

Ancient mitochondrial DNA analysis of human remains in the era of next-generation sequencing

Dina W. Ibrahim^{a,*}, Haidan Mostafa El-Shorbagy^{b,c}, Ismail A. Abdelhamid^d, Yehia Z. Gad^{a,e}

^a Ancient DNA Lab, National Museum of Egyptian Civilization, Cairo, Egypt

^b Zoology Department, Faculty of Science, Cairo University, Giza, Egypt.

^c Faculty of Biotechnology, October University for Modern Science and Art, 6 October, Giza, Egypt

^d Chemistry Department, Faculty of Science, Cairo University, Giza 12613, Egypt

^e Department of Medical Molecular Genetics, Division of Human Genetics and Genome Research, National Research Centre, El-Behouth St., Dokki, Giza 12622, Egypt

ARTICLE INFO

Keywords:

Ancient mitochondrial DNA
mtDNA variations
Miscoding lesions
Paleogenetics
And Bioinformatic tools

ABSTRACT

The analysis of ancient DNA (aDNA) has become important in recent years, providing new insights on human evolution and population history. Since the emergence of ancient DNA analysis in the mid-eighties, mitochondrial DNA (mtDNA) analysis has been an essential tool in paleogenomic research. Compared to nuclear DNA, its small, circular, multicopy genome is collectively less susceptible to chemical damage and fragmentation and has more potential for retrievability. Despite its longevity, mtDNA's maternal inheritance with no recombination allows researchers to reconstruct maternal lineages and thus partially track human ancestry and migratory patterns. The accuracy and authenticity of mtDNA data retrieved from paleontological and archaeological remains have improved significantly over the past three decades, driven by advances ranging from enhanced contamination control to next-generation genome sequencing and target enrichment. The current review highlights the challenges of recovering and authenticating ancient mtDNA, including DNA fragmentation, post-mortem DNA damage, contamination, heteroplasmy, and nuclear mitochondrial sequences (Numts), which can complicate research results. Within this context, the review addresses the rapidly changing sequencing technologies and the use of specialized bioinformatics software that tackle these challenges and enable researchers to reconstruct almost complete mtDNA genomes from human remains. Although mtDNA analysis has limitations and provides only a partial view of the human genomic architecture, it remains an important tool for analyzing maternal ancestry and generating authenticated genomic information of ancient human populations.

1. Introduction

The first ancient DNA extract from biological remains was performed by Higuchi & his colleagues in 1984. They examined the dried muscle of the quagga, a zebra-like species (*Equus quagga*) that went extinct in 1883. They were able to obtain many clones from the very low molecular weight DNA of quagga, and its mitochondrial DNA (mtDNA) was also sequenced.¹ The quagga publication encouraged many researchers to study DNA in other human and non-human remains.

The genetic material recovered from ancient samples revealed a typical pattern of damage at read termini, low copy numbers of endogenous aDNA molecules, high fragmentation, and the presence of exogenous DNA.² The high fragmentation is due to degradation over time, while the typical damage at read termini results from hydrolysis

and oxidation that affect the biological sample. These factors influence the accurate reconstruction of the endogenous ancient genome.

The rapid development of automated DNA sequencing technology has revolutionized the field of human archaeogenetics by moving from analyzing parts of the mitochondrial DNA (mtDNA) genome in ancient samples to reconstructing the complete mtDNA genome in the early 2000s.^{3–6} This was followed by analyzing the whole-genomic data of several ancient individuals.^{7–9}

Over the past few years, several aDNA studies have provided insights into the historical migration paths and relationships between various human populations, shedding light on the large-scale population dynamics of our ancestors.^{10–13} Such information is crucial for reconstructing the investigated pedigrees and understanding human evolution, thereby answering many historical questions.¹⁴

* Corresponding author at: Ancient DNA Laboratory, National Museum of Egyptian Civilization (NMEC), Ain El Sira, Al-Fustat, Cairo 11765, Egypt.

E-mail address: dina43wagdi@gmail.com (D.W. Ibrahim).

<https://doi.org/10.1016/j.jgeb.2026.100828>

Received 14 July 2026; Received in revised form 23 August 2026; Accepted 17 September 2026

Available online 29 September 2026

1687-157X/© 2026 The Authors. Published by Elsevier Inc. on behalf of Academy of Scientific Research and Technology. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Practically speaking, this kind of data might also be a valuable resource for “storytelling” in outreach, science communication, and museum exhibits. The current review aims to address the use of ancient mitochondrial DNA analysis in the era of next-generation sequencing (NGS) and the generation of whole-genomic data for ancient populations.

2. Mitochondrial DNA: structure and characteristics

2.1. Overview of the mitochondrial genome

The mitochondrion is an organelle that has its circular genomic DNA (Margulis, 1970) (Fig. 1). A human cell can contain several thousand copies of a small circular mitochondrial DNA (mtDNA).^{15,16} Human mtDNA includes about 16,600 base pairs and encodes 13 proteins that are required for oxidative phosphorylation.

Most of the mitochondrial genome is coding DNA, unlike nuclear DNA (nDNA), which contains non-coding regions involved in alternative splicing, gene regulation, and other less well-known functions.^{17–19} Consequently, mutations in mtDNA often influence function, thus compromising mitochondrial performance. The intron-free structure of mtDNA simplifies mitochondrial gene regulation, which is governed by transcription factors and epigenetic modifications. Understanding these regulatory processes is essential for assessing the impact of mtDNA variants on complex traits or diseases.^{19,20}

Being highly abundant per cell and maternally inherited with little or no recombination, mtDNA is believed to be a good representative of the maternal lineages in ancient remains.^{21–23} Thus, it can partially be used as a molecular marker to study genetic flow & phylogeny.^{24,25}

2.2. mtDNA Mutation Rates, Haplogroups, and Variants

2.2.1. Mutation rates

Although mtDNA is present in hundreds to thousands of copies per cell according to cell type, the sequence evolution rate of mtDNA genes is quite rapid, possibly 10–20 times that of comparable nuclear DNA (nDNA) genes.^{26–28} This is the result of multiplying a mtDNA mutant fixation rate by a mutation rate, which is possibly 100–1000 times higher than that of nDNA genes: $E = \mu F$ (where E is the rate of sequence evolution, μ is the mutation rate, and F is the fixation rate).²⁸

The mutation rate of mtDNA is higher than that of the nuclear genome due to the lack of proofreading enzymes. So, mtDNA is highly susceptible to mutations, including pathogenic mutations.^{19,28,29} However, not all mtDNA variations are highly deleterious due to their haploid inheritance and their small effective population size.³⁰ In addition, highly deleterious variants are removed by negative selection within populations, which is not fully understood, but prevents their

existence in the mtDNA of an organism.^{31–34}

2.2.2. Haplogroup nomenclature

Mitochondrial haplogroups consist of sets of related haplotypes defined by distinct combinations of mtDNA single-nucleotide polymorphisms (SNPs) relative to a reference sequence, each originating from a common ancestor.^{19,35} Haplogroup nomenclature consists of a combination of numbers and letters. The macro-haplogroup “L” includes haplogroups L0–L5 and their sub-haplogroups and represents the most ancient African mtDNA sequences, dating back to 130,000–200,000 years before present (YBP). About 65,000 YBP. Haplogroup L3 gave rise to the macrohaplogroups M and N. Those migrated out of Africa, where macrohaplogroup M moved to Asia, generating M1–M80, including C, D, G, and Z. Macrohaplogroup N also moved to Asia, generating A, B, F, and O, but also moved towards Europe, generating I, X, W, and R. The haplogroup R gave rise to the remaining European haplogroups H, V, J, U, and T. Then, haplogroups A, C, and D, which originated in north-eastern Siberia, reached America (~ 20,000 YBP). Also, recent migrations to America brought haplogroups B and X.^{28,36} Nowadays, more than 5400 haplogroups and subgroups are described in the phylogenetic tree, established by van Oven et al. which is updated and available online at <http://www.phylotree.org> (build 17)³⁷ (Fig. 2).

Certain algorithms in several tools, including Haplogrep3,^{38–40} HaploFind,⁴¹ MitoMaster,⁴² Phylotree,³⁷ and MitoTool,⁴³ are used to classify mtDNA haplogroups.

These tools rely on selecting a phylogenetic tree to infer haplogroups. Over time, these tools have been optimized to bypass oversights in haplogroup classification. Additionally, haplogroup classification has been enhanced throughout the years by modifying phylogenetic trees.^{19,37}

It's also critical to note that the mitochondrial background (haplogroup) from which a mtDNA variant originated implies its effects. The classic example, haplogroups N and M, both contain the mutant 3394C (Y30H). The mutant 3394C (Y30H) in haplogroup M is necessary for populations living at high altitudes to adapt to high altitude. However, in haplogroup N, it increased the penetrance of milder complex I variants and decreased complex I activity, which is believed to be linked to Leber hereditary optic neuropathy (LHON).⁴⁴

In several studies, scientists have found haplogroup-specific susceptibility to certain disorders.^{45,46} For example, type 2 diabetes is associated with the Asian haplogroup N9a,⁴⁷ and autism is associated with some European mtDNA haplogroups.⁴⁸

2.2.3. mtDNA variants

There are two main types of mtDNA variants: *homoplasmic* and *heteroplasmic*. Normally, an individual contains only one type of mtDNA termed homoplasmy. Among the thousands of copies of mtDNA that exist within a cell, some will undoubtedly experience mutations. Thus, some of the mutated copies of the genome may be passed on to subsequent generations, resulting in a mixture of distinct mitochondrial DNA (mtDNA) genomes. However, before a mature oocyte forms, a genetic bottleneck occurs that maintains homoplasmy as the norm.⁴⁹ The bottleneck reduces the chance of passing on a mixture of wild-type and mutant mtDNA genomes by allowing only a small number of mtDNA molecules to pass through the oocyte during its formation.⁵⁰ Heteroplasmy refers to the presence of multiple mtDNA variants within a single cell. This occurs when a mother transmits both a normal (wild-type) and a mutated version of her mitochondrial genome to her offspring. An individual will therefore possess two versions of mtDNA, usually differing by one base.⁵¹ However, instances of heteroplasmic mtDNA differing at two or even three positions have been reported.⁵²

Homoplasmic variations are further classified into three types: *recent deleterious maternally inherited*, *somatic*, and *ancient adaptive variants*. The ancient adaptive variants are those required for adaptation to different environmental conditions during migration. So, these variants are suggested to be relatively specific to each geographic location and could

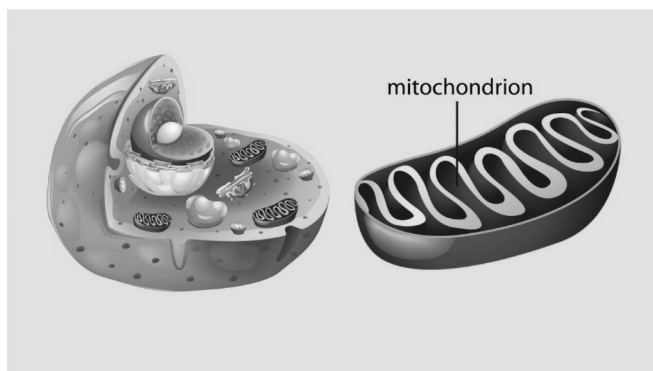


Fig. 1. Structure of the human mitochondrion.

Schematic illustration of the human mitochondrion showing the major structural components of the human mitochondrion.

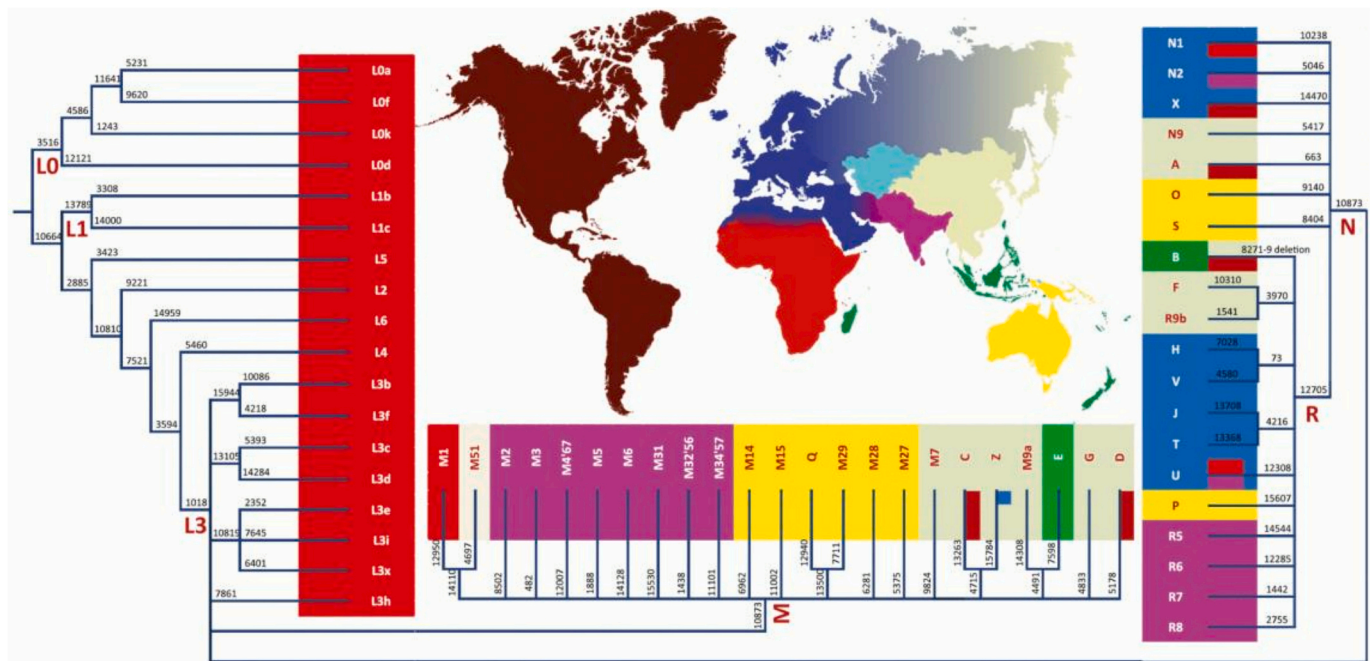


Fig. 2. mtDNA haplogroup tree and distribution map. Haplogroup labels are reported according to the <http://www.phylotree.org/> nomenclature (van Oven M (2009)). Only a single branch defining marker, preferably from the coding region, is shown. The main geographic features of haplogroup distribution are highlighted with color.

thus be used as mtDNA ancestry markers.⁵³

Somatic variants (base substitutions or deletions) accumulate over time in the mtDNA due to both genetic and environmental factors.⁵⁴ It has become increasingly clear that single-nucleotide substitutions, single-base insertions, and single-base deletions are the most common types of mtDNA mutations observed, and that most of us carry a significant number of low-level mtDNA variants in our cells.⁵⁵ Previously, the most frequently detected somatic mtDNA mutations were large deletions that result in smaller, circular mtDNA molecules.^{55,56} Recent advances in next-generation sequencing techniques for whole-genome analysis have enabled the analysis of mtDNA mutational spectra across many tissues in great depth. The findings revealed that somatic mtDNA mutations increased with aging.^{55,57–60} The most common mutations found in these investigations are transition mutations, which are correlated with deamination of cytidine and adenosine, or misincorporation by DNA polymerase γ as the main mutagenic processes in mtDNA.^{58–60}

Overall, findings disputed the idea that oxidative damage is a primary factor in mtDNA mutagenesis. They suggested that the majority of accumulated mtDNA point mutations are due to replication errors and/or spontaneous base hydrolysis.⁵⁵

A *heteroplasmic variant*, which is present in a subset of the mtDNA molecules, can arise in both somatic and germline cells. However, during the replication process (meiotic or mitotic cell division), there is a random shift in heteroplasmic allele frequency, resulting in a variation in the proportion of mutant-to-normal mtDNA in different tissues and organs.^{28,61} The process is not fully understood, but nuclear genes may influence its regulation. Pathogenic heteroplasmic mitochondrial variants are those associated with genetic disorders.⁶² According to reports, over 600 distinct point mutations are currently linked to various human diseases (source: [Mitomap.org](https://www.mitomap.org)).

Recently, with the rapid development of high-throughput sequencing technologies, it has become possible to sequence the Whole genome, leading to a better understanding of mitochondrial diseases. There are hundreds of pathogenic homoplasmic and heteroplasmic mitochondrial variants reported through next-generation sequencing.⁶³ In 1988, a disease-linked mtDNA mutation was first

discovered: mutation 11,778 G > A linked to Leber hereditary optic neuropathy disease.⁶⁴ Although the prevalence of mtDNA variations linked to diseases has been extensively studied in modern populations, their systematic investigation in ancient populations remains limited. For collecting data on mtDNA variations, including their clinical significance and current population frequencies, several publicly accessible databases are available⁶³:

- The GenBank database is an annotated collection of publicly accessible DNA sequences and was used to gather data on variant frequencies in modern populations.⁶⁵
- MitoMap is a database of the human mitochondrial genome that includes 14,383 SNVs, 49135 full-length sequences, and 72,235 control region sequences.⁶⁶ If a variant satisfies the criteria set forth, it is classified as “confirmed pathogenic”.
- MITOMASTER is used to analyze human mitochondrial SNVs or nucleotide sequences.⁴² The MITOMAP-MITOMASTER website is a comprehensive catalogue of mitochondrial variants, which is publicly available at <https://www.mitomap.org/MITOMAP>.
- Mito TIP is an in silico tool that predicts the probability that a novel single-nucleotide variant will cause disease.⁶⁷
- HmtVar is a variant-focused database that provides information on mtDNA variability and pathogenicity.⁶⁸

Several ancient DNA studies have reported mtDNA variants that have been associated with disease or classified as pathogenic based on evidence from modern populations. For example, 25 Thracian mtDNA samples,⁶⁹ dating from 3000 to 2000 BCE, were whole-genome sequenced. Of them, 608 mtDNA variants were found, and only one (m.15326 A > G (rs2853508)) was identified as likely pathogenic and linked to breast cancer.⁷⁰ The MitoMap database (2019, update nr. 3) was used to estimate the frequency of this variant across all analyzed samples, yielding a 0.98 population frequency. The high frequency means that this variant is common and thus probably not disease-related.⁶⁶

In another study, analyzing ancient mtDNA identified six variants as “confirmed pathogenic” in contemporary patients. The very first of these

was found in a Neolithic sample from 5800 to 5400 BCE that carried the MT-TS1 gene mutation m.7510 T > C. A sample from a Bronze Age site in Germany (2029–1911 BCE) revealed another variant, m.3243 A > G (rs199474657), which is one of the most prevalent mtDNA variants that putatively associated with mitochondrial disease.⁶³

2.3. Sequence miscoding and postmortem damage in ancient remains

Miscoding lesions are a distinctive feature of aDNA studies.⁷¹ The observed miscoding lesions have been suggested to result from two mechanisms. The first is due to frequent errors by DNA polymerase during PCR amplification, in which incorrect nucleotides are accidentally incorporated. Such errors can lead to modifications of template molecules during the early stages of PCR, when the initial concentration of template molecules is low. The second mechanism is the creation of errors due to postmortem biochemical damage to the initial template molecules.⁷¹

The moment a cell dies, damage to the DNA molecule begins, with reactive oxygen species and other chemicals, as well as biomolecular processes—particularly the uncontrolled activity of lysosomal nucleases—blaming.⁷² Because compartmentalization keeps these destructive enzymes under control throughout the cell's life, endogenous DNA remains intact within the cell. Furthermore, during cell life, various repair mechanisms can correct errors in the DNA sequence and prevent endogenous DNA degradation caused by reactive oxygen species and other chemical agents.⁷³

A variety of different types of damage can occur when damage to the DNA molecule is not detected during cell life or happens after death, including single- and double-stranded breaks, oxidative or hydrolytic lesions, and cross-linking of DNA to other biomolecules such as proteins, sugars, fatty acids, RNA, and DNA.⁷⁴ According to earlier research, different types of damage can affect the DNA molecule and often result in sequence variation across duplicates from the same person.²

Both hydrolytic and oxidative attacks on the phosphodiester backbone, as well as enzymatic processes, are responsible for the extreme reduction in the size of ancient DNA sequence fragments.⁷⁵ Apurinic/aprimidinic (AP) sites are created when the N-glycosyl bond, which connects the deoxyribose sugar to the nitrogenous base, is hydrolytically attacked (Fig. 3).

During in vitro amplification, AP sites themselves can cause strand breakage or inhibit DNA polymerase activity.⁷⁵ The study of the processes that impact biological material after death is known as taphonomy.⁷⁶ Numerous variables, such as time since deposition, the taphonomy of the environment in which the remains are recovered, and the storage conditions of the samples, could affect the degree of depurination or depyrimidination noticed within a particular sample.

DNA preservation or degradation over time is affected by i) biotic factors (bioturbation, soil bacteria, fauna, fungi, and other flora). ii) abiotic factors (radiation, physical damage, chemistry, pH, temperature, oxygen content, water content, soil type, porosity, weather processes, and humic substances). The formation of cross-links between DNA and other molecules will be broken down, altered, and catalyzed by these taphonomic variables.⁷⁷

In addition to producing AP sites, oxidative damage can cause mutagenic effects, leading to base damage or DNA lesions in the DNA backbone. This can lead to a fragile strand that may break, thus contributing to the fragmentation process. Reactive oxygen species, particularly hydroxyl radicals and hydrogen peroxide, are the main offenders behind oxidative damage.⁷² Although oxidative attack can occur at the double bonds of both purines and pyrimidines, resulting in ring fragmentation, the most common outcome of DNA oxidation is the production of 8-hydroxyguanine from guanine, which preferentially base pairs with adenine instead of cytosine, causing the G/C bond to be transversed to a T/A bond during the subsequent amplification.^{78,79} The attack can also occur within the deoxyribose molecule, breaking the sugar ring.⁷⁹ Amplification cannot proceed beyond these locations when base or sugar ring fragmentation occurs because DNA polymerase is blocked.⁷² Apart from 8-hydroxyguanine, several other oxidative damage products are produced, which can facilitate base mispairing during amplification.⁷⁹

Base substitutions can be categorized into two types: transitions and transversions [Table 1].⁸⁰

As shown in Table 1, C↔T or A↔G substitutions are predominantly caused by Hydrolytic deamination, a characteristic molecular damage pattern in ancient DNA.

Because DNA strands are complementary, a lesion on one strand directly results in a substitution on that strand; however, the substitution is also apparent on the complementary DNA strand following

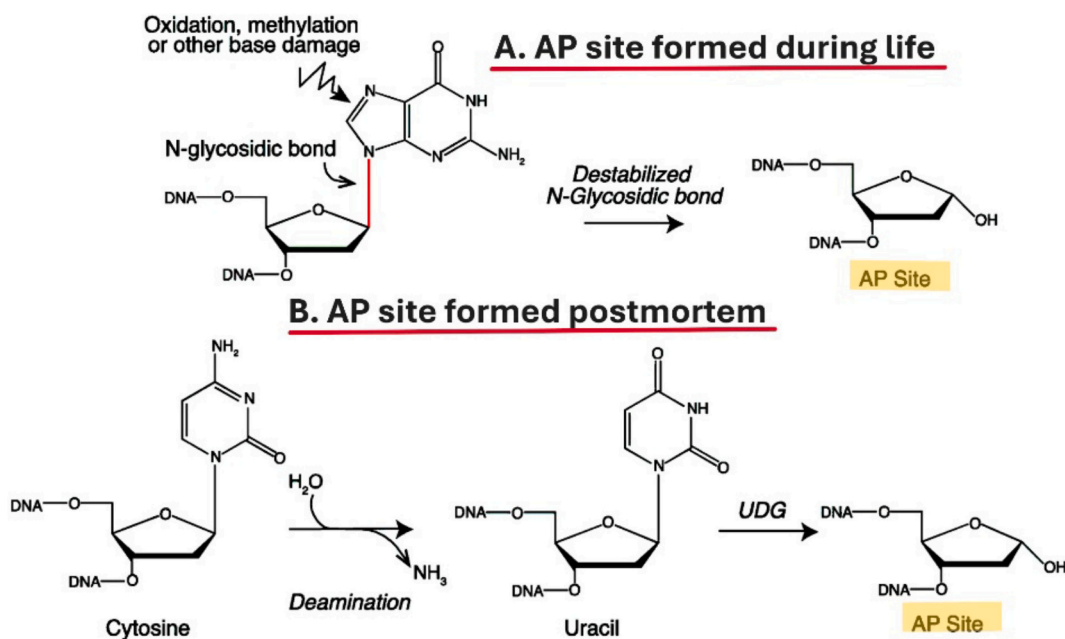


Fig. 3. Mechanisms of AP site formation. A. Destabilization of N-glycosidic bond after base damage generates AP site. B. Spontaneous or catalyzed deamination of cytosine generates uracil. Subsequent action of UDG produces an AP site.

Table 1
Types of base substitutions and the principal damaging agents causing the substitutions.

Transition	Transversion
It is the substitution of a purine for another purine base or a pyrimidine for another pyrimidine.	It is the substitution of a purine for a pyrimidine or a pyrimidine for a purine.
There is one possible transition. ($C \leftrightarrow T$ or $A \leftrightarrow G$).	There are two possible transversions. 1. Type 1 ($A \rightarrow C$, $T \rightarrow G$, $A \leftrightarrow T$) 2. Type 2 ($C \rightarrow A$, $G \rightarrow T$, $C \leftrightarrow G$)
In living organisms, it is less likely to cause an amino acid sequence change. So, this may remain a silent mutation.	In the living organism, it is more likely to cause amino acid sequence changes that affect the subsequently formed protein.
This type of damage is predominantly caused by Hydrolytic deamination.	This type of damage is predominantly caused by oxidative damage.
The most common type of point mutation.	This is less common than transition.

amplification. For example, depending on which DNA strand the substitution originated from, a type 2 transition can be considered either a $G \rightarrow A$ or a $C \rightarrow T$ substitution. According to reports, these type 2 transitions are the most prevalent type of postmortem damage.⁸⁰ Previous research has provided a detailed description of the different kinds of miscoding lesions found in ancient or degraded DNA (aDNA) templates. Regardless of the temperature and moisture content of the depositional environment from which the remains were recovered, a common theme across these studies is that the primary miscoding lesion during analysis is hydrolytic deamination leading to $C \rightarrow T$ transitions.^{2,80} According to certain studies using uracil-N-glycosylase (UNG), the deamination of cytosine is the only process that results in the $C \rightarrow T$ miscoding lesion.⁸¹ Although deamination is the primary candidate for this lesion, oxidative damage remains a plausible alternative, potentially generating the thymine analogues 5-hydroxycytosine and 5-hydroxyuracil.

This is especially true if the sample comes from an environment where hydrolytic activity is almost absent.⁷⁹ Nonetheless, studies of ancient hair samples have revealed that oxidative damage accounts for a larger proportion of total damage (Gilbert et al., 2006), which may be due to lower water content in these samples. This shift in oxidative damage-induced misincorporations, in turn, lowers the degree of hydrolytic damage-induced misincorporations.

In addition to oxidative damage, hydrolytic damage can result in base modification via deamination of the nitrogenous base. The pyrimidines, cytosine (C), and 5-methylcytosine are the main targets of this kind of damage. These two nucleotides can be modified to yield either uracil (U) or thymine (T). When the DNA molecule is amplified, these chemical products will subsequently cause a $C \rightarrow T$ transition.^{2,75,81,82} Deamination can also occur in purine nucleotides, adenine [A] and guanine [G]; however, this occurs to a far lesser extent. Adenine deamination generates Hypoxanthine, which preferentially base pairs with cytosine, resulting in a transition mutation from A to G. However, Guanine deamination results in Xanthine, which subsequently base pairs with cytosine.⁷⁵ In this case, conventional base pairing resumes in the subsequent amplification step, so there is no apparent mutation in the sequence.

For a comprehensive list of all 12 potential base substitutions, see Table 1, where base substitutions can be divided into two “types” for both transitions and transversions.

Analysis shows that modern DNA and potential contaminant sequences exhibit a significantly lower rate of nucleotide misincorporation compared to ancient samples. In other words, errors in nucleotide incorporation are much more prevalent in ancient DNA, while modern and contaminant sequences tend to display far greater fidelity. On the other hand, 85% of the detected nucleotide misincorporations in ancient sequences are type 2 transitions.⁸³ The primary cause of nucleotide misincorporations in ancient DNA is cytosine deamination.

Additionally, the apparent presence of G-to-A substitutions in sequencing data doesn't directly reflect a damage signal but is introduced during the “fill-in” step of double-stranded library protocols (e.g., Meyer & Kircher, 2010). During this step, damaged overhangs are filled before adapter ligation, transforming terminal $C \rightarrow T$ lesions on one strand into apparent $G \rightarrow A$ on the complementary strand in the sequencing output. In contrast, library preparation methods that avoid this fill-in step (e.g., single-stranded library protocols⁸⁴) preserve the authentic ancient DNA damage signature, which is characterized by $C \rightarrow T$ substitutions due to cytosine deamination at fragment termini. Thus, the observed substitution patterns strongly rely on the library preparation strategy rather than only on intrinsic DNA damage.

The changes from G to A and from C to T are distributed differently and unevenly throughout DNA molecules. It was suggested that single-stranded ends of molecules have a higher prevalence of cytosine deamination than the middle of molecules.²

3. Implications of Genetic Drift and Selection on the Mitochondrial Genome

Genetic drift and selection-driven changes in haplotype frequencies significantly affect the mtDNA molecule. It can be explained by a variety of factors, including uniparental transmission (no recombination) and the mitochondrial bottleneck process, in which mtDNA replicates from a small number of mtDNA templates, which magnifies the dilution of minor parental alleles in early development.^{19,85,86} A process known as Muller's Ratchet occurs when genetic drift triggers an irreversible accumulation of deleterious mutations in an asexual population's genome.^{19,87} As a result, heteroplasmic drift may reduce the effectiveness of mtDNA mutations as lineage markers.

There is a distinct pattern of negative and positive selection across mutation types and heteroplasmy levels. By measuring the rates of non-synonymous to synonymous mutations and comparing their normalized ratio (dN/dS), it was found that there is strong negative selection (dN/dS close to 0) on nonsense mutations at all ages and heteroplasmy levels, confirming the rapid elimination of protein-truncating mtDNA variants. However, missense mutations may be positively selected (dN/dS > 1) at the organelle level but negatively selected (dN/dS < 1) at the cellular level, resulting in antagonistic selection landscapes.⁸⁸ So, the dN/dS ratio should be interpreted carefully when assessing positive and negative selections because these values may not accurately reflect fitness effects when conventional population-genetic assumptions are not applicable. This is particularly relevant to mitochondrial genomes.⁸⁹

Genetic studies would be affected by the conflict regarding the irregular occurrence of parental leakage and the maternal inheritance of mtDNA, leading to misinterpretations of mtDNA-disease risks or associated traits.¹⁹

4. Retrieval and authentication of ancient DNA

4.1. Challenges in DNA preservation and retrieval

Ancient DNA is naturally fragmented. Typically, short fragments of ancient DNA range in size from 10 to 150 bp.⁹⁰ However, the possibility of incorrect mis-mappings of endogenous reads and non-endogenous (exogenous or environmental) contamination increases when ultra-short reads (≤ 35 bp) are included.⁹¹ Larger fragments, however, might be present in “younger” and better-preserved samples.⁹²

Being highly abundant per cell and maternally inherited, with little or no recombination, mtDNA is believed to be better preserved in ancient remains.^{21–23} Thus, it can be used as a molecular marker to study genetic flow & phylogeny. However, mtDNA is much more susceptible to damage due to the lack of protective proteins, such as histones, and its low-efficiency repair mechanism.^{93–95}

In the early Ancient DNA studies, the initial target was to sequence the hypervariable regions of mtDNA.⁹⁶ To determine potential maternal

relationships, a limited number of markers were used to analyze SNPs within the mtDNA coding region and hypervariable regions I and II (HVR—I, HVR-II), which typically define unique profiles and haplogroups.

Custom primers were used to amplify mtDNA markers, and restriction enzymes and/or Sanger sequencing were used to detect coding-region SNPs in highly degraded samples.⁹⁷

Authentication of the results is another issue affecting the PCR approach. Exogenous human contamination can introduce a shared contaminant across multiple samples, leading to false-positive matches and erroneous kinship attributions.

Precautions are typically taken to control potential contamination. Over the past twenty years, Sanger-type sequencing (STS), especially of the control region (CR), has been the accepted method for analyzing mitochondrial DNA (mtDNA) in forensic cases involving missing persons and human identification.^{98,99} However, NGS can overcome the primary limitations of STR processing by enabling whole mitochondrial genome (mtGenome) sequencing alongside superior sensitivity and throughput.

Although the control region (CR) usually receives most of the attention in forensic mtDNA analysis, using the entire mitochondrial genome (mtGenome) can yield higher discrimination.⁹⁹ Even though STS processing works well, it is a laborious, expensive, and time-consuming procedure, especially when analyzing the entire mtGenome.⁹⁹

To overcome many of the drawbacks associated with PCR-based methodologies, NGS techniques have been used recently to analyze aDNA.¹⁴ The primer-independent approach enables recovery of very short molecules and the reconstruction of target sequences from samples that were too damaged or highly fragmented for PCR analysis. With the use of high-throughput sequencing and enrichment techniques, data can be obtained even from minimal DNA.¹⁰⁰ Methods have been developed to enhance the recovery and processing of short, damaged DNA fragments from ancient remains. For instance, ancient DNA-optimized extraction protocols involve silica spin columns,¹⁰¹ magnetic beads,¹⁰² and in-solution silica binding¹⁰³ to maintain short DNA molecules.¹⁰⁴

In contrast, several commercial capture kits are now available for targeting mtDNA using either DNA or RNA baits that tile across the mitochondrial genome, ensuring comprehensive coverage. These include Twist Mitochondrial Panel (Twist Bioscience, USA) and myBaits® Mito hybridization capture kit (Arbor Biosciences, Ann Arbor, MI, USA). These panels enable researchers to enrich the Mt. genome, perform sequencing, and analyze genetic variation, including heteroplasmy-associated variants and disease-related variants. In addition to species identification, phylogenetics, barcoding, population genetics, and other popular applications.

Another challenge is the presence of nuclear sequences of mitochondrial origin (NumtS), which can affect the analysis of mtDNA (den Tex et al., 2010; Samaniego Castruita et al., 2015; Diroma, Maria Angela et al., 2021). These sequences that have been incorporated into the nuclear genome are non-functional and non-expressive. They present a problem for mtDNA amplification and subsequent analysis as they are not currently filtered out by many mtDNA analysis tools, such as mtDNA-server.¹⁰⁵ So, researchers should be aware of this potential confounding factor. On the other hand, computational techniques for NumtS identification in aDNA samples have been developed, enabling NumtS identification in aDNA sequencing.¹⁰⁶ A standard procedure based on utilizing only reads mapping uniquely to a genomic reference containing the nuclear and mitochondrial genomes together proved feasible for humans and was subsequently advised in aDNA research as a means of excluding putative NumtS.^{107–110}

The accurate reconstruction of mtDNA sequences from ancient samples remains challenging due to damage, fragmentation, and/or aDNA contamination, which may result in false-positive base calls at crucial positions, leading to incorrect haplogroup designations and biased downstream analysis. Misalignments & incorrect base calling may also result from the presence of poly-C tracts in the hypervariable

regions I & II when length heteroplasmy occurs.^{111,112}

4.2. Authentication of ancient DNA

The DNA fragments retrieved from ancient samples presented characteristics of authentic aDNA, such as an increase in nucleotide misincorporation at the fragment termini and short fragment length (< 100 bp). Notably, the frequency of guanine-to-adenine substitutions at the 3' termini and cytosine-to-thymine substitutions at the 5' termini was increased, indicating postmortem DNA damage. These damage patterns were assessed by mapDamage or complementary tools, which confirmed typical deamination signals.^{113,114}

A powerful tool for verifying the accuracy of whole-genome sequencing data is the mitochondrial genome. Although mitochondrial haplogroups are not unique to individuals, we can confirm the identity of a sample by comparing a complete mtDNA sequence to previous sequencing results or the expected haplotype and detect potential contamination. When the same mitochondrial DNA haplogroup is replicated across several independent sequencing experiments, it's a strong indication that the sample has been appropriately recognized, with no mislabeling or mix-up.¹¹⁵ This step is crucial for ensuring the accuracy of the results. In contrast, the presence of exogenous DNA may be indicated by the presence of uncommon or multiple mitochondrial DNA haplogroups within a single sample. Therefore, mtDNA can also be used to detect contamination. To distinguish authentic endogenous DNA from modern contaminants, researchers search for signs of DNA damage, check sequence consistency, and analyze mitochondrial DNA haplogroups. They often use specialized computational tools, such as Schmutzi¹¹⁶ to detect contamination in both whole-genome sequencing and mitochondrial data. Furthermore, mitochondrial DNA is essential for quality control. Scientists assess the depth of sequencing, the consistency of results, the completeness of the genome, and how well it matches established haplogroups for verifying the accuracy of sequencing and bioinformatics workflows. Collectively, mitochondrial DNA is an essential tool for maintaining the accuracy and reliability of whole-genome sequencing (WGS) studies.

5. Analytical tools & programs

Using the NGS technique made it possible to significantly increase the number of loci and individuals successfully typed, leading to the production of kinship estimates with higher resolution and a more accurate reconstruction of historical societies.¹¹⁷ Specialized bioinformatics pipelines were created to enhance the analysis of ancient molecules.¹¹⁸ Additionally, specific tools were created for the handling and analysis of molecule degradation. As an example, the damageProfiler,¹¹⁹ mapDamage,¹²⁰ and mapDamage2.0 (RRID:SCR_001240)¹²¹ packages enable the description of length patterns and misincorporation. PMD tool facilitates the decontamination of ancient genomes by calculating damage patterns and identifying degraded sequences. ContamMix, ContamLD,¹²² AuthenticT,¹²³ ANGSD,¹²⁴ and Schmutzi¹¹⁶ tools are used for estimating the contamination within human mtDNA. Schmutzi and ANGSD tools were developed to recognize and reduce the effects of contamination and damage on ancient DNA fragments and, if feasible, reconstruct high-quality ancient mitogenomes from publicly accessible raw genomic data. Schmutzi¹¹⁶ is a widely used tool for identifying the mitochondrial consensus sequence, detecting deamination patterns (endogenous aDNA marker), and estimating exogenous mitochondrial contamination. However, the ANGSD¹²⁴ tool could reconstruct the mitogenome of a very low-quality sample that couldn't be called by Schmutzi.¹²⁵

Mitochondrial Haplogroup classification could be performed by algorithms such as HaploGrep. HaploGrep provides an all-in-one solution for mtDNA profile quality assessment (ranging from 0 to 1) across population genetics, clinical genetics, and forensics. HaploGrep leverages the most up-to-date version of Phylotree to analyze sequences

generated by both the Schmutzi and ANGSD pipelines. Access is straightforward and free via <http://haplogrep.uibk.ac.at>.³⁸ The fully updated Haplogrep 3 release⁴⁰ introduces several advanced features, including enhanced quality control for VCF files, reducing errors caused by poor-quality variant calls. This framework enables the rapid detection of artificial or rare mutations, as well as recombinant artifacts and missing variants. Several bioinformatics tools have been developed and applied for mitochondrial DNA analysis, as summarized in [Table 2].

Next-generation sequencing technology evolves so quickly that bioinformatics pipelines need to be updated frequently. The main steps of the workflows are well known. However, new tools and software are continuously being developed. Tools and software updating will have significant impacts on data analysis, altering what researchers have found. Choosing the appropriate tool depends on how much detail you need while considering available time, funds, and computing resources.

6. Ethical considerations in ancient human remains research

Ethical considerations in ancient human remains are the principles used to protect individuals, communities, researchers, and scientific knowledge. Such considerations should be applied across all stages, from initial recovery to data analysis to retain their cultural and archaeological context. While ancient human remains serve as a scientific resource for biological and archaeological information, they represent the material remains of individuals and may possess spiritual, or cultural importance for descendant communities.

Consequently, an ethical framework for the scientific study of ancient human remains has been proposed, assessing the relationship between ancient human remains and researchers.¹²⁶ It keeps a balance between posthumous rights and the advancement of scientific understanding, the level of invasiveness and appropriateness of sampling and processing, as well as the relevant transport and storage, and the appropriate management of data generation, interpretation, and publication. Such considerations are highly relevant to ancient DNA studies, in which analysis may require destructive sampling of skeletal remains or teeth, highlighting the importance of preserving the remaining specimens while limiting sample removal, and ensuring proper handling and storage throughout the research process.

Consistent with these considerations, Current ethical frameworks have recognized the importance of clarity, proper handling of human remains, and active collaboration with archaeologists, curators, and communities with cultural or ancestral connections to the remains. Such considerations are highlighting the importance of recognizing ancient human remains as representing individuals rather than treating them solely as biological research resources.^{127–131}

Before undertaking sampling, the physical structure and condition of the remains should be documented, and all collected samples should be appropriately labeled, securely stored, and systematically tracked. Destructive sampling in biomolecular investigations should be scientifically supported and should involve the minimum amount of material necessary to achieve the study objectives. Preserving DNA extracts and sequencing libraries, when allowed, may help minimize the need for repeated destructive sampling in future studies.

7. Conclusion

This review addresses the role of mitochondrial DNA in the analysis of ancient samples. The major strength of utilizing mtDNA over nuclear DNA (nuDNA) is its higher retrievability potential with no recombination, due to its high copy number in the cell. Despite this, there are certain restrictions and drawbacks to limiting the study of human remains to mtDNA analysis only.

Advances in laboratory protocols and sequencing technologies have made it possible to more effectively retrieve mtDNA sequences from skeletal remains found in archaeological sites with improved data quality and authenticity, even from highly degraded samples.

Table 2
The bioinformatic tools used for mtDNA analysis.

Bioinformatic tool	Advantages	Limitations
<i>Burrows-Wheeler Alignment tool (BWA):</i> an alignment package that efficiently aligns short sequencing reads against a large reference sequence.	- It is a Fast and Effective tool, allowing accurate alignment of short sequencing reads. - It supports both color space reads and base space reads. - BAM files generated by BWA are compatible with most downstream tools. - The BWA aln algorithm performs particularly well with highly fragmented ancient DNA and short Illumina reads. - It supports mismatch tolerance and gapped alignment.	- It aligns reads to a linear reference genome which can reduce variant detection and coverage near the artificial breakpoint. This can be solved by using CircularMapper with BWA. - For Ancient DNA studies certain parameters need to be optimized to improve the mapping of damaged reads. - It is less effective for detecting complex mitochondrial structural variants, large deletions, and rearrangements. - It doesn't perform downstream analysis to detect heteroplasmy, estimate contamination, and generate consensus sequence.
<i>PhyloTree:</i> is a powerful JavaScript library designed for visualizing phylogenetic trees which is often used by researchers to classify mtDNA haplogroups.	- It provides an updated and extensive way to classify mtDNA haplogroups which are essential for tracing maternal ancestry & population Genetics. - It is updated regularly from newly sequenced mitogenome sequences. - It includes numerous interactive features. - Researchers can find and use it through (http://www.phylotree.org).³⁷ - It serves as a tool to educate others. - It is Capable of handling multiple selection categories for comparative analysis.	- It is a java-based tool which may be a barrier for some researchers. - It depends on external databases and studies that need updates often to maintain accuracy. - Loading a fully annotated mtDNA tree may slow down the browser performance.
<i>MITOMAP:</i> is a comprehensive database of mtDNA variations and their association with disease and evolution.	- It combines many details about human mtDNA, including sequence data, pathogenic mutations, population variations, and gene-gene interactions. - Easy Access. - Frequent Updates. - It has an integration with MITOMASTER, an analytical system that uses information from MITOMAP for the analysis of new information submitted by researchers.^{42,66} - It serves as a tool to educate others for a better understanding of mitochondrial genomics. - It is widely used in population genetics	- It is focused on human mtDNA. - Some populations are underrepresented, which limits global applicability. - There is a need for standardization in Data interpretation. - It is difficult to predict the pathogenicity of novel mutations.

(continued on next page)

Table 2 (continued)

Bioinformatic tool	Advantages	Limitations
<i>HaploGrep</i> : a bioinformatic tool designed to classify mtDNA haplogroups. ^{38–40}	- and forensic investigations. - It is a Fast and Effective tool as it utilizes Phylotree database to provide exact haplogroup classifications. <i>HaploGrep 3</i> was enhanced to support the partial FASTA sequence analysis. - It is user-friendly - It works with multiple input file formats. - It adds checks to reduce misclassification. <i>HaploGrep 3</i> has an integrated variant annotation for immediate interpretation of variants.	- It relies on the Phylotree database. The database needs frequent updates to stay useful. - Since it focuses on mtDNA analysis, it may not perform well with other DNA types. - It is difficult to handle complex heteroplasmy. - Beginners in bioinformatics tools might take some time to add new phylogenetic trees. - Its Performance depends on complete variant information.
<i>damage Profiler</i> : A tool built in Java made to identify damage patterns in ancient DNA. ¹¹⁹	- It is Fast, interactive, and easy to use java-based program for coherent calculation of damage patterns in mapped NGS reads from various mapping tools. - It can be applied to single- or multi-reference mapping files, which is useful in metagenomic studies. - It is simple to integrate into existing analysis pipelines, as in nf-core/eager. - It can work with various file formats such as BAM, SAM, and CRAM. - It generates detailed outputs like damage plots, fragment length distributions, and base frequency tables - It allows customization of visualization options. - It is Available as an open-source code on GitHub.	- It Requires Java to function. - It does not perform Bayesian estimation of damage parameters, which may be required for deeper statistical analysis. - There may be limitations in the graphical interface for advanced users. - Although it is focused on analyzing damage patterns in aDNA, it requires another tool to quantify modern contamination.
<i>mapDamage</i> : a tool for analyzing damage patterns in sequences generated from aDNA using next-generation sequencing platforms. ^{120,121}	- The tool is useful for distinguishing ancient DNA from modern contaminants based on cytosine deamination. - It provides useful summary tables and extracts information for authenticating aDNA based on DNA damage signatures. - It uses Bayesian statistical models to determine important characteristics of aDNA molecules. - It can work well with different DNA library protocols.	- Users need to install SAMtools and R. This setup may not work well for everyone. - Although it is focused on analyzing damage patterns in aDNA, it requires another tool to quantify modern contamination. - It offers few options for customizing output plots. - It can slow down processing of large datasets.

Table 2 (continued)

Bioinformatic tool	Advantages	Limitations
<i>PMDtools</i> is a software tool designed to detect postmortem damage patterns and hence decontaminate the ancient genomes.	- It allows quantitative comparisons of damage levels across different samples. - It can be accessed freely through GitHub. - It can distinguish ancient DNA from modern contamination. - It uses a likelihood framework incorporating damage patterns, base quality scores, and biological polymorphisms for the accurate identification of aDNA. - It allows users to customize their filtering options. - It can be integrated into existing pipelines and supports SAM-formatted input files. - free-to-use software - It uses Markov Chain Monte Carlo (MCMC) methods to estimate contamination levels with high accuracy. - free-to-use software. - It generates clear documentation, including plots for contamination estimates, that helps in the interpretation. - It can be integrated into existing pipelines and works with various sequencing data formats like BAM & FASTA files. - It provides accurate estimates of contamination rates with low standard errors, even with low levels of contamination (<10%). - It can effectively distinguish between authentic ancient DNA and contaminant DNA by measuring the breakdown of LD - Its contamination estimates are based on aDNA damage patterns. - It is an open-source tool. - It is applicable for both sexes, unlike ANGSD which is applicable for X-Chromosome only. - It relies on patterns of post-mortem damage to make an accurate estimate of contamination in single-stranded DNA (ssDNA) libraries.	- It relies on external programs like SAMtools to work with BAM files and uses R to create plots that might not be suitable for non-technical users. - It requires an input BAM file with MD tags, a specific feature in the output file of many aligners. - It specializes in detecting ancient DNA damage. - It doesn't quantify contamination levels.
<i>ContamMix</i> : is a tool used for estimating contamination in ancient mtDNA using a mixture-model approach.	- free-to-use software. - It uses Markov Chain Monte Carlo (MCMC) methods to estimate contamination levels with high accuracy. - free-to-use software. - It generates clear documentation, including plots for contamination estimates, that helps in the interpretation. - It can be integrated into existing pipelines and works with various sequencing data formats like BAM & FASTA files. - It provides accurate estimates of contamination rates with low standard errors, even with low levels of contamination (<10%). - It can effectively distinguish between authentic ancient DNA and contaminant DNA by measuring the breakdown of LD - Its contamination estimates are based on aDNA damage patterns. - It is an open-source tool. - It is applicable for both sexes, unlike ANGSD which is applicable for X-Chromosome only. - It relies on patterns of post-mortem damage to make an accurate estimate of contamination in single-stranded DNA (ssDNA) libraries.	- It might not be accurate with Low-Quality Consensus Sequences. - It can take time to learn as it is command-line based. - The method assumes that contamination is less than 50%, which might not always be the case. - It may ignore sites with ambiguous characters. - The tool focuses on estimating mtDNA contamination and not the nuclear DNA contamination. - It requires Python 3 and R, as well as other tools like PMDtools and SAMtools, which might not be ideal for all users - It can be time-consuming as it requires many pre-processing steps. - Its accuracy decreases with low coverage (<0.5×). - It specializes in analyzing ancient DNA. - It may require significant computational resources, especially for low-coverage whole-genome shotgun sequences. - It is hard to learn.
<i>ContamLD</i> : is a tool designed to estimate nuclear aDNA contamination by measuring the breakdown of linkage disequilibrium (LD) due to the introduction of contaminant DNA. ¹²²	- It provides accurate estimates of contamination rates with low standard errors, even with low levels of contamination (<10%). - It can effectively distinguish between authentic ancient DNA and contaminant DNA by measuring the breakdown of LD - Its contamination estimates are based on aDNA damage patterns. - It is an open-source tool. - It is applicable for both sexes, unlike ANGSD which is applicable for X-Chromosome only. - It relies on patterns of post-mortem damage to make an accurate estimate of contamination in single-stranded DNA (ssDNA) libraries.	- It requires Python 3 and R, as well as other tools like PMDtools and SAMtools, which might not be ideal for all users - It can be time-consuming as it requires many pre-processing steps. - Its accuracy decreases with low coverage (<0.5×). - It specializes in analyzing ancient DNA. - It may require significant computational resources, especially for low-coverage whole-genome shotgun sequences. - It is hard to learn.
<i>AuthentiCT</i> : a command-line tool designed to estimate the proportion of present-day DNA contamination in ancient DNA datasets generated from single-	It relies on patterns of post-mortem damage to make an accurate estimate of contamination in single-stranded DNA (ssDNA) libraries.	- It needs Python 3.6 and several dependencies. - It specializes in libraries of ssDNA and it is not suitable with UDG-treated samples or double-stranded DNA (dsDNA).

(continued on next page)

Table 2 (continued)

Bioinformatic tool	Advantages	Limitations
stranded DNA libraries. ¹²³	- It doesn't require high data and can measure contamination with just 10,000 mapped sequences which is useful for poorly preserved specimens. - It is a free-to-use tool - It is an easy-to-use tool.	It requires SAM file with an MD field as an input file.
ANGSD: is software developed for analyzing NGS data. ¹²⁴	- It can handle large datasets efficiently. - It can handle different input formats, such as BAM and CRAM. - It is ideal for low-depth sequencing data as it works with genotype likelihoods instead of called genotypes. - It can perform population genetic analyses, association mapping, and summary statistics - It allows you to process multiple tasks at once and get results quickly and efficiently. - It is a Free and open-source software.	- Results can be changed based on quality filtering. - It is not recommended for high-depth data as it may be less accurate. - It doesn't support indels in any analysis. - It isn't easy to learn as it is a Command-line interface. - It has mostly been tested and proven to work on GNU/Linux systems, but it can also compile on OSX with some limitations.
Schmutzi: an iterative approach to estimate present-day human contamination in aDNA datasets, with a focus on reconstructing endogenous mtDNA. ¹¹⁶	- It uses patterns of post-mortem damage for accurate contamination estimation. - It can rebuild the endogenous mitochondrial genome, accounting for contamination and deamination. - It uses a Bayesian maximum a posteriori (MAP) estimate, which gives strong and reliable results. - Free and Open Source. - It generates detailed Documentation.	- It relies on External Tools - It may underestimate contamination levels if there are multiple present-day human contaminants. - Accuracy might not be achieved in low-coverage samples. - It is designed for mtDNA analysis and might not work well for other kinds of DNA analysis. - It may require significant computational resources, especially for large datasets. - It runs better with Linux and macOS.

In paleogenomic studies, the standard nowadays for mitochondrial DNA analysis is to recover the whole mitochondrial genomes either through deep sequencing for well-preserved samples or through targeted capture enrichment for more degraded samples.

Although contamination and postmortem damage remain as existing challenges, advances in methodology and integrative approaches are expected to improve the resolution and reliability of mtDNA-based inferences.

CRedit authorship contribution statement

Dina W. Ibrahim: Conceptualization, Writing – original draft, Investigation, Visualization, Writing – review & editing. **Haidan Mostafa El-Shorbagy:** Writing – review & editing. **Ismail A. Abdelhamid:** Writing – review & editing. **Yehia Z. Gad:** Supervision, Conceptualization, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The study was supported by the National Museum of Egyptian Civilization.

Consent to participate

Not applicable.

Consent for publication

Not applicable.

Ethical approval

Not applicable.

Funding

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

References

- Higuchi R, Bowman B, Freiberger M, Ryder OA, Wilson AC. DNA sequences from the quagga, an extinct member of the horse family. *Nature*. 1984;312(5991):282–284. <https://doi.org/10.1038/312282a0>.
- Briggs AW, Stenzel U, Johnson PLF, et al. Patterns of damage in genomic DNA sequences from a Neandertal. *Proc Natl Acad Sci USA*. 2007;104(37):14616–14621. <https://doi.org/10.1073/pnas.0704665104>.
- Vernesi C, Caramelli D, Dupanloup I, Bertorelle G, Lari M, Cappellini E, Moggi-Cecchi J, Chiarelli B, Castri L, Casoli A, Mallegni F, Lalueza-Fox C, & Barbujani G. (2004). The Etruscans: a population-genetic study. *Am J Hum Genet*, 74(4), 694–704. <https://doi.org/https://doi.org/10.1086/383284>. Published online 2004. doi:10.1086/383284.
- Haak W, Forster P, Bramanti B, et al. Ancient DNA from the first European farmers in 7500-year-old Neolithic sites. *Science*. 2005;310(5750):1016–1018. <https://doi.org/10.1126/science.1118725>.
- Sampietro ML, Caramelli D, Lao O, et al. The genetics of the pre-Roman Iberian Peninsula: a mtDNA study of ancient Iberians. *Ann Hum Genet*. 2005;69(Pt 5):535–548. <https://doi.org/10.1111/j.1529-8817.2005.00194.x>.
- Tommasi A, Boscolo Agostini R, Villani G, et al. Fifteen millennia of human mitogenome evolution in Sicily. *Sci Adv*. 2025;11(46), eady1674. <https://doi.org/10.1126/sciadv.ady1674>.
- Mathieson I, Lazaridis I, Rohland N, et al. Genome-wide patterns of selection in 230 ancient Eurasians. *Nature*. 2015;528(7583):499–503. <https://doi.org/10.1038/nature16152>.
- Patterson N, Isakov M, Booth T, et al. Large-scale migration into Britain during the middle to late bronze age. *Nature*. 2022;601(7894):588–594. <https://doi.org/10.1038/s41586-021-04287-4>.
- Morez Jacobs A, Irish JD, Cooke A, et al. Whole-genome ancestry of an old kingdom Egyptian. *Nature*. 2025;644(8077):714–721. <https://doi.org/10.1038/s41586-025-09195-5>.
- Nielsen R, Akey JM, Jakobsson M, Pritchard JK, Tishkoff S, Willerslev E. Tracing the peopling of the world through genomics. *Nature*. 2017;541(7637):302–310. <https://doi.org/10.1038/nature21347>.
- Olalde I, Brace S, Allentoft ME, et al. The beaker phenomenon and the genomic transformation of Northwest Europe. *Nature*. 2018;555(7695):190–196. <https://doi.org/10.1038/nature25738>.
- Bergström A, McCarthy SA, Hui R, et al. Insights into human genetic variation and population history from 929 diverse genomes. *Science*. 2020;367(6484), eaay5012. <https://doi.org/10.1126/science.aay5012>.
- Uricoechea Patiño D, Collins A, García OJR, et al. High mitochondrial haplotype diversity found in three pre-Hispanic groups from Colombia. *Genes (Basel)*. 2023;14(10), 1853. <https://doi.org/10.3390/genes14101853>.
- Vai S, Amorim CEG, Lari M, Caramelli D. Kinship determination in archeological contexts through DNA analysis. *Front Ecol Evol*. 2020;8, 83. <https://doi.org/10.3389/fevo.2020.00083>.
- Miller FJ. Precise determination of mitochondrial DNA copy number in human skeletal and cardiac muscle by a PCR-based assay: lack of change of copy number

- with age. *Nucleic Acids Res.* 2003;31(11):61e–661. <https://doi.org/10.1093/nar/gng060>.
16. Copeland WC. Defects of Mitochondrial DNA replication. *J Child Neurol.* 2014;29(9):1216–1224. <https://doi.org/10.1177/0883073814537380>.
 17. Taanman JW. The mitochondrial genome: structure, transcription, translation and replication. *Biochimica et Biophysica Acta (BBA) - Bioenergetics.* 1999;1410(2):103–123. [https://doi.org/10.1016/S0005-2728\(98\)00161-3](https://doi.org/10.1016/S0005-2728(98)00161-3).
 18. Zardoya R. Recent advances in understanding mitochondrial genome diversity. *Fl000Res.* 2020;9:270. <https://doi.org/10.12688/fl000research.21490.1>.
 19. Ferreira T, Rodríguez S. Mitochondrial DNA: inherent complexities relevant to genetic analyses. *Genes.* 2024;15(5), 617. <https://doi.org/10.3390/genes15050617>.
 20. Habbane M, Montoya J, Rhouda T, Sbaoui Y, Radallah D, Emperador S. Human Mitochondrial DNA: particularities and diseases. *Biomedicines.* 2021;9(10), 1364. <https://doi.org/10.3390/biomedicines9101364>.
 21. Kivisild T. Maternal ancestry and population history from whole mitochondrial genomes. *Investig Genet.* 2015;6(1):3. <https://doi.org/10.1186/s13323-015-0022-2>.
 22. Diallo MY, Čížková M, Kulichová I, et al. Circum-Saharan prehistory through the lens of mtDNA diversity. *Genes.* 2022;13(3), 533. <https://doi.org/10.3390/genes13030533>.
 23. Morris AG, Salie T, Mitnik A, et al. Reconstructing ancient southern African mitochondrial genomes at Faraoskop. *Archaeol Anthropol Sci.* 2025;17(1), 11. <https://doi.org/10.1007/s12520-024-02130-1>.
 24. Pakendorf B, Stoneking M. MITOCHONDRIAL DNA AND HUMAN EVOLUTION. *Annu Rev Genom Hum Genet.* 2005;6(1):165–183. <https://doi.org/10.1146/annurev.genom.6.080604.162249>.
 25. Lankheet I, Chowdhury A, Tellgren-Roth C, et al. Revisiting the African mtDNA landscape: a continental update from complete mitochondrial genomes. *Evolutionary Biology* Preprint posted online April 7, 2025. doi:<https://doi.org/10.1101/2025.04.05.647361>.
 26. Brown WM, Prager EM, Wang A, Wilson AC. Mitochondrial DNA sequences of primates: tempo and mode of evolution. *J Mol Evol.* 1982;18(4):225–239. <https://doi.org/10.1007/BF01734101>.
 27. Wallace DC, Ye J, Neckelmann SN, Singh G, Webster KA, Greenberg BD. Sequence analysis of cDNAs for the human and bovine ATP synthase β subunit: mitochondrial DNA genes sustain seventeen times more mutations. *Curr Genet.* 1987;12(2):81–90. <https://doi.org/10.1007/BF00434661>.
 28. Wallace DC, Chalkia D. Mitochondrial DNA genetics and the heteroplasmy conundrum in evolution and disease. *Cold Spring Harb Perspect Biol.* 2013;5(11), a021220. <https://doi.org/10.1101/cshperspect.a021220>.
 29. Hertweck KL, Dasgupta S. The landscape of mtDNA modifications in cancer: a tale of two cities. *Front Oncol.* 2017;7, 262. <https://doi.org/10.3389/fonc.2017.00262>.
 30. Neiman M, Taylor DR. The causes of mutation accumulation in mitochondrial genomes. *Proc Biol Sci.* 2009;276(1660):1201–1209. <https://doi.org/10.1098/rspb.2008.1758>.
 31. Fan W, Waymire KG, Narula N, et al. A mouse model of mitochondrial disease reveals germline selection against severe mtDNA mutations. *Science.* 2008;319(5865):958–962. <https://doi.org/10.1126/science.1147786>.
 32. Stewart JB, Freyer C, Elson JL, et al. Strong purifying selection in transmission of mammalian mitochondrial DNA. *Hurst LD, ed PLoS Biol.* 2008;6(1), e10. <https://doi.org/10.1371/journal.pbio.0060010>.
 33. Wai T, Teoli D, Shoubridge EA. The mitochondrial DNA genetic bottleneck results from replication of a subpopulation of genomes. *Nat Genet.* 2008;40(12):1484–1488. <https://doi.org/10.1038/ng.258>.
 34. Árnadóttir ER, Moore KHS, Guðmundsdóttir VB, et al. The rate and nature of mitochondrial DNA mutations in human pedigrees. *Cell.* 2024;187(15):3904–3918. e8. <https://doi.org/10.1016/j.cell.2024.05.022>.
 35. Wallace DC. Bioenergetics in human evolution and disease: implications for the origins of biological complexity and the missing genetic variation of common diseases. *Philos Trans R Soc B.* 2013;368(1622), 20120267. <https://doi.org/10.1098/rstb.2012.0267>.
 36. Wallace DC. Mitochondrial DNA Variation in Human Radiation and Disease. *Cell.* 2015;163(1):33–38. <https://doi.org/10.1016/j.cell.2015.08.067>.
 37. Van Oven M, Kayser M. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum Mutat.* 2009;30(2):E386–E394. <https://doi.org/10.1002/humu.20921>.
 38. Kloss-Brandstätter A, Pacher D, Schönherr S, et al. HaploGrep: a fast and reliable algorithm for automatic classification of mitochondrial DNA haplogroups. *Hum Mutat.* 2011;32(1):25–32. <https://doi.org/10.1002/humu.21382>.
 39. Weissensteiner H, Pacher D, Kloss-Brandstätter A, et al. HaploGrep 2: mitochondrial haplogroup classification in the era of high-throughput sequencing. *Nucleic Acids Res.* 2016;44(W1):W58–W63. <https://doi.org/10.1093/nar/gkw233>.
 40. Schönherr S, Weissensteiner H, Kronenberg F, Forer L. Haplogrep 3 - an interactive haplogroup classification and analysis platform. *Nucleic Acids Res.* 2023;51(W1):W263–W268. <https://doi.org/10.1093/nar/gkad284>.
 41. Vianello D, Sevin F, Castellani G, Lomartire L, Capri M, Franceschi C. HAPLOFIND: A new method for high-throughput mtDNA Haplogroup assignment. *Hum Mutat.* 2013;34(9):1189–1194. <https://doi.org/10.1002/humu.22356>.
 42. Brandon MC, Ruiz-Pesini E, Mishmar D, et al. MITOMASTER: a bioinformatics tool for the analysis of mitochondrial DNA sequences. *Hum Mutat.* 2009;30(1):1–6. <https://doi.org/10.1002/humu.20801>.
 43. Fan L, Yao YG. An update to MitoTool: using a new scoring system for faster mtDNA haplogroup determination. *Mitochondrion.* 2013;13(4):360–363. <https://doi.org/10.1016/j.mito.2013.04.011>.
 44. Ji F, Sharpley MS, Derbeneva O, et al. Mitochondrial DNA variant associated with Leber hereditary optic neuropathy and high-altitude Tibetans. *Proc Natl Acad Sci USA.* 2012;109(19):7391–7396. <https://doi.org/10.1073/pnas.1202484109>.
 45. Rishishwar L, Jordan IK. Implications of human evolution and admixture for mitochondrial replacement therapy. *BMC Genomics.* 2017;18(1):140. <https://doi.org/10.1186/s12864-017-3539-3>.
 46. Wallace DC. A Mitochondrial paradigm of metabolic and degenerative diseases, aging, and Cancer: A Dawn for evolutionary medicine. *Annu Rev Genet.* 2005;39(1):359–407. <https://doi.org/10.1146/annurev.genet.39.110304.095751>.
 47. Fuku N, Park KS, Yamada Y, et al. Mitochondrial Haplogroup N9a confers resistance against type 2 diabetes in Asians. *Am J Hum Genet.* 2007;80(3):407–415. <https://doi.org/10.1086/512202>.
 48. Chalkia D, Singh LN, Leipzig J, et al. Association between Mitochondrial DNA Haplogroup variation and autism Spectrum disorders. *JAMA Psychiatry.* 2017;74(11):1161. <https://doi.org/10.1001/jamapsychiatry.2017.2604>.
 49. Marchington DR, Macaulay V, Hartshorne GM, Barlow D, Poulton J. Evidence from human oocytes for a genetic bottleneck in an mtDNA disease. *Am J Hum Genet.* 1998;63(3):769–775. <https://doi.org/10.1086/302009>.
 50. Chinnery PF, Johnson MA, Wardell TM, et al. The epidemiology of pathogenic mitochondrial DNA mutations. *Ann Neurol.* 2000;48(2):188–193. [https://doi.org/10.1002/1531-8249\(200008\)48:2<188::aid-ana8>3.0.co;2-p](https://doi.org/10.1002/1531-8249(200008)48:2<188::aid-ana8>3.0.co;2-p).
 51. Meltan T, Nelson K. Forensic mitochondrial DNA analysis: two years of commercial casework experience in the United States. *Croat Med J.* 2001;42(3):298–303.
 52. Irwin JA, Saunier JL, Niederstätter H, et al. Investigation of Heteroplasmy in the human Mitochondrial DNA control region: A synthesis of observations from more than 5000 global population samples. *J Mol Evol.* 2009;68(5):516–527. <https://doi.org/10.1007/s00239-009-9227-4>.
 53. Gonçalves VF, Carvalho CMB, Bortolini MC, Bydlowski SP, Pena SDJ. The phylogeography of African Brazilians. *Hum Hered.* 2008;65(1):23–32. <https://doi.org/10.1159/000106059>.
 54. Wallace DC. A mitochondrial bioenergetic etiology of disease. *J Clin Invest.* 2013;123(4):1405–1412. <https://doi.org/10.1172/JCI61398>.
 55. Szczepanowska K, Trifunovic A. Different faces of mitochondrial DNA mutators. *Biochimica et Biophysica Acta (BBA) - Bioenergetics.* 2015;1847(11):1362–1372. <https://doi.org/10.1016/j.bbabi.2015.05.016>.
 56. Pikó L, Hougham AJ, Bulpitt KJ. Studies of sequence heterogeneity of mitochondrial DNA from rat and mouse tissues: evidence for an increased frequency of deletions/additions with aging. *Mech Ageing Dev.* 1988;43(3):279–293. [https://doi.org/10.1016/0047-6374\(88\)90037-1](https://doi.org/10.1016/0047-6374(88)90037-1).
 57. Wallace DC. Mitochondrial DNA mutations in disease and aging. *Environ Mol Mutagen.* 2010;51(5):440–450. <https://doi.org/10.1002/em.20586>.
 58. Ameer A, Stewart JB, Freyer C, et al. Ultra-deep sequencing of mouse mitochondrial DNA: mutational patterns and their origins. *Barsh GS, ed PLoS Genet.* 2011;7(3), e1002028. <https://doi.org/10.1371/journal.pgen.1002028>.
 59. Kennedy SR, Salk JJ, Schmitt MW, Loeb LA. Ultra-sensitive sequencing reveals an age-related increase in somatic mitochondrial mutations that are inconsistent with oxidative damage. *Van Houten B, ed PLoS Genet.* 2013;9(9), e1003794. <https://doi.org/10.1371/journal.pgen.1003794>.
 60. Itsara LS, Kennedy SR, Fox EJ, et al. Oxidative stress is not a major contributor to somatic mitochondrial DNA mutations. *Turnbull DM, ed PLoS Genet.* 2014;10(2): e1003974. doi:<https://doi.org/10.1371/journal.pgen.1003974>.
 61. Stewart JB, Chinnery PF. The dynamics of mitochondrial DNA heteroplasmy: implications for human health and disease. *Nat Rev Genet.* 2015;16(9):530–542. <https://doi.org/10.1038/nrg3966>.
 62. Sequeira A, Martin MV, Rollins B, et al. Mitochondrial mutations and polymorphisms in psychiatric disorders. *Front Gene.* 2012;3. <https://doi.org/10.3389/fgene.2012.00103>.
 63. Toncheva D, Serbezov D, Karachanak-Yankova S, Nesheva D. Ancient mitochondrial DNA pathogenic variants putatively associated with mitochondrial disease. *Caramelli D, ed PLoS One.* 2020;15(9), e0233666. <https://doi.org/10.1371/journal.pone.0233666>.
 64. Wallace DC, Singh G, Lott MT, et al. Mitochondrial DNA mutation associated with Leber's hereditary optic neuropathy. *Science.* 1988;242(4884):1427–1430. <https://doi.org/10.1126/science.3201231>.
 65. Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. GenBank. *Nucleic Acids Res.* 2016;44(D1):D67–D72. <https://doi.org/10.1093/nar/gkv1276>.
 66. Lott MT, Leipzig JN, Derbeneva O, et al. mtDNA variation and analysis using Mitomap and Mitomaster. *Curr Protoc Bioinformatics.* 2013;44(1). <https://doi.org/10.1002/0471250953.bio123s44>.
 67. Sonney S, Leipzig J, Lott MT, et al. Predicting the pathogenicity of novel variants in mitochondrial tRNA with MitoTIP. *Poisot T, ed PLoS Comput Biol.* 2017;13(12): e1005867. doi:<https://doi.org/10.1371/journal.pcbi.1005867>.
 68. Preste R, Vitale O, Clima R, Gasparre G, Attimonelli M. HmtVar: a new resource for human mitochondrial variations and pathogenicity data. *Nucleic Acids Res.* 2019;47(D1):D1202–D1210. <https://doi.org/10.1093/nar/gky1024>.
 69. Modi A, Nesheva D, Sarno S, et al. Ancient human mitochondrial genomes from bronze age Bulgaria: new insights into the genetic history of Thracians. *Sci Rep.* 2019;9(1):5412. <https://doi.org/10.1038/s41598-019-41945-0>.
 70. Landrum MJ, Lee JM, Riley GR, et al. ClinVar: public archive of relationships among sequence variation and human phenotype. *Nucl Acids Res.* 2014;42(D1): D980–D985. <https://doi.org/10.1093/nar/gkt1113>.
 71. Pääbo S. Ancient DNA: extraction, characterization, molecular cloning, and enzymatic amplification. *Proc Natl Acad Sci USA.* 1989;86(6):1939–1943. <https://doi.org/10.1073/pnas.86.6.1939>.

72. Pääbo S, Poinar H, Serre D, et al. Genetic analyses from ancient DNA. *Annu Rev Genet.* 2004;38(1):645–679. <https://doi.org/10.1146/annurev.genet.37.110801.143214>.
73. Barnes DE, Lindahl T. Repair and genetic consequences of endogenous DNA Base damage in mammalian cells. *Annu Rev Genet.* 2004;38(1):445–476. <https://doi.org/10.1146/annurev.genet.38.072902.092448>.
74. Poinar HN, Hofreiter M, Spaulding WG, et al. Molecular Coprospect: dung and diet of the extinct ground sloth *Nothotheriops shastensis*. *Science.* 1998;281(5375):402–406. <https://doi.org/10.1126/science.281.5375.402>.
75. Lindahl T. Instability and decay of the primary structure of DNA. *Nature.* 1993;362(6422):709–715. <https://doi.org/10.1038/362709a0>.
76. Blau S. Taphonomy: Definition. In: Smith C, ed. *Encyclopedia of Global Archaeology*. New York: Springer; 2014:7235. https://doi.org/10.1007/978-1-4419-0465-2_132.
77. Eriksen AMH, Nielsen TK, Matthiesen H, et al. Bone biodeterioration—the effect of marine and terrestrial postdepositional environments on early diagenesis and bone bacterial community. *Franzetti A, ed. PLoS ONE.* 2020;15(10), e0240512. <https://doi.org/10.1371/journal.pone.0240512>.
78. Shibutani S, Takeshita M, Grollman AP. Insertion of specific bases during DNA synthesis past the oxidation-damaged base 8-oxodG. *Nature.* 1991;349(6308):431–434. <https://doi.org/10.1038/349431a0>.
79. Dizdaroğlu M, Jaruga P, Birincioglu M, Rodriguez H. Free radical-induced damage to DNA: mechanisms and measurement 1,2 1 This article is part of a series of reviews on “Oxidative DNA Damage and Repair.” The full list of papers may be found on the homepage of the journal. 2Guest Editor: Miral Dizdaroğlu. *Free Radic Biol Med.* 2002;32(11):1102–1115. [https://doi.org/10.1016/S0891-5849\(02\)00826-2](https://doi.org/10.1016/S0891-5849(02)00826-2).
80. Hansen AJ, Willerslev E, Wiuf C, Mourier T, Arctander P. Statistical evidence for miscoding lesions in ancient DNA templates. *Mol Biol Evol.* 2001;18(2):262–265. <https://doi.org/10.1093/oxfordjournals.molbev.a003800>.
81. Gilbert MTP, Hansen AJ, Willerslev E, et al. Characterization of genetic miscoding lesions caused by postmortem damage. *Am J Hum Genet.* 2003;72(1):48–61. <https://doi.org/10.1086/345379>.
82. Brotherton P, Endicott P, Sanchez JJ, et al. 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 Res.* 2007;35(17):5717–5728. <https://doi.org/10.1093/nar/gkm588>.
83. Olivieri C, Ermini L, Rizzi E, et al. Characterization of nucleotide misincorporation patterns in the Iceman’s mitochondrial DNA. Gilbert MTP, ed. *PLoS ONE.* 2010;5(1), e8629. <https://doi.org/10.1371/journal.pone.0008629>.
84. Gansauge MT, Meyer M. Single-stranded DNA library preparation for the sequencing of ancient or damaged DNA. *Nat Protoc.* 2013;8(4):737–748. <https://doi.org/10.1038/nprot.2013.038>.
85. Marchington DR, Hartshorne GM, Barlow D, Poulton J. Homopolymeric tract heteroplasmy in mtDNA from tissues and single oocytes: support for a genetic bottleneck. *Am J Hum Genet.* 1997;60(2):408–416.
86. Chinnery PF, Thorburn DR, Samuels DC, et al. The inheritance of mitochondrial DNA heteroplasmy: random drift, selection or both? *Trends Genet.* 2000;16(11):500–505. [https://doi.org/10.1016/S0168-9525\(00\)02120-X](https://doi.org/10.1016/S0168-9525(00)02120-X).
87. Otten ABC, Smeets HJM. Evolutionary defined role of the mitochondrial DNA in fertility, disease and ageing. *Hum Reprod Update.* 2015;21(5):671–689. <https://doi.org/10.1093/humupd/dmv024>.
88. Campbell P, Chapman MS, Przybyla M, et al. Mitochondrial mutation, drift and selection during human development and ageing. Preprint posted online June 26. 2023. <https://doi.org/10.21203/rs.3.rs-3083262/v1>.
89. Williams MJ, Zapata L, Werner B, Barnes CP, Sottoriva A, Graham TA. Measuring the distribution of fitness effects in somatic evolution by combining clonal dynamics with dN/dS ratios. *eLife.* 2020;9, e48714. <https://doi.org/10.7554/eLife.48714>.
90. Sawyer S, Krause J, Guschanski K, Savolainen V, Pääbo S. Temporal patterns of nucleotide misincorporations and DNA fragmentation in ancient DNA. Lalueza-Fox C, ed. *PLoS One.* 2012;7(3), e34131. <https://doi.org/10.1371/journal.pone.0034131>.
91. De Filippo C, Meyer M, Prüfer K. Quantifying and reducing spurious alignments for the analysis of ultra-short ancient DNA sequences. *BMC Biol.* 2018;16(1):121. <https://doi.org/10.1186/s12915-018-0581-9>.
92. Hofreiter M, Serre D, Poinar HN, Kuch M, Pääbo S. Ancient DNA. *Nat Rev Genet.* 2001;2(5):353–359. <https://doi.org/10.1038/35072071>.
93. Kujoth GC, Hiona A, Pugh TD, et al. Mitochondrial DNA mutations, oxidative stress, and apoptosis in mammalian aging. *Science.* 2005;309(5733):481–484. <https://doi.org/10.1126/science.1112125>.
94. Ishikawa K, Takenaga K, Akimoto M, et al. ROS-generating mitochondrial DNA mutations can regulate tumor cell metastasis. *Science.* 2008;320(5876):661–664. <https://doi.org/10.1126/science.1156906>.
95. Nadalutti CA, Ayala-Peña S, Santos JH. Mitochondrial DNA damage as driver of cellular outcomes. *Am J Physiol Cell Physiol.* 2022;322(2):C136–C150. <https://doi.org/10.1152/ajpcell.00389.2021>.
96. Moeder KP, Weber AW, Bamford FJ, et al. Matrilineal affinities and prehistoric Siberian mortuary practices: a case study from Neolithic Lake Baikal. *J Archaeol Sci.* 2005;32(4):619–634. <https://doi.org/10.1016/j.jas.2004.12.002>.
97. Esparza Á, Palomo-Díez S, Velasco-Vázquez J, Delibes G, Arroyo-Pardo E, Salazar-García DC. Familiar kinship? Palaeogenetic and isotopic evidence from a triple burial of the Cogotas I archaeological culture (bronze age, Iberian Peninsula). *Oxford J Archaeology.* 2017;36(3):223–242. <https://doi.org/10.1111/oja.12113>.
98. Parson W, Gusmão L, Hares DR, et al. DNA Commission of the International Society for forensic genetics: revised and extended guidelines for mitochondrial DNA typing. *Forensic Sci Int Genet.* 2014;13:134–142. <https://doi.org/10.1016/j.fsigen.2014.07.010>.
99. Peck MA, Brandhagen MD, Marshall C, Diegoli TM, Irwin JA, Sturk-Andreaggi K. Concordance and reproducibility of a next generation mtGenome sequencing method for high-quality samples using the Illumina MiSeq. *Forensic Sci Int Genet.* 2016;24:103–111. <https://doi.org/10.1016/j.fsigen.2016.06.003>.
100. Rohland N, Mallick S, Mah M, Maier R, Patterson N, Reich D. Three assays for in-solution enrichment of ancient human DNA at more than a million SNPs. *Genome Res.* 2022;32(11–12):2068–2078. <https://doi.org/10.1101/gr.276728.122>.
101. Dabney J, Knapp M, Glocke I, et al. Complete mitochondrial genome sequence of a middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc Natl Acad Sci USA.* 2013;110(39):15758–15763. <https://doi.org/10.1073/pnas.1314445110>.
102. Rohland N, Glocke I, Aximu-Petri A, Meyer M. Extraction of highly degraded DNA from ancient bones, teeth and sediments for high-throughput sequencing. *Nat Protoc.* 2018;13(11):2447–2461. <https://doi.org/10.1038/s41596-018-0050-5>.
103. Rohland N, Hofreiter M. Ancient DNA extraction from bones and teeth. *Nat Protoc.* 2007;2(7):1756–1762. <https://doi.org/10.1038/nprot.2007.247>.
104. Kapp JD, Green RE, Shapiro B. A fast and efficient single-stranded genomic library preparation method optimized for ancient DNA. Fleischer RC, ed. *J Hered.* 2021;112(3):241–249. <https://doi.org/10.1093/jhered/esab012>.
105. Weissensteiner H, Forer L, Fuchsberger C, et al. mtDNA-server: next-generation sequencing data analysis of human mitochondrial DNA in the cloud. *Nucleic Acids Res.* 2016;44(W1):W64–W69. <https://doi.org/10.1093/nar/gkw247>.
106. Cihlar JC, Strobl C, Lagacé R, Muenzler M, Parson W, Budowle B. Distinguishing mitochondrial DNA and NUMT sequences amplified with the precision ID mtDNA whole genome panel. *Mitochondrion.* 2020;55:122–133. <https://doi.org/10.1016/j.mito.2020.09.001>.
107. Picardi E, Pesole G. Mitochondrial genomes gleaned from human whole-exome sequencing. *Nat Methods.* 2012;9(6):523–524. <https://doi.org/10.1038/nmeth.2029>.
108. Calabrese C, Simone D, Diroma MA, et al. MToolBox: a highly automated pipeline for heteroplasmy annotation and prioritization analysis of human mitochondrial variants in high-throughput sequencing. *Bioinformatics.* 2014;30(21):3115–3117. <https://doi.org/10.1093/bioinformatics/btu483>.
109. Peltzer A, Jäger G, Herbig A, et al. EAGER: efficient ancient genome reconstruction. *Genome Biol.* 2016;17(1):60. <https://doi.org/10.1186/s13059-016-0918-z>.
110. Santibanez-Koref M, Griffin H, Turnbull DM, Chinnery PF, Herbert M, Hudson G. Assessing mitochondrial heteroplasmy using next generation sequencing: A note of caution. *Mitochondrion.* 2019;46:302–306. <https://doi.org/10.1016/j.mito.2018.08.003>.
111. Bendall KE, Sykes BC. Length heteroplasmy in the first hypervariable segment of the human mtDNA control region. *Am J Hum Genet.* 1995;57(2):248–256. PMID: 7668250; PMCID: PMC1801530.
112. Stewart J, Fisher C, Aagaard P, et al. Length variation in HV2 of the human Mitochondrial DNA control region. *J Forensic Sci.* 2001;46(4):862–870. <https://doi.org/10.1520/JFS15059J>.
113. Orlando L, Allaby R, Skoglund P, et al. Ancient DNA analysis. *Nat Rev Methods Primers.* 2021;1(1), 14. <https://doi.org/10.1038/s43586-020-00011-0>.
114. Garrido Marques A, Rubinacci S, Malaspina AS, Delaneau O, Sousa Da Mota B. Assessing the impact of post-mortem damage and contamination on imputation performance in ancient DNA. *Sci Rep.* 2024;14(1), 6227. <https://doi.org/10.1038/s41598-024-56584-3>.
115. Templeton JEL, Brotherton PM, Llamas B, et al. DNA capture and next-generation sequencing can recover whole mitochondrial genomes from highly degraded samples for human identification. *Investig Genet.* 2013;4(1), 26. <https://doi.org/10.1186/2041-2223-4-26>.
116. Renaud G, Slon V, Duggan AT, Kelso J, Schmutzki: estimation of contamination and endogenous mitochondrial consensus calling for ancient DNA. *Genome Biol.* 2015;16(1):224. <https://doi.org/10.1186/s13059-015-0776-0>.
117. Chyleński M, Ehler E, Somel M, et al. Ancient mitochondrial genomes reveal the absence of maternal kinship in the burials of Çatalhöyük people and their genetic affinities. *Genes.* 2019;10(3), 207. <https://doi.org/10.3390/genes10030207>.
118. Kircher M, Sawyer S, Meyer M. Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. *Nucleic Acids Res.* 2012;40(1):e3. <https://doi.org/10.1093/nar/gkr771>.
119. Neukamm J, Peltzer A, Nieselt K. DamageProfiler: fast damage pattern calculation for ancient DNA. Kelso J, ed. *Bioinformatics.* 2021;37(20):3652–3653. <https://doi.org/10.1093/bioinformatics/btab190>.
120. Ginolhac A, Rasmussen M, Gilbert MTP, Willerslev E, Orlando L. mapDamage: testing for damage patterns in ancient DNA sequences. *Bioinformatics.* 2011;27(15):2153–2155. <https://doi.org/10.1093/bioinformatics/btr347>.
121. Jönsson H, Ginolhac A, Schubert M, Johnson PLF, Orlando L. mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics.* 2013;29(13):1682–1684. <https://doi.org/10.1093/bioinformatics/btt193>.
122. Nakatsuka N, Harney É, Mallick S, Mah M, Patterson N, Reich D. ContamLD: estimation of ancient nuclear DNA contamination using breakdown of linkage disequilibrium. *Genome Biol.* 2020;21(1):199. <https://doi.org/10.1186/s13059-020-02111-2>.
123. Peyrégne S, Peter BM. AuthenticT: a model of ancient DNA damage to estimate the proportion of present-day DNA contamination. *Genome Biol.* 2020;21(1):246. <https://doi.org/10.1186/s13059-020-02123-y>.

124. Korneliusson TS, Albrechtsen A, Nielsen R. ANGSD: analysis of next generation sequencing data. *BMC Bioinformatics*. 2014;15(1):356. <https://doi.org/10.1186/s12859-014-0356-4>.
125. Heraclides A, Fernández-Domínguez E. Mitochondrial DNA consensus calling and quality filtering for constructing ancient human mitogenomes: comparison of two widely applied methods. *Int J Mol Sci*. 2022;23(9), 4651. <https://doi.org/10.3390/ijms23094651>.
126. Kreissl Lonfat BM, Kaufmann IM, Rühli F. A code of ethics for evidence-based research with ancient human remains. *Anat Rec*. 2015;298(6):1175–1181. <https://doi.org/10.1002/ar.23126>.
127. Wagner JK, Colwell C, Claw KG, et al. Fostering responsible research on ancient DNA. *Am J Hum Genet*. 2020;107(2):183–195. <https://doi.org/10.1016/j.ajhg.2020.06.017>.
128. Alpaslan-Roodenberg S, Anthony D, Babiker H, et al. Ethics of DNA research on human remains: five globally applicable guidelines. *Nature*. 2021;599(7883): 41–46. <https://doi.org/10.1038/s41586-021-04008-x>.
129. Gibbon VE, Thompson JC, Alves S. Informed proxy consent for ancient DNA research. *Commun Biol*. 2024;7(1):815. <https://doi.org/10.1038/s42003-024-06413-0>.
130. Robbins Schug G, Halcrow SE, De La Cova C. They are people too: the ethics of curation and use of human skeletal remains for teaching and research. *American Journal of Biological Anthropology*. 2025;186(2), e70013. <https://doi.org/10.1002/ajpa.70013>.
131. Nilsson Stutz L, Peyroteo Stjerna R, Tarlow S. Ethical entanglements: human remains, museums and ethics in a European perspective. *Antiquity*. 2025;99(403), e8. <https://doi.org/10.15184/aqy.2024.177>.