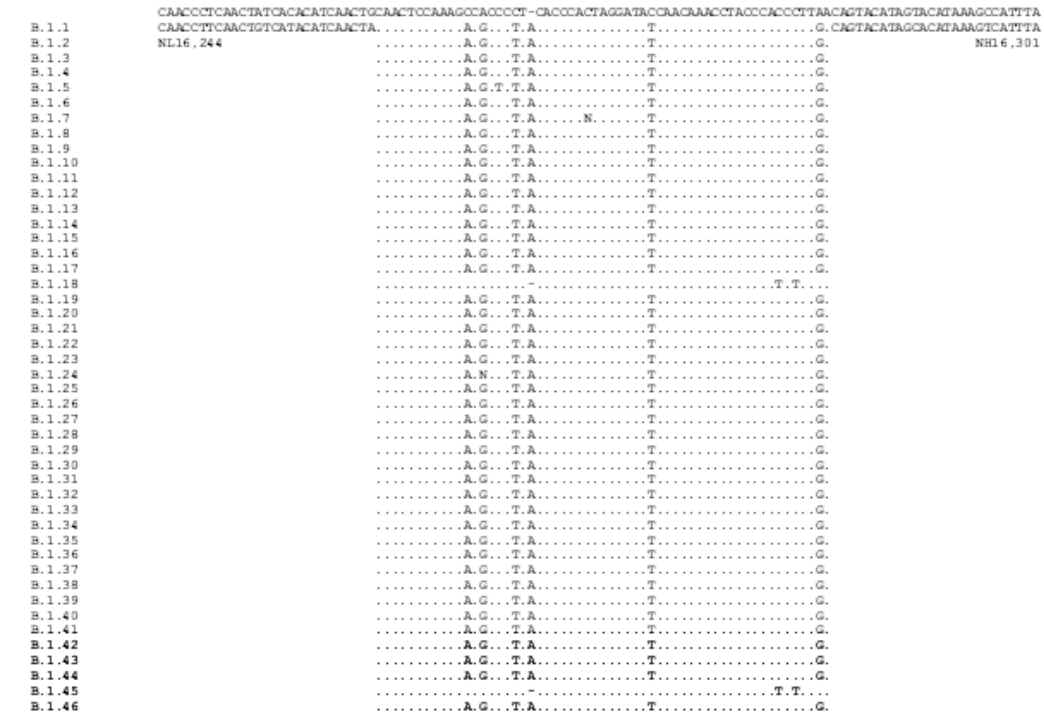
APPENDIX B. Supplementary Information

B1. Supplemetary information to chapter 3.2.

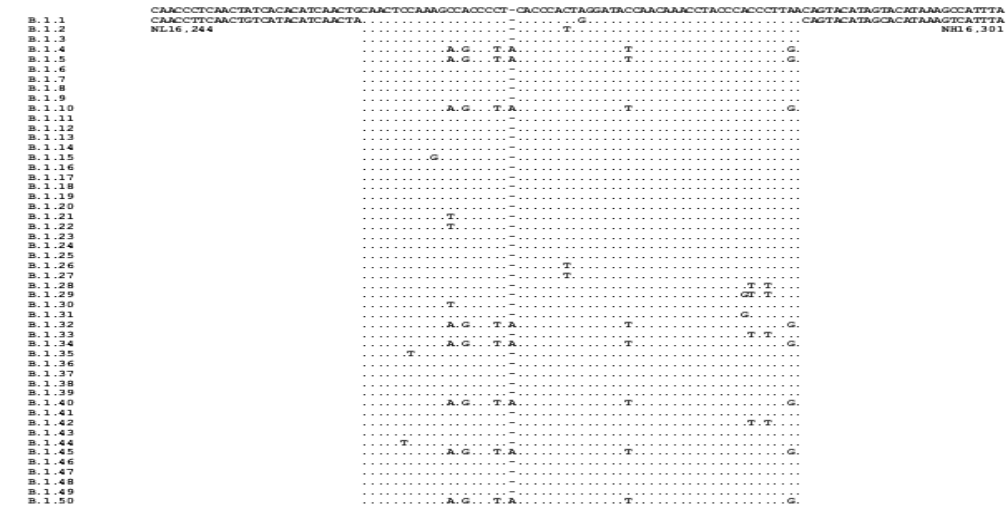
Supplementary Table 1:

Clone sequences for the different Neandertal samples amplified with and without blocking primers.

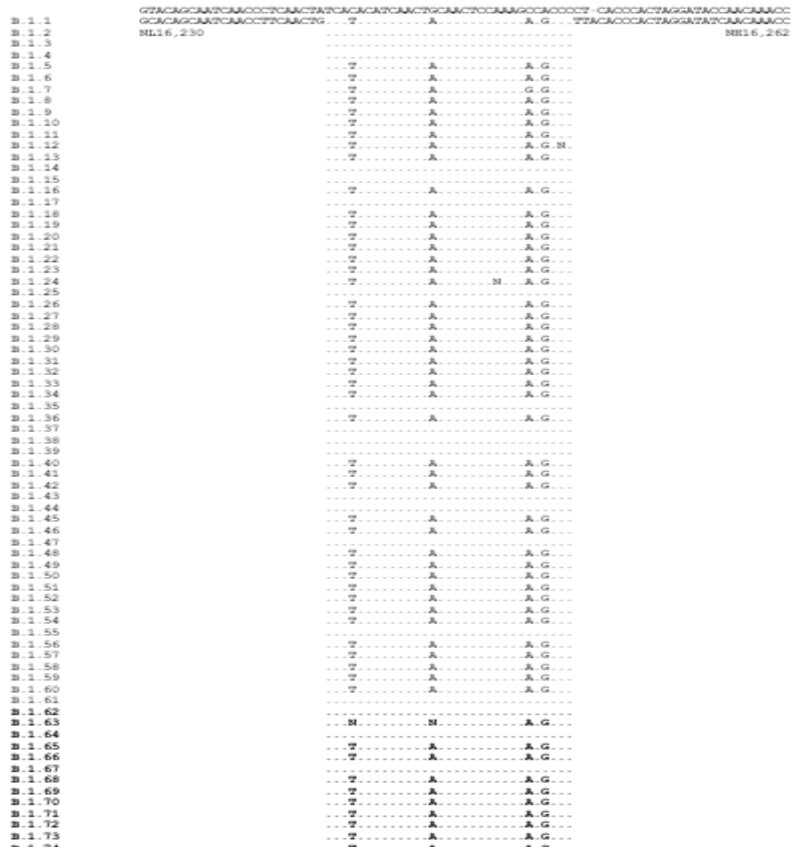
SD 1253 (with blocking primers)



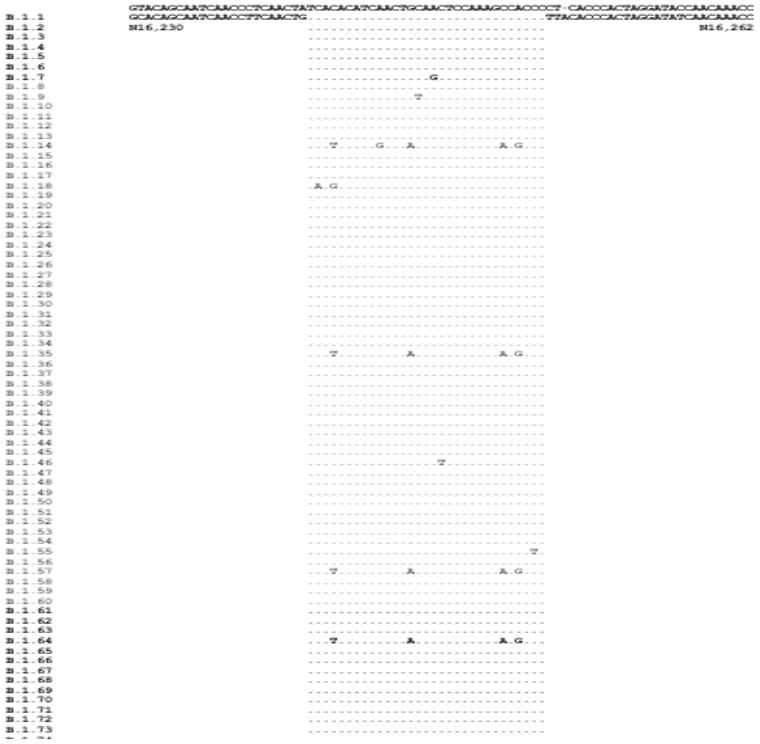
SD 1253 (without blocking primers)



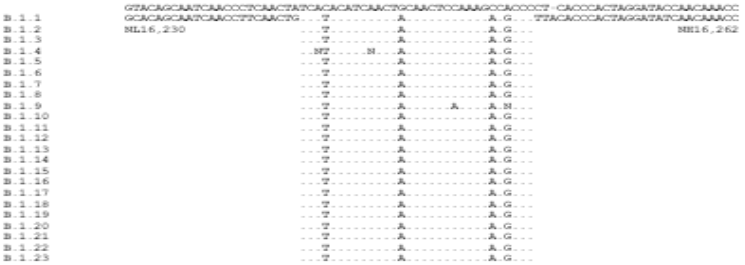
SD 441 (with blocking primers)



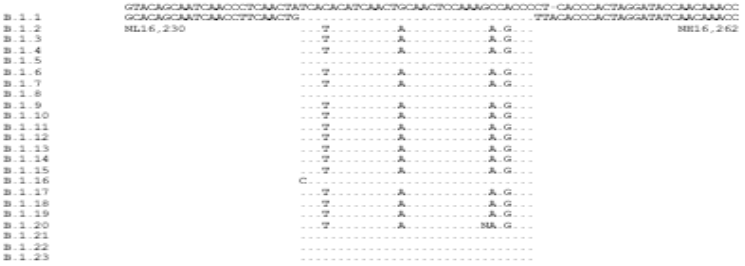
SD 441 (without blocking primers)



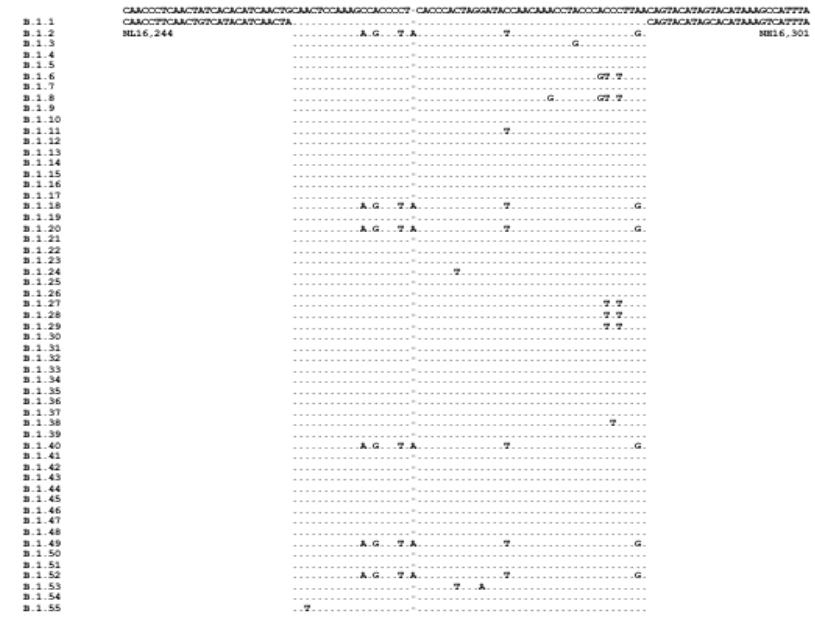
SD 1161 (with blocking primers)



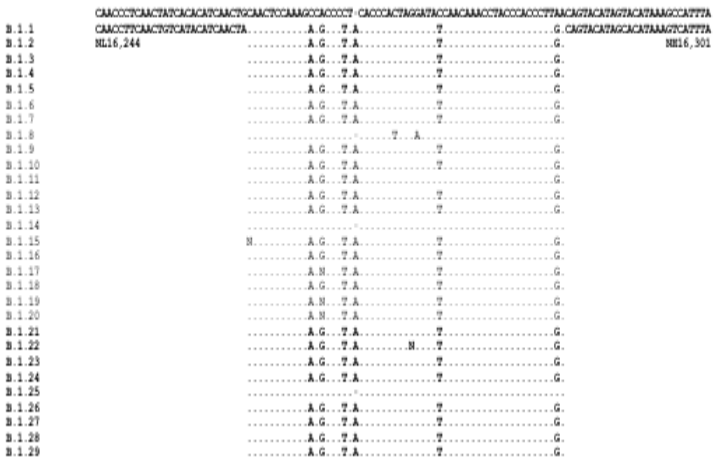
SD 1161 (without blocking primers)



SD 1351c (without blocking primers)



SD 1351c (with blocking primers)



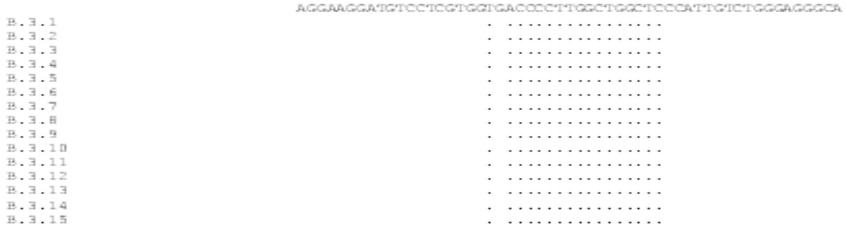
B2. Supplementary information to chapter 3.4.

Barcelona:

```
Sidrón 1253 PCR1

EXON 6; 261; P20144 R20163
                                AGGAAGGATGTCCTCGTGGTGACCCCTTGGCTGGCTCCCATTGTCTGGGAGGGCA
B.1.1                           . ................
B.1.2                           . ................
B.1.3                           . ................
B.1.4                           . ................
B.1.5                           . ................
B.1.6                           . ................
B.1.7                           . ................
B.1.8                           . ................
B.1.9                           . ................
B.1.10                          . ................
B.1.11                          . ................
B.1.12                          . ................
B.1.13                          . ................
B.1.14                          . ................
B.1.15                          . ................

Sidrón 1253 PCR1

EXON 6; 297; P20173 R20194
                                                    TTGGCTGGCTCCCATTGTCTGGGAGGGCACATTCAACATCGACATCCTCAACGAGCAGTTC
B.1.1                                               ...................
B.1.2                                               ...................
B.1.3                                               ...................
B.1.4                                               ...................
B.1.5                                               ...................
B.1.6                                               ...................
B.1.7                                               ...................
B.1.8                                               ...................
B.1.9                                               ...................
B.1.10                                              ...................
B.1.11                                              ...................
B.1.12                                              ...................

Sidrón 1253 PCR2

EXON 6; 261; P20144 R20163


                                AGGAAGGATGTCCTCGTGGTGACCCCTTGGCTGGCTCCCATTGTCTGGGAGGGCA
B.2.1                           . ................
B.2.2                           . ................
B.2.3                           . ................
B.2.4                           . ................
B.2.5                           . ................
B.2.6                           . ................
B.2.7                           . ................
B.2.8                           . ................
B.2.9                           . ................
B.2.10                          . ................
```

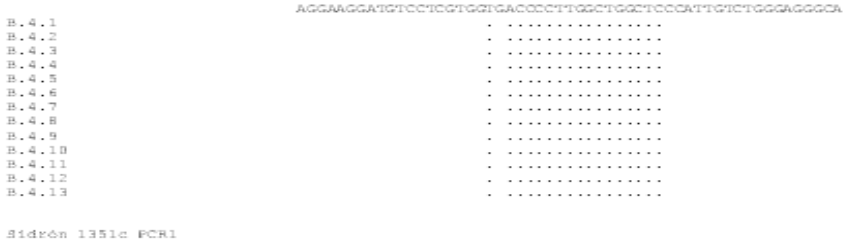
Sidron 1253 PCR3

EXON 6; 261; P20144 R20163



Sidron 1253 PCR4

EXON 6; 261; P20144 R20163



Sidrón 1351c PCR1

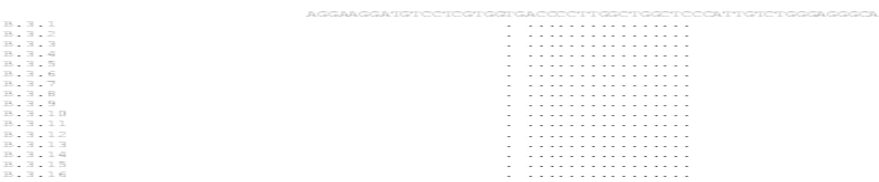
EXON 6; 261; P20144 R20163

Sidrón 1351c PCR2

EXON 6; 261; P20144 R20163

Sidrón 1351c PCR3

EXON 6; 261; P20144 R20163



Sidrón 1351c PCR4

EXON 6; 261; P20144 R20163

Sidrón 1253

Mitochondrial control (NL16,135-NH16,169)

```
TGAATATTGTACGGTACCATAAATACTTGACCACCTGTAGTACATAAAAACCCAATCCACATCAAAACCC
B.1.1                ...T........T.....C..............
B.1.2                ...T........T.....C..............
B.1.3                ...T........T.....C..............
B.1.4                ...T........T.....C..............
B.1.5                ...T........T.....C..............
B.1.6                ...T........T.....C..............
B.1.7                ...T........T.....C..............
B.1.8                ...T........T.....C..............
B.1.9                ...T........T.....C..............
B.1.10               ...T........T.....C..............
B.1.11               ...T........T.....C..............
B.1.12               ...T........T.....C..............
```

Sidrón 1253

Mitochondrial control (NL16,182-NH16,223)

```
AACCCAATCCACATCAAAACCCCCTCCCC--ATGCTTACAAGCAAGTACAGCAATCAACCCTCAACTATCACACATCAACTGC
B.2.1             C.....C....--................C.............
B.2.2             C.....C.....................C.............
B.2.3             C.....C....--................C..N..........
B.2.4             C.....C.----.................C.............
B.2.5             C.....C....--.........N.....C.............
B.2.6             C.....C...---................C.............
B.2.7             C.....C.....................C.............
B.2.8             C.....C....--................C.............
B.2.9             C.....C....--................C.............
```

Sidrón 1253

Y2 control

```
            AACTTAATCAGATTTAGGACACAAAAGCTACTACATAATGAAAAAGAGAGCTGGTGACTT
B.1.1                                 ......A.........
B.1.2                                 ......A.........
B.1.3                                 ......A.........
B.1.4                                 ......A.........
B.1.5                                 ......A.........
B.1.6                                 ......A.........
B.1.7                                 ......A.........
B.1.8                                 ......A.........
```

Sidrón 1253

Mitochondrial control

```
mtDNA 6,267       TGCTCGCATCTGCTATAGTGGARGCCGGMGCAGGAACAGGTTGAACAGTCTACCCTCCCTTAGCA
B.1.1                                 ..A.....C..............
B.1.2                                 ..A.....C............A..
B.1.3                                 ..A.....C..............
B.1.4                                 ..A.....C..............
B.1.5                                 ..A.....C..............
B.1.6                                 ..A.....C..............
B.1.7                                 ..A.....C..............
B.1.8                                 ..A.....C..AA..........
B.1.9                                 ..A...T.C..............
B.1.10                                ..A.....C..............
B.1.11                                ..A.....C..............
B.1.12                                ..A.....C..............
B.1.13                                ..A.....C..............
B.1.14                                ..A.....C..............
B.1.15                                ..A.....C..............
B.1.16
```

Sidrón 1351C

Mitochondrial control

```
mtDNA 6,267        TGCTCGCATCTGCTATAGTGGARGCCGGMGCAGGAACAGGTTGAACAGTCTACCCTCCCTTAGCA
B.1.1                                 ..A.T...C..................
B.1.2                                 ..A.....C..................
B.1.3                                 ..A.....C..................
B.1.4                                 ..A.....C..................
B.1.5                                 ..A.T...C..................
B.1.6                                 ..A.....C..................
B.1.7                                 ..A.....C..................
B.1.8                                 ..A.....C..................
B.1.9                                 ..A.....C..................
B.1.10                                ..A.....C..................
B.1.11                                ..A.....C..................
B.1.12                                A.A.....CA.A..........N
B.1.13                                ..A.....C..................
B.1.14                                ..........................
```

Leipzig:

Sidrón 1253 PCR1

Mitochondrial control

```
mtDNA 6,267        TGCTCGCATCTGCTATAGTGGARGCCGGMGCAGGAACAGGTT GAACAGTCTACCCTCCCTTAGCA
Consensus CV4 13                      ..A.....C..................  ...
E9MLHLIB1AINIX                        ..........................  ...
E9MLHLIB1AJ42A RC                     ..........................  ...
E9MLHLIB1AQP4F RC                     ..........................  ...
E9MLHLIB1AQDKI                        ..........................  ...
E9MLHLIB1ATPHD RC                     ..........................  ...
E9MLHLIB1AMRQ7 RC                     ..........................  ...
E9MLHLIB1AMOHD                        ..........................  ...
E9MLHLIB1AJCCA                        ..........................  ...
```

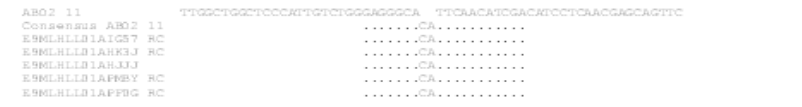
Sidrón 1253 PCR1

EXON 6; 261; P20144 R20163

```
ABO1 10            AGGAAGGATGTCCTCGTGGTG ACCCCTTGGCTGGCTCCCATTGTCTGGGAGGGCAT
Consensus ABO1 10                     .  ..................
E9MLHLIB1APHGG                        .. ..................
E9MLHLIB1AR6CR                        .. ..................
E9MLHLIB1ANI24                        .. ..................
E9MLHLIB1AOTN2                        .. ..................
E9MLHLIB1AR594                        .. ..................
E9MLHLIB1AO6PM                        .. ..................
E9MLHLIB1AMIYB                        .. ..................
E9MLHLIB1AI3TG RC                     .. ..................
E9MLHLIB1AMFNX                        .. ..................
E9MLHLIB1AIRS7                        .. ..................
E9MLHLIB1AID2N RC                     .. ..................
E9MLHLIB1ALVT1                        .. .... ..............
E9MLHLIB1ALTRW RC                     .. ..................
E9MLHLIB1AOQQX                        .. ..................
E9MLHLIB1ARTFF RC                     .. ..................
E9MLHLIB1ATPFW                        .. ..................
E9MLHLIB1AOI9X                        .. ..................
E9MLHLIB1AOY2L                        .. ..................
E9MLHLIB1AO2T1 RC                     .. ..................
E9MLHLIB1AMODJ RC                     .. ..................
E9MLHLIB1AKV6O RC                     .CG.... ..............
E9MLHLIB1ATN1P RC                     .  ..................
E9MLHLIB1AMZXF                        .  ..................
E9MLHLIB1AIQ6I RC                     .  ..................
E9MLHLIB1AG9EW RC                     .  ..................
E9MLHLIB1ANT72 RC                     .  ..................
E9MLHLIB1ALV22                        .  ..................
E9MLHLIB1AHII2 RC                     .  ..................
E9MLHLIB1ATGL0 RC                     .  ..................
E9MLHLIB1AH5T4                        .  ..................
E9MLHLIB1ALY5C                        .  ..................
E9MLHLIB1AK2KH RC                     .  ..................
E9MLHLIB1ATCK2 RC                     .  ..................
E9MLHLIB1ATCNV RC                     .  ..................
E9MLHLIB1AKYDE RC                     .  ..................
E9MLHLIB1ASN7E RC                     .  ..................
E9MLHLIB1AMFEX RC                     .  ..................
E9MLHLIB1AOJYC                        .  ..................
E9MLHLIB1AKE5O RC                     .  ..................
E9MLHLIB1ARQDW RC                     .  ..................
E9MLHLIB1ARBUD                        .  ..................
E9MLHLIB1AJTBP RC                     .  ..................
E9MLHLIB1AQBNB RC                     .  ..................
E9MLHLIB1AQXDC RC                     .  ..................
E9MLHLIB1AIO4P                        .  ..................
E9MLHLIB1AN494 RC                     .  ..................
E9MLHLIB1AHMDC RC                     .  ..................
```

Sidrón 1253 PCR1

EXON 6; 297; F20173 R20194



Sidrón 1253 PCR2

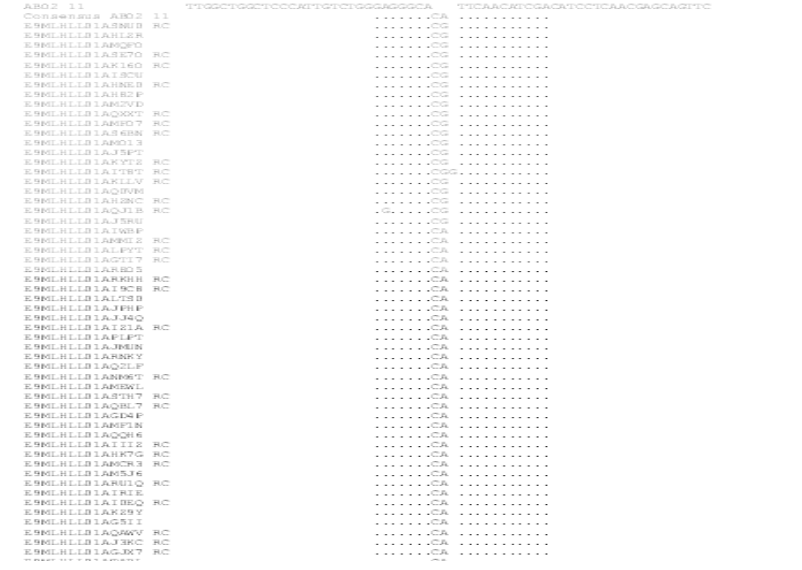
Mitochondrial control

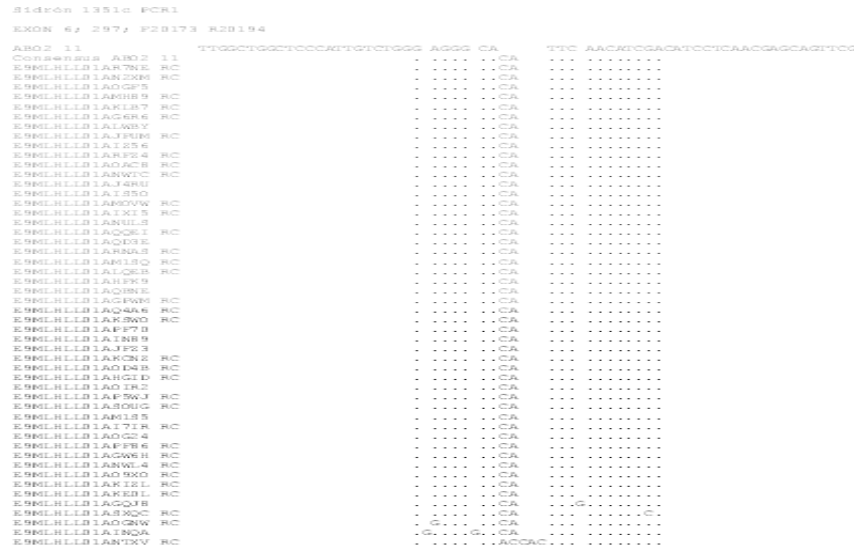
Sidrón 1253 PCR2

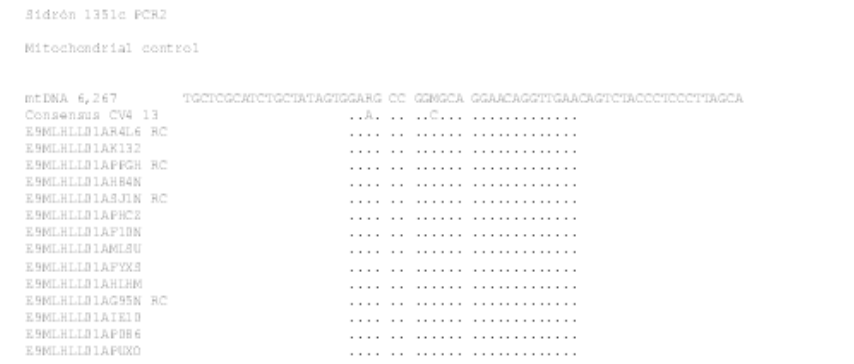
EXON 6; 261; F20144 R20163

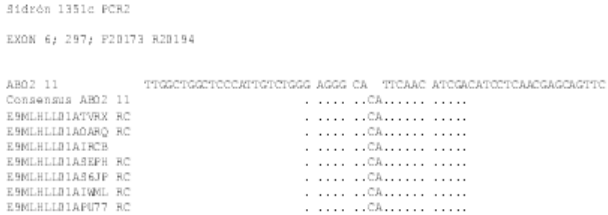
Sidrón 1253 PCR2

EXON 6; 297; F20173 R20194




Sidrón 1351c PCR1
EXON 6; 297; P20173 R20194
ABO2 11 TTGGCTGGCTCCCAATTGTCTGGG AGGG CA TTC AACATCGACATCCTCAACGAGCAGTTCG
Consensus ABO2 11
E9MLHLLD1AR7NE RC
E9MLHLLD1AN2KM RC
E9MLHLLD1AOGP5
E9MLHLLD1AMHB9 RC
E9MLHLLD1AKLB7 RC
E9MLHLLD1AG6R6 RC
E9MLHLLD1ALWBY
E9MLHLLD1AJPUM RC
E9MLHLLD1AI2S6
E9MLHLLD1ARP24 RC
E9MLHLLD1AOJCB RC
E9MLHLLD1ANWTC RC
E9MLHLLD1AJ4RU
E9MLHLLD1AI3SO
E9MLHLLD1AMOVW RC
E9MLHLLD1AIXI5 RC
E9MLHLLD1ANUL3
E9MLHLLD1AQQK1 RC
E9MLHLLD1AQD3E
E9MLHLLD1ARNA3 RC
E9MLHLLD1AM18Q RC
E9MLHLLD1AIQE3 RC
E9MLHLLD1AHFK9
E9MLHLLD1AQBNE
E9MLHLLD1ACPM9 RC
E9MLHLLD1AQ4A6 RC
E9MLHLLD1AK3W0 RC
E9MLHLLD1AFF70
E9MLHLLD1AINB9
E9MLHLLD1AJF23
E9MLHLLD1AKCN2 RC
E9MLHLLD1AOD4B RC
E9MLHLLD1AHG1D RC
E9MLHLLD1AOIK2
E9MLHLLD1AP5WJ RC
E9MLHLLD1ASOUG RC
E9MLHLLD1AM1S5
E9MLHLLD1AI71R RC
E9MLHLLD1AOG24
E9MLHLLD1APFB6 RC
E9MLHLLD1ACWEH RC
E9MLHLLD1ANWL4 RC
E9MLHLLD1AO9XO RC
E9MLHLLD1AKIEL RC
E9MLHLLD1AKEDL RC
E9MLHLLD1AGQJ8 RC
E9MLHLLD1A8XQC RC
E9MLHLLD1AOGNW RC
E9MLHLLD1AINQA RC
E9MLHLLD1ANTXV RC


Sidrón 1351c PCR2
Mitochondrial control
mtDNA 6,267 TGCTCGCATCTGCTATAGTGGARG CC GSMGCA GGAACAGGTTGAACAGTCTACCCTCCCTTAGCA
Consensus CV4 13
E9MLHLLD1AR4L6 RC
E9MLHLLD1AK132
E9MLHLLD1APPGH RC
E9MLHLLD1AH84N
E9MLHLLD1A8J1N RC
E9MLHLLD1APHC2
E9MLHLLD1AF1DN
E9MLHLLD1AML8U
E9MLHLLD1APYX8
E9MLHLLD1AHIHM
E9MLHLLD1AG95N RC
E9MLHLLD1AIE1D
E9MLHLLD1APD86
E9MLHLLD1APUX0


Sidrón 1351c PCR2
EXON 6; 297; P20173 R20194
ABO2 11 TTGGCTGGCTCCCAATTGTCTGGG AGGG CA TTCAAC ATCGACATCCTCAACGAGCAGTTC
Consensus ABO2 11
E9MLHLLD1ATVRX RC
E9MLHLLD1AOARQ RC
E9MLHLLD1AIRCB
E9MLHLLD1A8EPH RC
E9MLHLLD1A86JP RC
E9MLHLLD1AIWML RC
E9MLHLLD1APU77 RC