ANCIENT DNA IN ARCHAEOLOGY

Alfonso Sánchez Hermosilla

Medical Doctor and Forensic Anthropologist

INTRODUCTION

As occurs in the field of Forensic Medicine and Forensic Sciences, and also in Anthropology, in Archaeology, with the passage of time, the information provided by genetic studies has become more important.

However, genetic research on ancient biological material of interest in archaeology is more complicated than research on modern material, whether in the field of clinical genetics or forensic genetics. The main reasons, especially when compared to recent DNA, also known as "fresh DNA", are the following: its scarcity, its high level of fragmentation, contamination with exogenous DNA, the presence of inhibitory substances and the molecular modifications that the original genetic material under investigation has undergone. It is therefore "low quality DNA". Circumstances that make it difficult to investigate and draw valid conclusions.

THE ANCIENT DNA

In Anglo-Saxon literature, it is common to find the term ***ancient DNA,*** which can be translated as ancient DNA (aDNA), and defines the DNA recovered from biological remains preserved naturally or artificially. The scientific community has tried to reach a consensus on the minimum age that a biological sample must have in order to be considered ancient, establishing limits, sometimes arbitrary, since in practice, to date, there does not seem to be a significant difference between DNA that is 100 years old and DNA that is more than 10,000 years old (Pääbo, 1989). In order to avoid this problem, most researchers do not make any distinction, and consider DNA to be all DNA that has undergone an autolytic and/or diagenetic process, as they are all in a very similar situation.

In other bibliographic sources, under the name of "ancient DNA", a series of chemical and genetic techniques are brought together to obtain information on the genome from ancient biological samples in order to verify their authenticity, but also to know their physico-chemical characteristics.

This is due to a very specific circumstance: DNA is a very unstable molecule from the perspective of chemistry.

The stability of DNA depends entirely on the **phosphate-ester chemical bridges** maintaining their morphology and stability, (Eglinton and Logan, 1991), these bridges are very easily destroyed spontaneously by hydrolysis processes, resulting in an extremely low preservation potential.

This type of chemical damage especially affects the purines, adenine and guanine, which causes a breakage of the double helical chain (Lindahl and Anderson, 1972).

Hydrolytic damage can also result in deamination of Cytosine to Uracil, and to a lesser extent to Thymine to Hypoxanthine.

On the other hand, DNA is also affected by other processes, such as **oxidative damage**, which mainly affects pyrimidines, Cytosine and Thymine. This phenomenon occurs spontaneously in aqueous DNA solutions, leading to an accumulation of modifications in the genetic material from the remains of living beings which, once dead, no longer have the necessary cellular repair mechanisms, especially when the DNA is fully hydrated (Lindhl, 1993).

It should be noted that it is not possible to completely dehydrate a DNA chain, as some water molecules are structurally necessary for the genetic material itself, which is why hydrolytic damage cannot be completely avoided, with the exception perhaps of some spores of primitive and simple beings, in which the DNA is naturally packaged, adopting a condensed and dehydrated morphology (Lindahl, 1993), something that does not occur with the genome usually present in archaeological artefacts.

Likewise, the amino acids in proteins that are associated with DNA structure, histones, also degrade over time, especially aspartic acid, which undergoes rapid **racemisation**, comparable to the rate of DNA **purification**, as do other amino acids, albeit at a slower rate.

Finally, it should be noted that the autolytic degradation of DNA begins immediately upon the death of the living being that possesses it, and progresses without stopping over time; this phenomenon is partly due to the autolytic enzymes present in the cellular structures themselves, but also in those of the organisms responsible for the cadaveric phenomena, i.e. bacteria, fungi, algae, insects, etc. In turn, these organisms, when they die, also add their own DNA, thus contaminating the genome of the organic matter on which they were feeding.

To ensure the preservation of ancient DNA in soft tissues, the cellular structure of the tissues, in particular the nuclear membrane, must be preserved.

It is very important to avoid contamination of a sample when removing it from its archaeological context. Not only at the site itself, but also during excavation, during transport, in the museum or institution where the sample is exhibited or stored, and finally in the genetics laboratory where the sample is studied. In reality, the biggest problem when working with genetic material is contamination with **DNA exogenous** to the sample, since the techniques used are so sensitive that a small amount of contaminating exogenous DNA from other organisms close to and different from those of the sample studied will also be amplified together with the **endogenous DNA**, thus providing erroneous results. This includes DNA from micro-organisms, fungi, putrefactive fauna, and even DNA transfer from one specimen to another of the same species.

Researchers themselves are a source of **contaminating DNA** to be taken into account, and it is brought in through skin shedding or aerosols from breathing. The same can be said of the genetics laboratory itself, where it is not always easy to detect and eradicate sources of contaminating DNA, which may come from previous experiments, known as "*carry-over* contamination". In order to minimise these risks, the following authenticity criteria have been proposed (Pääbo et al. 2004):

1. Analysis in a dedicated laboratory of ancient DNA.
2. Separation of pre-PCR and post-PCR working samples.
3. Use of exclusive instruments and equipment.
4. Removal of the surface layer of the sample, as this is the layer with the highest level of contamination.
5. Analysis by a single researcher.
6. Parallel processing of extraction and amplification targets.
7. Use of different primers to amplify the same fragment.
8. Inverse correlation between amplicon size and the intensity of the amplification product visualised on an agarose gel.
9. Phylogenetic sense. That is, in non-human samples, the sequences obtained are similar to those of other members of the same or related species, and in human samples, the comparison of the sequences obtained is similar to those of other researchers, archaeologists or anthropologists.
10. Carrying out several extractions of the same specimen and several amplifications of the same extract.
11. Replication of the entire experimental process in an independent laboratory.
12. Biochemical preservation assays for other molecules in the sample. The most commonly used method, where possible, is the analysis of the amino acids in the sample (total content and degree of racemisation).
13. Quantification of the number of template DNA molecules in the extracts. This allows an assessment of the probability that the mutations observed in the DNA sequences are a direct consequence of postmortem oxidative modifications of the original DNA. In general, it is considered that when the number of original DNA molecules exceeds 1,000, the probability of this phenomenon is low.
14. Bacterial cloning of the amplification products and sequencing of multiple clones. This methodology allows to detect any heterogeneity present in the amplification products and to identify the source of heterogeneity: contamination, molecular modifications and/or polymerase errors.

Moreover, **mitochondrial DNA** (mDNA) is not associated with histones, but with *binding proteins,* which gives it a certain level of protection compared to nuclear DNA (nDNA), according to Kelnman and Moran, 1996. In addition, the mitochondrial membrane contains a large number of enzymes that have a very high affinity for DNA (Cho et al. 1998), which also helps to protect mDNA from degenerative phenomena.

And all this without forgetting the damage caused by the environmental conditions to which the biological remains have been subjected from the time of the death of the individual until their arrival at the genetics laboratory.

It is for all these reasons that the ancient DNA present in the samples of archaeological interest is highly degraded, showing frequent molecular cross-linking, partly due to hydrolytic damage, but also showing modifications in its pyrimidine bases, mainly due to oxidative damage. In any case, even if the molecular morphology of the DNA and its physico-chemical integrity are altered, its components can persist for long periods of time and are susceptible to investigation.

The difficulties caused by the analysis of ancient DNA are as follows:

- Low *Copy Number (LCN*) of DNA molecules from the sample.
- The fragmentation and postmortem alterations that this DNA undergoes.
- The scarcity of intact fragments endogenous to the sample, which is often contaminated with exogenous DNA, either from the external environment or from people who have been in contact with the sample at some point in its history, or even the researchers, archaeologists, anthropologists, laboratory personnel, etc. themselves.
- Impurities provided by the archaeological site context, mainly soil, often have an inhibitory effect on the chemical reactions that make DNA amplification and sequencing possible.
- The presence of **PCR inhibitors**, many of them still unknown in nature, present both in the sample itself and in the soil; some of these inhibitory agents have been described, including humic or fulvic acids, porphyrin residues, type I collagen, Maillard products or even the damage to the old DNA itself when it is extensive and reaches such a degree that the DNA degradation products themselves inhibit PCR.

In order to deal with all these problems, a set of standardised procedures known as **Authenticity Criteria** have been established that contemplate the application of a series of preventive measures to be developed, not only in the laboratory, but also in terms of the DNA isolation methodology, the infrastructures themselves, the instruments and the application of specific techniques. Thus, a specific ancient DNA result will only be validated by the scientific community when these authenticity criteria are observed (Pääbo et al. 2004).

Sometimes, environmental conditions, or special circumstances that facilitate preservation, can reduce the rate of DNA degradation, but it is impossible to completely stop the damage to this fragile molecule. This imposes a time limit on DNA preservation under all conditions. Circumstances that favour the preservation of the ancient genome are:

1. Rapid dehydration of biological material after death or deposition on the archaeological site, which reduces hydrolytic damage.
2. Anaerobic conditions in the medium in which the sample is kept throughout its history.
3. High ionic strengths.
4. High presence of tannins or humic acids in the soil.
5. Association with proteins of chromosomal or even other origin, which can promote DNA preservation and slow down the degradative activity of micro-organisms.

In reality, it is the specific environmental conditions in each individual case that largely determine the degree of preservation of ancient DNA.

HISTORY OF ANCIENT DNA RESEARCH

The first studies on ancient DNA date back to the 1980s and involved research on small genetic sequences from the skin of mummified animals and humans. From the very beginning, the difficulty of this research became apparent (Higuchi et al. 1984; Pääbo, 1985, 1989), as most of the genome obtained had a contaminating origin from microbial and fungal agents, while a small part of it originated from biological material of archaeological interest, which was also badly damaged.

The invention **of the PCR** (Polymerase Chain Reaction) **technique** represented a turning point in the initial situation, as it allowed the amplification of small DNA fragments, allowing samples that could not provide any genetic information with the previous methodology to be studied satisfactorily.

Another technique that has sometimes been useful is **Bacterial Cloning**, which involves cloning multiple copies of a DNA fragment by inserting it into a self-replicating genetic element such as a plasmid or circular DNA molecule, and by heat shock inserting it into certain bacteria that are easy to grow in laboratory media. Each bacterium receives a single strand of the amplified DNA that is inserted into its genome. The problem is that molecular damage to the DNA can prevent its integration into the bacteria, and even the bacteria's own molecular damage repair mechanisms can cause changes to the DNA.

However, the biggest problem today comes from contamination with DNA foreign to the original sample, which is aggravated when the research is carried out on ancient human DNA, which can be contaminated not only at the perimortem moments of the subject or subjects to be investigated, but also at any time from that moment until the sample arrives at the genetics laboratory. This possible contamination cannot always be detected satisfactorily. It is for this reason, among others, that the term "*Dirty DNA*" is frequently used in the literature.

The methodology is now widely established and accepted in the scientific community, allowing genetic identification of biological species, sometimes even individual diagnosis, i.e. if different biological samples belong to the same individual and not to a close relative, as well as population studies and palaeoecological changes.

In the future, both the methodology available and that which may be developed will undoubtedly find new fields of research, including **Exobiology**, as well as **Paleoexobiology**, in the event that life forms alien to planet earth can be found, and that these, in turn, possess genetic material consisting of DNA and RNA.

METHODOLOGY FOR EXTRACTING ANCIENT DNA

The standard method agreed by the scientific community to extract ancient DNA from a sample of archaeological interest consists of the following steps:

1. **Cleaning and crushing of the sample in** order to extract the genetic material it contains.
2. **Amplification of this genetic material by PCR and post-PCR processing**. PCR *(Polymerase Chain Reaction)* is an amplification method that allows a part of the genome to be selected and copied numerous times - virtually infinitely, as it is possible to obtain millions of copies of a sequence of interest *(Target)* - by *in vitro* synthesis, so that from very small samples with very little genetic material, a sufficient amount of genetic material is obtained for the purpose of being studied. To this end, the DNA to be studied is mixed with an enzyme capable of copying the DNA, this enzyme must be able to withstand high temperatures, for example, *Taq polymerase*, two small DNA fragments called *primers* are added to this mixture, and individual nucleotides (adenine, cytosine, thymine and guanine), together with a pH-stabilising compound consisting of various salts called 'buffer', and magnesium chloride - both of which are necessary for Taq to work.
3. In the first PCR step, the two DNA strands **are separated** (denatured) by subjecting the aforementioned mixture to a temperature of 94ºC. In a second phase, called annealing, the two primers, which have sequences complementary to two flanking regions of the DNA section to be selected and copied, are joined to the DNA under study. This process is carried out at a temperature of 50-60°C. Then, in a third step, Taq polymerase copies each of the two DNA strands by adding nucleotides from the primers attached to each of them.
4. This process or cycle is then repeated 30-40 times to achieve exponential copying of the DNA fragment under study. This procedure is called **amplification**.
5. Finally, the method known as **sequencing** is applied, a technique that allows the sequence of nucleotides that make up the fragment to be analysed to be known. Sequencing can be direct, analysing the product obtained by PCR without any additional treatment, or it can also be done by **cloning** the fragments obtained by amplification. The scientific community gives preference to amplification for the simple reason that it is possible to detect possible molecular damage and amplification errors, although it is also a more laborious procedure.

This process can be applied to any ancient biological material, but its yield is higher when using teeth or bones, where DNA preservation is better. On the other hand, the yield is lower on biological material of plant origin, partly also because its DNA content is lower than in material of animal or human origin.

Using this methodology it is possible to study the following types of genetic material:

1. mDNA from old samples by automatic sequencing.
2. Y-chromosome microsatellites (STRs) by fragment analysis (ALF).
3. Nuclear microsatellites (STRs) of non-sex chromosomes by fragment analysis (LFA).
4. Amplification of the amelogenin gene, present on the X and Y chromosomes (sex chromosomes) by fragment analysis (ALF).
5. Quantification of the number of DNA copies present in extracts obtained from biological samples by *Real Time PCR*.
6. Bacterial cloning of amplification products and sequencing of multiple clones.

The ideal material for the study of ancient DNA are dental pieces with an intact external appearance, without visible structural damage, as well as long bones (tibia and femur) but also others such as ribs, radius, fibula, talus, on the condition that they have a thick cortex in their diaphysis, and abundant compact bone tissue inside. In any case, clean skeletal pieces without external fissures shall be used.

APPLICATIONS OF ANCIENT DNA STUDIES IN ARCHAEOLOGY AND ANTHROPOLOGY

Using the available methodology, it is possible to make a sex diagnosis of genetic material from ancient biological samples, to establish family relationships between individuals who may be related, to estimate the minimum number of different individuals present in the same archaeological context, e.g. necropolis or mass graves, as well as population studies to determine whether there is genetic continuity between samples from different periods in the same archaeological site or geographical region, thus assessing the impact of possible migratory movements.

Molecular sex diagnosis is useful in those cases in which the preservation of bone remains, human or animal, does not allow the sex of the individual to be determined by anatomical diagnosis, which occurs when the remains are very fragmented, or belong to infantile individuals, or in male subjects with a gracile skeleton or in female subjects with a robust skeleton. For this purpose, the sequence of the **amelogenin gene**, located on both X and Y chromosomes, is determined, but which in certain parts of its sequence has a differential length. These small differences make it possible to establish the presence of a single type of chromosome if the sample comes from female individuals (XX), or two different types if it comes from male individuals (XY) (Akane et al. 1992; Sullivan et al. 1993). In the case that we find a group of individuals in a site, the determination of the percentage of sexes in the population can contribute to support, or refute, different hypotheses in the field of anthropology or archaeology.

In the event that multiple burials are found at a site, it can be useful to determine **Biological Parentage Relationships**, which can help to understand possible burial patterns, as well as specific social behavioural characteristics. For this purpose, genetic markers are used, especially autosomal markers, of nuclear origin, although information provided by haploid inheritance markers such as mitochondrial DNA, but also nuclear DNA, specifically

Y-chromosome polymorphisms, is also useful. Of these, the autosomal systems are the most likely to provide information. However, as there are two copies per cell nucleus, one from the biological father and the other from the mother, it is a major problem to recover them when the biological sample is degraded, which unfortunately happens very often. On the other hand, mitochondrial DNA is less polymorphic than nuclear DNA, but it only provides information exclusively from maternal inheritance, although sometimes, exceptionally, the presence of mitochondrial DNA of maternal origin, and to a lesser extent, also of paternal origin, has been described in the same individual. For this purpose, family relationships can be established in terms of mathematical probability, as is the case in forensic cases.

Personal Identification has been used to identify historical figures, most notably in the case of research on alleged members of the Tsarist royal family, corpses or not (Gill et al. 1994), especially in the case of Princess Anastasia, using a set of autosomal and mitochondrial DNA markers (Gill et al. 1994; Coble et al. 2009). The results are also expressed in terms of mathematical probability.

An increasing number of publications refer to the determination of blood grouping (ABO) with genetic methods, and even to the study of point mutations (SNPs) that allow the determination of phenotypic characteristics.

Another possible application is **Palaeoepidemiology**, as it is possible to genetically characterise infectious diseases that often cannot be detected by other means in biological samples. For this, the amplification and detection of specific sequences of certain infectious agents can help to determine the cause of death of an individual or group of individuals, and even corroborate other hypotheses of archaeological and historical interest. It is also possible to detect genetic diseases and their transmission in specific populations.

It is also possible to carry out studies on the **genetic characterisation of ancient populations**, which makes it possible to differentiate some populations from others, their genetic variability, as well as to carry out temporal studies of genetic changes and to draw up hypotheses on their causes: genetic mutations, migrations or mixing of populations. It is also possible, always in terms of probability, to attribute a population and geographical origin to isolated individuals.

It goes without saying that the recovery of DNA from ancient remains is not limited to the study of human remains, but also animal and plant remains, and can even verify archaeological hypotheses relating to the introduction of agriculture and livestock farming in different population, geographical and historical contexts. To this end, research has been carried out on seeds of the genus *Triticum L.*, insect pupae that accompanied mummies, coprolites, plant pollens, breeds of livestock interest to distinguish between ovids and caprids, and even extinct species. Some of the remains investigated come from insects trapped in amber thousands or even millions of years ago. It is also possible to

carry out studies on the diet of ancient populations (Palaeodiet), or even on the diet of extinct animals.

BIBLIOGRAPHY

- Archaeometry and Archaeological Analysis. **Ancient DNA identification.** https://www.ucm.es/arqueoanálisis/identificación-adn-antiguo [accessed 2017-05/15].
- Brenes G., Bustillos C., Cadena A., Díaz N., Sánchez G. ***Adn Antiguo.*** https://www.clubensayos.com/Ciencia/Adn-antiguo/475811.html [Accessed 2017-05-15].
- Laboratory of Forensic Genetics and Population Genetics of the Complutense University of Madrid. ***Problematic of ancient DNA studies.*** http://www.bing.com/search?q=%22aden+antique%22&src=IE-TopResult&FORM=IE10TR [Accessed 2015-05-16].
- Palomo Díez S., Gomes C. ***Introduction to ancient DNA in Archaeology and Anthropology.*** https://www.reserarchgate.net/publication/281971544 [Accessed 2017-05-15].
- Pérez P. ***Ancient DNA and its study in the field of archaeology.*** http://mediosantropologicos2.blogspto.com.es/2012/03/adn-antiguo-y-su-estudio-en-antropologia [Accessed 2017-02-17] review web page.
- Sánchez Hermosilla, A. ***Is it possible to clone Jesus Christ?*** Linteum...
- Sze-wah L., Stephen C.Y. ***Compatibility of DNA IQ™ , QIAamp®DNA Investigator, and QIAsymphony® DNA Investigator®with various fingerprint treatments.*** http://link.springer.com/article/10.1007/s00414-016-1447-8 [Accessed 2017-05-16].
- Willerslev E., Cooper A. ***Ancient DNA.*** https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1634942/ [Accessed 2017-05-15].